



**HAL**  
open science

# Technical Design Report for the Phase-II Upgrade of the ATLAS TDAQ System

N. Berger, A.M. Burger, O. Dartsis, M. Delmastro, L. Di Ciaccio, P.J. Falke, S. Falke, C. Goy, T. Guillemin, T. Hryn'ova, et al.

► **To cite this version:**

N. Berger, A.M. Burger, O. Dartsis, M. Delmastro, L. Di Ciaccio, et al.. Technical Design Report for the Phase-II Upgrade of the ATLAS TDAQ System. CERN-LHCC-2017-020, CERN. 2018. hal-03990850

**HAL Id: hal-03990850**

**<https://hal.science/hal-03990850>**

Submitted on 15 Feb 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



**ATLAS**  
EXPERIMENT

ATL-TDR-029 · LHCC-2017-020

0100101101110101 0100101101110101  
0100101101110101 0100101101110101

# ATLAS

TDAQ Phase-II Upgrade



**Technical Design Report**







# Technical Design Report for the Phase-II Upgrade of the ATLAS Trigger and Data Acquisition System

---

The ATLAS Collaboration

Reference: CERN-LHCC-2017-020  
ATLAS-TDR-029  
Created: 15 June 2018  
Last modified: 15 June 2018  
**Prepared by: The ATLAS Collaboration**

© 2018 CERN for the benefit of the ATLAS Collaboration.  
Reproduction of this article or parts of it is allowed as specified in the CC-BY-4.0 license.





## **Abstract**

This Technical Design Report documents the plans to upgrade the ATLAS Trigger and Data Acquisition system for the High Luminosity LHC (HL-LHC). The HL-LHC is expected to start operations in the middle of 2026, to ultimately reach a peak instantaneous luminosity of  $\mathcal{L} = 7.5 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ , corresponding to approximately 200 inelastic proton-proton collisions per bunch crossing, and to deliver more than ten times the integrated luminosity of the LHC Runs 1-3 combined (up to  $4000 \text{ fb}^{-1}$ ).

Meeting these requirements poses significant challenges to the Trigger and to the Data Acquisition system to fully exploit the physics potential of the HL-LHC. A baseline architecture, based on a single-level hardware trigger with a maximum rate of 1 MHz and  $10 \mu\text{s}$  latency, is proposed and documented. With the help of a hardware-based tracking sub-system as co-processor, software-based reconstruction follows to achieve further rejection. Up to 10 kHz event data are sent into storage.

The Report describes in detail the physics motivations, the requirements, the fundamental parameters, the technical design implementation, and the expected performance of the proposed upgrade. The Report also documents the organisation of the Upgrade Project, its management structure, planning and scheduling with a review of the major milestones, and costing information.



---

## ATLAS Institutions

### **Argentina**

Departamento de Física, Universidad de Buenos Aires, Buenos Aires  
Instituto de Física La Plata, Universidad Nacional de La Plata and CONICET, La Plata

### **Armenia**

Yerevan Physics Institute, Yerevan

### **Australia**

Department of Physics, University of Adelaide, Adelaide  
School of Physics, University of Sydney, Sydney  
School of Physics, University of Melbourne, Victoria

### **Austria**

Institut für Astro- und Teilchenphysik, Leopold-Franzens-Universität, Innsbruck  
Fachhochschule Wiener Neustadt, Wiener Neustadt

### **Azerbaijan**

Institute of Physics, Azerbaijan Academy of Sciences, Baku

### **Belarus**

B.I. Stepanov Institute of Physics, National Academy of Sciences of Belarus, Minsk  
Research Institute for Nuclear Problems of Byelorussian State University, Minsk

### **Brazil**

*Brazil Cluster:* Departamento de Engenharia Elétrica, Universidade Federal de Juiz de Fora (UFJF), Juiz de Fora; Universidade Federal do Rio De Janeiro COPPE/EE/IF, Rio de Janeiro; Universidade Federal de São João del Rei (UFSJ), São João del Rei; Instituto de Física, Universidade de São Paulo, São Paulo

### **Canada**

Department of Physics, Simon Fraser University, Burnaby BC  
Department of Physics, University of Alberta, Edmonton AB  
Department of Physics, McGill University, Montreal QC  
Group of Particle Physics, University of Montreal, Montreal QC  
Department of Physics, Carleton University, Ottawa ON  
Department of Physics, University of Toronto, Toronto ON  
Department of Physics, University of British Columbia, Vancouver BC  
TRIUMF, Vancouver BC; Department of Physics and Astronomy, York University, Toronto ON  
Department of Physics and Astronomy, University of Victoria, Victoria BC

### **CERN**



---

European Organization for Nuclear Research, Geneva, Switzerland

### **Chile**

*Chile Cluster:* Departamento de Física, Pontificia Universidad Católica de Chile, Santiago;  
Departamento de Física, Universidad Técnica Federico Santa María, Valparaíso

### **China**

*China IHEP-NJU-THU Cluster:* Institute of High Energy Physics, Chinese Academy of Sciences, Beijing; Physics Department, Tsinghua University, Beijing; Department of Physics, Nanjing University, Nanjing

*China USTC-SDU-SJTU Cluster:* Department of Modern Physics and State Key Laboratory of Particle Detection and Electronics, University of Science and Technology of China, Hefei; Institute of Frontier and Interdisciplinary Science and Key Laboratory of Particle Physics and Particle Irradiation (MOE), Shandong University, Qingdao; School of Physics and Astronomy, Shanghai Jiao Tong University, KLPPAC-MoE, SKLPPC, Shanghai; Tsung-Dao Lee Institute, Shanghai

*Hong Kong Cluster:* Department of Physics, Chinese University of Hong Kong, Shatin, N.T., Hong Kong; Department of Physics, University of Hong Kong, Hong Kong; Department of Physics and Institute for Advanced Study, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong

### **Colombia**

Centro de Investigaciones, Universidad Antonio Nariño, Bogota

### **Czech Republic**

Palacký University, RCPTM, Joint Laboratory of Optics, Olomouc

Charles University, Faculty of Mathematics and Physics, Prague

Czech Technical University in Prague, Prague

Institute of Physics, Academy of Sciences of the Czech Republic, Prague

### **Denmark**

Niels Bohr Institute, University of Copenhagen, Copenhagen

### **France**

LAPP, Université Grenoble Alpes, Université Savoie Mont Blanc, CNRS/IN2P3, Annecy

LPC, Université Clermont Auvergne, CNRS/IN2P3, Clermont-Ferrand

IRFU, CEA, Université Paris-Saclay, Gif-sur-Yvette

LPSC, Université Grenoble Alpes, CNRS/IN2P3, Grenoble INP, Grenoble

CPPM, Aix-Marseille Université, CNRS/IN2P3, Marseille

LAL, Université Paris-Sud, CNRS/IN2P3, Université Paris-Saclay, Orsay

LPNHE, Sorbonne Université, Paris Diderot Sorbonne Paris Cité, CNRS/IN2P3, Paris

### **Georgia**

*Georgia Cluster:* E. Andronikashvili Institute of Physics, Iv. Javakhishvili Tbilisi State University, Tbilisi; High Energy Physics Institute, Tbilisi State University, Tbilisi

---

## Germany

Institut für Physik, Humboldt Universität zu Berlin, Berlin  
Physikalisches Institut, Universität Bonn, Bonn  
Lehrstuhl für Experimentelle Physik IV, Technische Universität Dortmund, Dortmund  
Institut für Kern- und Teilchenphysik, Technische Universität Dresden, Dresden  
Physikalisches Institut, Albert-Ludwigs-Universität Freiburg, Freiburg  
II. Physikalisches Institut, Justus-Liebig-Universität Giessen, Giessen  
II. Physikalisches Institut, Georg-August-Universität Göttingen, Göttingen  
Deutsches Elektronen-Synchrotron DESY, Hamburg and Zeuthen  
Kirchhoff-Institut für Physik, Ruprecht-Karls-Universität Heidelberg, Heidelberg;  
Physikalisches Institut, Ruprecht-Karls-Universität Heidelberg, Heidelberg  
Institut für Physik, Universität Mainz, Mainz  
Fakultät für Physik, Ludwig-Maximilians-Universität München, München  
Max-Planck-Institut für Physik (Werner-Heisenberg-Institut), München  
Department Physik, Universität Siegen, Siegen  
Fakultät für Mathematik und Naturwissenschaften, Fachgruppe Physik, Bergische  
Universität Wuppertal, Wuppertal  
Fakultät für Physik und Astronomie, Julius-Maximilians-Universität Würzburg,  
Würzburg

## Greece

National Centre for Scientific Research "Demokritos", Agia Paraskevi  
Physics Department, National and Kapodistrian University of Athens, Athens  
Department of Physics, Aristotle University of Thessaloniki, Thessaloniki  
Physics Department, National Technical University of Athens, Zografou

## Israel

Department of Physics, Technion, Israel Institute of Technology, Haifa  
Department of Particle Physics, Weizmann Institute of Science, Rehovot  
Raymond and Beverly Sackler School of Physics and Astronomy, Tel Aviv University, Tel  
Aviv

## Italy

INFN Sezione di Bologna; Dipartimento di Fisica e Astronomia, Università di Bologna,  
Bologna  
INFN e Laboratori Nazionali di Frascati, Frascati  
INFN Sezione di Genova; Dipartimento di Fisica, Università di Genova, Genova  
INFN Sezione di Lecce; Dipartimento di Matematica e Fisica, Università del Salento, Lecce  
INFN Sezione di Milano; Dipartimento di Fisica, Università di Milano, Milano  
INFN Sezione di Napoli; Dipartimento di Fisica, Università di Napoli, Napoli  
INFN Sezione di Pavia; Dipartimento di Fisica, Università di Pavia, Pavia  
INFN Sezione di Pisa; Dipartimento di Fisica E. Fermi, Università di Pisa, Pisa

---

INFN Gruppo Collegato di Cosenza, Laboratori Nazionali di Frascati; Dipartimento di Fisica, Università della Calabria, Rende  
INFN Sezione di Roma; Dipartimento di Fisica, Sapienza Università di Roma, Roma  
INFN Sezione di Roma Tor Vergata; Dipartimento di Fisica, Università di Roma Tor Vergata, Roma  
INFN Sezione di Roma Tre; Dipartimento di Matematica e Fisica, Università Roma Tre, Roma  
INFN-TIFPA; Università degli Studi di Trento, Trento  
INFN Gruppo Collegato di Udine, Sezione di Trieste, Udine; ICTP, Trieste; Dipartimento di Chimica, Fisica e Ambiente, Università di Udine, Udine

### **Japan**

Research Center for Advanced Particle Physics and Department of Physics, Kyushu University, Fukuoka  
Faculty of Applied Information Science, Hiroshima Institute of Technology, Hiroshima  
Graduate School of Science, Kobe University, Kobe  
Faculty of Science, Kyoto University, Kyoto  
Kyoto University of Education, Kyoto  
Department of Physics, Shinshu University, Nagano  
Nagasaki Institute of Applied Science, Nagasaki  
Graduate School of Science and Kobayashi-Maskawa Institute, Nagoya University, Nagoya  
Faculty of Science, Okayama University, Okayama  
Graduate School of Science, Osaka University, Osaka  
Department of Physics, Tokyo Institute of Technology, Tokyo  
Graduate School of Science and Technology, Tokyo Metropolitan University, Tokyo  
International Center for Elementary Particle Physics and Department of Physics, University of Tokyo, Tokyo  
Waseda University, Tokyo  
Division of Physics and Tomonaga Center for the History of the Universe, Faculty of Pure and Applied Sciences, University of Tsukuba, Tsukuba  
KEK, High Energy Accelerator Research Organization, Tsukuba

### **Morocco**

*Morocco Cluster:* Centre National de l’Energie des Sciences Techniques Nucleaires (CNESTEN), Rabat; Faculté des Sciences Ain Chock, Réseau Universitaire de Physique des Hautes Energies - Université Hassan II, Casablanca; Faculté des Sciences Semlalia, Université Cadi Ayyad, LPHEA-Marrakech; Faculté des Sciences, Université Mohamed Premier and LPTPM, Oujda; Faculté des sciences, Université Mohammed V, Rabat

### **Netherlands**

Nikhef National Institute for Subatomic Physics and University of Amsterdam, Amsterdam



---

Institute for Mathematics, Astrophysics and Particle Physics, Radboud University  
Nijmegen/Nikhef, Nijmegen

**Norway**

Department for Physics and Technology, University of Bergen, Bergen  
Department of Physics, University of Oslo, Oslo

**Poland**

Institute of Nuclear Physics Polish Academy of Sciences, Krakow  
AGH University of Science and Technology, Faculty of Physics and Applied Computer  
Science, Krakow; Marian Smoluchowski Institute of Physics, Jagiellonian University,  
Krakow

**Portugal**

*Portugal Cluster:* Laboratório de Instrumentação e Física Experimental de Partículas - LIP;  
Departamento de Física, Faculdade de Ciências, Universidade de Lisboa, Lisboa;  
Departamento de Física, Universidade de Coimbra, Coimbra; Departamento de Física,  
Universidade do Minho, Braga; Departamento de Física Teórica y del Cosmos,  
Universidad de Granada, Granada (Spain); Dep Física and CEFITEC of Faculdade de  
Ciências e Tecnologia, Universidade Nova de Lisboa, Caparica

**Romania**

*Romania Cluster:* Transilvania University of Brasov, Brasov; Horia Hulubei National  
Institute of Physics and Nuclear Engineering, Bucharest; National Institute for Research  
and Development of Isotopic and Molecular Technologies, Physics Department,  
Cluj-Napoca; Department of Physics, Alexandru Ioan Cuza University of Iasi, Iasi;  
University Politehnica Bucharest, Bucharest; West University in Timisoara, Timisoara

**Russia**

D.V. Skobeltsyn Institute of Nuclear Physics, M.V. Lomonosov Moscow State University,  
Moscow  
Institute for Theoretical and Experimental Physics (ITEP), Moscow  
National Research Nuclear University MEPhI, Moscow  
P.N. Lebedev Physical Institute of the Russian Academy of Sciences, Moscow  
Novosibirsk State University Novosibirsk; Budker Institute of Nuclear Physics, SB RAS,  
Novosibirsk  
State Research Center Institute for High Energy Physics, NRC KI, Protvino  
Konstantinov Nuclear Physics Institute of National Research Centre "Kurchatov Institute",  
PNPI, St. Petersburg  
Tomsk State University, Tomsk

**JINR**

Joint Institute for Nuclear Research, Dubna, Russia

**Serbia**

---

Institute of Physics, University of Belgrade, Belgrade

**Slovak Republic**

*Slovak Republic Cluster:* Faculty of Mathematics, Physics and Informatics, Comenius University, Bratislava; Department of Subnuclear Physics, Institute of Experimental Physics of the Slovak Academy of Sciences, Kosice

**Slovenia**

Department of Experimental Particle Physics, Jožef Stefan Institute and Department of Physics, University of Ljubljana, Ljubljana

**South Africa**

*South Africa Cluster:* Department of Physics, University of Cape Town, Cape Town; Department of Mechanical Engineering Science, University of Johannesburg, Johannesburg; School of Physics, University of the Witwatersrand, Johannesburg

**Spain**

Institut de Física d'Altes Energies (IFAE), Barcelona Institute of Science and Technology, Barcelona  
Departamento de Física Teórica C-15 and CIAFF, Universidad Autónoma de Madrid, Madrid  
Instituto de Física Corpuscular (IFIC), Centro Mixto Universidad de Valencia - CSIC, Valencia

**Sweden**

Fysiska institutionen, Lunds universitet, Lund  
Department of Physics, Stockholm University; Oskar Klein Centre, Stockholm  
Physics Department, Royal Institute of Technology, Stockholm  
Department of Physics and Astronomy, University of Uppsala, Uppsala

**Switzerland**

Albert Einstein Center for Fundamental Physics and Laboratory for High Energy Physics, University of Bern, Bern  
Département de Physique Nucléaire et Corpusculaire, Université de Genève, Genève

**Taiwan**

Department of Physics, National Tsing Hua University, Hsinchu  
Institute of Physics, Academia Sinica, Taipei

**Turkey**

*Ankara Cluster:* Department of Physics, Ankara University, Ankara; Istanbul Aydin University, Istanbul; Division of Physics, TOBB University of Economics and Technology, Ankara  
*Bogazici Cluster:* Bahcesehir University, Faculty of Engineering and Natural Sciences, Istanbul; Istanbul Bilgi University, Faculty of Engineering and Natural Sciences, Istanbul;

---

Department of Physics, Bogazici University, Istanbul; Department of Physics Engineering,  
Gaziantep University, Gaziantep

### **United Kingdom**

School of Physics and Astronomy, University of Birmingham, Birmingham  
Department of Physics and Astronomy, University of Sussex, Brighton  
Cavendish Laboratory, University of Cambridge, Cambridge  
Department of Physics, University of Warwick, Coventry  
Particle Physics Department, Rutherford Appleton Laboratory, Didcot  
SUPA - School of Physics and Astronomy, University of Edinburgh, Edinburgh  
Department of Physics, Royal Holloway University of London, Egham  
SUPA - School of Physics and Astronomy, University of Glasgow, Glasgow  
Physics Department, Lancaster University, Lancaster  
Oliver Lodge Laboratory, University of Liverpool, Liverpool  
Department of Physics and Astronomy, University College London, London  
School of Physics and Astronomy, Queen Mary University of London, London  
School of Physics and Astronomy, University of Manchester, Manchester  
Department of Physics, Oxford University, Oxford  
Department of Physics and Astronomy, University of Sheffield, Sheffield

### **United States of America**

Physics Department, SUNY Albany, Albany NY  
Department of Physics and Astronomy, University of New Mexico, Albuquerque NM  
Department of Physics and Astronomy, Iowa State University, Ames IA  
Department of Physics, University of Massachusetts, Amherst MA  
Department of Physics, University of Michigan, Ann Arbor MI  
High Energy Physics Division, Argonne National Laboratory, Argonne IL  
Department of Physics, University of Texas at Arlington, Arlington TX  
Department of Physics, University of Texas at Austin, Austin TX  
Physics Division, Lawrence Berkeley National Laboratory and University of California,  
Berkeley CA  
Department of Physics, Indiana University, Bloomington IN  
Department of Physics, Boston University, Boston MA  
Laboratory for Particle Physics and Cosmology, Harvard University, Cambridge MA  
Enrico Fermi Institute, University of Chicago, Chicago IL  
Ohio State University, Columbus OH  
Physics Department, Southern Methodist University, Dallas TX  
Department of Physics, Northern Illinois University, DeKalb IL  
Department of Physics, Duke University, Durham NC  
Department of Physics and Astronomy, Michigan State University, East Lansing MI  
Center for High Energy Physics, University of Oregon, Eugene OR  
University of Iowa, Iowa City IA  
Department of Physics and Astronomy, University of California Irvine, Irvine CA

---

Nevis Laboratory, Columbia University, Irvington NY  
Department of Physics, University of Wisconsin, Madison WI  
Department of Physics and Astronomy, Tufts University, Medford MA  
Department of Physics, Yale University, New Haven CT  
Department of Physics, New York University, New York NY  
Homer L. Dodge Department of Physics and Astronomy, University of Oklahoma,  
Norman OK  
Department of Physics, University of Pennsylvania, Philadelphia PA  
Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh PA  
Physics Department, University of Texas at Dallas, Richardson TX  
Louisiana Tech University, Ruston LA  
Santa Cruz Institute for Particle Physics, University of California Santa Cruz, Santa Cruz  
CA  
Department of Physics, University of Washington, Seattle WA  
SLAC National Accelerator Laboratory, Stanford CA  
Department of Physics, Oklahoma State University, Stillwater OK  
Departments of Physics and Astronomy, Stony Brook University, Stony Brook NY  
Department of Physics, University of Arizona, Tucson AZ  
Physics Department, Brookhaven National Laboratory, Upton NY  
Department of Physics, University of Illinois, Urbana IL  
Department of Physics, Brandeis University, Waltham MA

# Contents

<b>Executive Summary</b>	<b>1</b>
<b>I General Overview of the TDAQ Phase-II Upgrade Project</b>	<b>3</b>
<b>1 Introduction</b>	<b>5</b>
1.1 LHC Roadmap	6
1.2 Physics Drivers for the HL-LHC Upgrades	9
1.3 ATLAS Upgrade Strategies for the HL-LHC	11
1.4 Baseline Architecture and Structure of the Upgrade Project	13
1.5 Outline of this Report	16
<b>2 Physics Motivation</b>	<b>19</b>
2.1 Physics Signatures with Single-Electron and Single-Muon Triggers	20
2.2 Physics Signatures with Two Leptons	24
2.3 Physics Signatures with Photons	24
2.4 Physics Signatures with Hadronically Decaying Tau Leptons	26
2.5 Physics Signatures with Jets	27
2.6 Physics Signatures with Missing Transverse Energy	30
2.7 Physics Signatures with Forward Electrons	31
2.8 Physics Signatures with Exotic Objects	32
2.9 Physics with an Inclusive Vector Boson Fusion (VBF) Trigger	33
2.10 <i>B</i> -Physics Signatures	34
2.11 Physics Signatures for Heavy Ion Collision	35
2.12 Summary of Requirements and Motivation for the Upgrade	37
<b>3 Challenges and Limitations of the Run 3 TDAQ system</b>	<b>39</b>
3.1 Overview of the Phase-I Trigger and Readout Architecture	40
3.2 Features and Limitations of the Run 3 Level-1 Trigger System	44
3.2.1 Level-1 Calorimeter Trigger Limitations	44
3.2.2 Level-1 Muon Trigger Limitations	53
3.2.3 Level-1 Topological Trigger Limitations	58
3.2.4 Level-1 Trigger Menu up to Run 3	59
3.2.5 Limitations in the Level-1 CTP and TTC systems	62
3.3 Limitations of the Run 3 DAQ System	63
3.3.1 Detector Readout	63

3.3.2	Dataflow, Storage and Networking . . . . .	65
3.4	Limitations of the Run 3 HLT System . . . . .	68
3.4.1	Estimates of Rejection Factors in the Run 3 HLT . . . . .	68
3.4.2	Fast Tracking . . . . .	68
3.4.3	Features and Limitations of Event Selection in the Run 3 HLT . . . . .	70
<b>4</b>	<b>Architectural and Functional Requirements</b>	<b>75</b>
4.1	Constraints from the Detectors . . . . .	75
4.1.1	New Inner Tracker . . . . .	75
4.1.2	Calorimeter Detectors . . . . .	77
4.1.3	Muon System . . . . .	77
4.1.4	Other Detectors . . . . .	79
4.2	Level-0 Trigger Requirements . . . . .	79
4.3	DAQ Requirements . . . . .	83
4.4	Event Filter Requirements . . . . .	84
<b>5</b>	<b>Description of the Baseline System</b>	<b>87</b>
5.1	Functional overview . . . . .	87
5.2	The Level-0 Trigger System . . . . .	90
5.2.1	Level-0 Calorimeter Trigger . . . . .	90
5.2.2	Level-0 Muon Trigger . . . . .	92
5.2.3	MUCTPI . . . . .	94
5.2.4	Global Trigger . . . . .	94
5.2.5	Central Trigger Processor . . . . .	95
5.2.6	Trigger, Timing and Control System . . . . .	96
5.2.7	Technical implementation and system size . . . . .	98
5.2.8	Level-0 Trigger Latency . . . . .	100
5.3	Data Acquisition System . . . . .	101
5.3.1	Readout . . . . .	101
5.3.2	Dataflow . . . . .	106
5.3.3	Network . . . . .	109
5.4	Event Filter System . . . . .	110
5.4.1	Use of Tracking in the Event Filter . . . . .	111
5.4.2	Event Filter System Overview . . . . .	113
5.4.3	Event Filter Farm Hardware . . . . .	114
5.4.4	Event Filter Software . . . . .	116
5.4.5	Hardware-based Tracking for the Trigger Subsystem . . . . .	117
<b>6</b>	<b>Expected Performance</b>	<b>123</b>
6.1	Performance Estimation Procedures . . . . .	123
6.2	Topological Clusters . . . . .	124
6.3	Electrons and Photons . . . . .	124
6.4	Muons . . . . .	128



6.5	Tau Leptons	132
6.6	Jets	135
6.7	Missing Transverse Energy	141
6.8	Boosted-object Triggers	145
6.9	Exotic-object Triggers	147
6.10	Inclusive Vector Boson Fusion Trigger	148
6.11	Example Run 4 Trigger Menu	149
6.12	Tracking Needs	151
6.13	Impact of the System on $HH \rightarrow 4b$	156
6.14	System Flexibility and Dataset Quality	158
<b>II</b>	<b>Detailed Description of System Components</b>	<b>161</b>
<b>7</b>	<b>Level-0 Calorimeter Trigger</b>	<b>163</b>
7.1	Evolution of the Hardware Calorimeter Trigger	163
7.1.1	Enhancements and Interface changes of the Phase-I Level-1 Calorimeter Trigger	163
7.1.2	Requirements for the Level-0 Calorimeter Trigger	164
7.2	Performance of the Level-0 Calorimeter Trigger	164
7.3	Architecture and Hardware Realisation	165
7.3.1	Overview	165
7.3.2	Input Signals	168
7.3.3	Processing System	170
7.3.4	The forward FEX System (fFEX)	171
7.3.5	Output Signals	173
7.3.6	Latency	174
7.3.7	Readout and Monitoring	174
7.4	Firmware	175
7.4.1	Algorithmic Firmware	175
7.4.2	Infrastructure Firmware	175
7.5	R & D Programme	175
7.6	Commissioning	176
<b>8</b>	<b>Level-0 Muon Trigger</b>	<b>177</b>
8.1	Introduction	177
8.2	Overview of the Current System and the Limitations	177
8.3	Overview of the Upgrade	179
8.3.1	Extension of the Latency and the Rate	180
8.3.2	Improvement of the Trigger Performance	181
8.4	Sector Logic	182
8.4.1	Trigger Scheme and Performance	183
8.4.2	Hardware Design	196

8.4.3	Firmware Design . . . . .	198
8.4.4	Detector Signal Inputs . . . . .	199
8.4.5	Realtime Output Data Format . . . . .	200
8.5	NSW Trigger Processor . . . . .	202
8.5.1	Motivation for the Upgrade . . . . .	202
8.5.2	Performance of the Segment Reconstruction . . . . .	204
8.5.3	Hardware Design . . . . .	204
8.5.4	Firmware Design . . . . .	205
8.5.5	Detector Signal Inputs . . . . .	206
8.5.6	Realtime Output Data Format . . . . .	207
8.6	MDT Trigger Processor . . . . .	207
8.6.1	Motivation for the Implementation . . . . .	207
8.6.2	Trigger Performance . . . . .	208
8.6.3	Hardware Overview . . . . .	212
8.6.4	Hit Extraction . . . . .	213
8.6.5	Segment Reconstruction . . . . .	216
8.6.6	Transverse Momentum Evaluation . . . . .	223
8.6.7	Realtime Output Data Format . . . . .	225
8.6.8	Readout Path . . . . .	225
8.6.9	Resource Estimates . . . . .	226
8.7	Latency Estimates . . . . .	226
8.8	R&D Items . . . . .	229
<b>9</b>	<b>Global Trigger</b>	<b>235</b>
9.1	Introduction . . . . .	235
9.1.1	Requirements for the Global Trigger System . . . . .	236
9.1.2	Overview of Global Trigger System Design . . . . .	237
9.2	Global Trigger Interfaces . . . . .	239
9.2.1	Calorimeter Subdetector Inputs . . . . .	240
9.2.2	L0Calo Inputs . . . . .	240
9.2.3	L0Muon Inputs . . . . .	240
9.2.4	Central Trigger Processor Outputs . . . . .	240
9.2.5	Front-End LInk eXchange (FELIX) Inputs and Outputs . . . . .	241
9.2.6	Configuration and Control, Interface to Detector Control System . . . . .	241
9.3	Trigger Strategy and Algorithms . . . . .	241
9.3.1	Algorithm Performance . . . . .	242
9.3.2	Algorithm Implementation . . . . .	242
9.3.3	Topological Clustering . . . . .	243
9.3.4	Anti- $k_t$ Jet Clustering . . . . .	249
9.4	Physical Realisation . . . . .	255
9.4.1	Architecture . . . . .	255
9.4.2	Required Number of Global Common Modules (GCMs) . . . . .	255
9.4.3	Required Number of Links on each GCM . . . . .	256

---

9.4.4	GCM Requirements and Preliminary Design	258
9.4.5	Fibre Plant	259
9.4.6	Production Firmware Deployment Module	260
9.5	Latency	261
9.5.1	Data arrival time at Multiplexer Processor	262
9.5.2	Processing latency and budget	262
9.6	Firmware	263
9.6.1	Firmware Blocks	264
9.6.2	FPGA Resource Usage	265
9.6.3	Firmware Development	266
<b>10</b>	<b>Central Trigger System</b>	<b>269</b>
10.1	Level-0 CTP	271
10.1.1	CTPMI: Machine Interface	272
10.1.2	CTPIN: Trigger Inputs	273
10.1.3	CTPCORE	274
10.2	Distribution of Timing, Trigger, and Control Signals (TTC)	278
10.2.1	Local Trigger Interface	279
10.3	Level-0 MUCTPI	281
10.4	Monitoring, Control, Configuration, Offline Software	283
10.5	Latency	284
10.6	Project Planning	285
<b>11</b>	<b>Data Acquisition</b>	<b>289</b>
11.1	Introduction	289
11.2	Overview of System Functionality	290
11.3	Event Data Format	290
11.4	Detector Readout	291
11.4.1	FELIX	291
11.4.2	Data Handler	294
11.4.3	R&D leading to the final design and implementation of the Readout System	296
11.5	Dataflow	298
11.5.1	Event Builder	299
11.5.2	Storage Handler	300
11.5.3	Event Aggregator	300
11.5.4	Storage System Sizing & Technology	300
11.5.5	Factors Affecting Implementation	301
11.5.6	System Workflow	302
11.6	Network	306
11.6.1	Network Domain Description	306
11.6.2	Network Technology Choice	307
11.6.3	Network Architecture	308

11.6.4	Control Network . . . . .	309
11.6.5	Network Installation . . . . .	310
11.7	Online Software . . . . .	310
11.7.1	System Overview . . . . .	310
11.7.2	EF Farm Management . . . . .	312
11.7.3	Operational Monitoring of the Online System . . . . .	316
11.7.4	Physics Monitoring . . . . .	318
11.7.5	System Administration . . . . .	323
11.8	Standalone DAQ Mode & Commissioning . . . . .	324
<b>12</b>	<b>Event Filter</b>	<b>329</b>
12.1	Event Filter Hardware . . . . .	331
12.2	Selection software . . . . .	332
12.3	Study of GPGPU usage in the Event Filter (EF) . . . . .	334
12.4	Model for CPU estimation . . . . .	337
<b>13</b>	<b>Hardware-based Tracking for the Trigger (HTT)</b>	<b>343</b>
13.1	Overview of the HTT Architecture . . . . .	345
13.2	Interface with Event Filter . . . . .	347
13.3	Comparison to FTK . . . . .	348
13.4	Functional description of HTT . . . . .	349
13.4.1	Data preparation . . . . .	349
13.4.2	Pattern matching . . . . .	349
13.4.3	Track fitting . . . . .	352
13.4.4	Duplicate removal . . . . .	354
13.5	HTT Performance Studies . . . . .	354
13.5.1	Associative Memory (AM) Pattern-Matching Performance . . . . .	354
13.5.2	Track-Fitting Performance . . . . .	356
13.5.3	Muon and Electron Track-finding Efficiencies . . . . .	358
13.5.4	Resolutions of Track Parameters . . . . .	358
13.5.5	Simulated HTT Data Size . . . . .	360
13.6	Description of the HTT Hardware and Firmware . . . . .	362
13.6.1	Dataflow requirements . . . . .	363
13.6.2	Tracking Processor (TP) . . . . .	367
13.6.3	Pattern Recognition Mezzanine (PRM) . . . . .	370
13.6.4	Associative Memory (AM) for Phase-II . . . . .	376
13.6.5	Track Fitter Mezzanine (TFM) . . . . .	383
13.6.6	The HTT Interface (HTTIF) Infrastructure . . . . .	386
13.6.7	Dataflow summary . . . . .	387
13.6.8	Size and power consumption of the HTT system . . . . .	387
13.7	Project Milestones . . . . .	390

---

<b>14 Evolution Scenario</b>	<b>393</b>
14.1 Criteria for Evolution	393
14.1.1 Uncertainty in Hadronic Trigger Rate Estimates	394
14.1.2 Uncertainty in ITk Pixel Detector Occupancies	395
14.2 Requirements for the Evolved System	396
14.2.1 ITk Pixel Detector	398
14.2.2 ITk Strip Detector	398
14.2.3 NSW	399
14.3 Overview of the Evolved System Architecture	399
14.4 Design Implications for the TDAQ Sub-systems	401
14.4.1 Level-0 Calo Trigger	401
14.4.2 Level-0 Muon Trigger	401
14.4.3 Global Trigger	401
14.4.4 Trigger, Timing, and Control	403
14.4.5 Central Trigger System	404
14.4.6 Summary of Hardware Trigger Latency Estimates in the Evolved System	407
14.4.7 Data Acquisition	407
14.4.8 Event Filter	409
14.4.9 HTT	411
14.5 Physics Opportunities with the Evolved System	416
14.6 Challenges of the Evolved System	418
<b>15 DCS and TDAQ interfaces</b>	<b>421</b>
15.1 Interfaces for On-Detector Front End (FE) Components	422
15.1.1 Exclusive DCS On-Detector Control Path	423
15.1.2 On-Detector Control via the Readout Path	424
15.2 Interfaces for Off-Detector Front-End Components	425
15.2.1 Devices with Proprietary Interfaces	426
15.2.2 Off-detector Electronics based on ATCA and VME Standards	426
15.2.3 Interfaces for Other Purpose-Built Off-Detector Front-Ends	427
15.3 Back-End and Middleware	427
15.3.1 Back-End Hardware and Software	427
15.3.2 Standard Middleware OPC UA	428
15.3.3 Network	430
15.4 DCS of TDAQ Sub-Systems	430
<b>16 System Integration, Installation and Commissioning</b>	<b>433</b>
16.1 Overview of the TDAQ Phase-II Integration and Installation	433
16.1.1 Location of the TDAQ Phase-II Upgrade Project's Components	434
16.1.2 Installation and Commissioning Planning	434
16.2 Laboratories and Facilities	437
16.2.1 Building 4 surface testing facility	437

16.2.2	Lab4 Testbed Infrastructure . . . . .	437
16.2.3	TDAQ Maintenance Facility . . . . .	438
16.2.4	Safety . . . . .	438
16.3	Production Firmware Deployment . . . . .	440
16.4	Validation and Initial System Integration at CERN of custom Hardware sub-systems . . . . .	441
16.4.1	Quality Assurance and Quality Control at the production sites . . . . .	441
16.4.2	Acceptance validation tests at CERN . . . . .	442
16.4.3	Level-0 Trigger Sub-system Integration Tests . . . . .	442
16.4.4	HTT specific Integration and Initial Commissioning . . . . .	443
16.4.5	DAQ Readout Sub-system: FELIX and Data Handlers . . . . .	443
16.5	Installation in USA15 . . . . .	444
16.6	Integration with the Detector Systems and Early Commissioning . . . . .	449
16.7	Installation and Commissioning Plans of DAQ/EF components in SDX1 . . . . .	450
<b>III Project Management and Organisation</b>		<b>455</b>
<b>17</b>	<b>Project Management and Organisation</b>	<b>457</b>
17.1	Overview of the ATLAS Upgrade Organisation . . . . .	457
17.1.1	ATLAS management . . . . .	458
17.1.2	Executive Board . . . . .	458
17.1.3	Upgrade Coordinator . . . . .	459
17.1.4	Upgrade Steering Committee . . . . .	459
17.1.5	Upgrade Advisory Board . . . . .	459
17.1.6	Interactions with External Committees and Organisations . . . . .	460
17.1.7	Upgrade Projects . . . . .	460
17.2	Organisation and management of the TDAQ systems and the TDAQ Upgrades	461
17.2.1	TDAQ Institutional Board . . . . .	462
17.2.2	TDAQ Management Team . . . . .	463
17.2.3	TDAQ Steering Group . . . . .	463
17.2.4	TDAQ Resources . . . . .	463
17.2.5	TDAQ system and TDAQ Upgrade Projects . . . . .	464
17.3	TDAQ UPR Organisation and Upgrade Project Management Plan . . . . .	465
17.3.1	UPR Management Team . . . . .	465
17.3.2	Upgrade Project Leader (UPL) . . . . .	467
17.3.3	Deputy Upgrade Project Leaders . . . . .	467
17.3.4	Extended TDAQ Steering Group . . . . .	468
17.3.5	UPR Level-0 Trigger Coordinator . . . . .	468
17.3.6	UPR DAQ Coordinator . . . . .	469
17.3.7	UPR EF & Performance Coordinator . . . . .	469
17.3.8	TDAQ Upgrade Project Office . . . . .	470
17.3.9	Organisation of the TDAQ UPR Systems and Sub-systems . . . . .	470

<b>18 Project Cost Estimates</b>	<b>477</b>
18.1 Overview of the Cost Management Plan . . . . .	477
18.2 CORE Costing Policy . . . . .	478
18.3 Costing Methodology . . . . .	479
18.3.1 Estimate Uncertainties . . . . .	481
18.4 UPR Project Breakdown Structure (PBS) . . . . .	482
18.5 UPR CORE Costing Tables . . . . .	486
18.5.1 Level-0 Trigger . . . . .	487
18.5.2 Data Acquisition System . . . . .	488
18.5.3 Event Filter . . . . .	488
18.6 Costing Profiles . . . . .	491
18.6.1 Spending profile of the Level-0 Trigger System . . . . .	493
18.6.2 Spending profile of the DAQ System . . . . .	493
18.6.3 Spending profile of the EF Trigger System . . . . .	494
<b>19 Planning and Schedule</b>	<b>497</b>
19.1 Overview of the Schedule Management Plan . . . . .	497
19.2 Design, Production Plan and Hardware/Firmware Integration . . . . .	499
19.2.1 Specification Review . . . . .	500
19.2.2 Preliminary Design Review . . . . .	501
19.2.3 Final Design Review . . . . .	501
19.2.4 Production Readiness Review . . . . .	502
19.2.5 Internal TDAQ Reviews . . . . .	503
19.3 Production Schedule and Milestones . . . . .	504
19.3.1 Level-0 Trigger System's Milestones . . . . .	507
19.3.2 DAQ System's Milestones . . . . .	509
19.3.3 EF System's Milestones . . . . .	509
<b>20 Resources Requirements and Institutional Responsibilities</b>	<b>513</b>
20.1 Resource Management Plan . . . . .	513
20.2 Required Manpower Estimate . . . . .	514
20.2.1 Level-0 Trigger System . . . . .	514
20.2.2 DAQ System . . . . .	514
20.2.3 EF System . . . . .	517
20.3 Participating Institute Responsibilities . . . . .	524
<b>21 Risk Analysis and Mitigation Strategies</b>	<b>533</b>
21.1 Risk Management Plan . . . . .	533
21.2 Risk Register . . . . .	535
21.2.1 Detector Readout Limitations . . . . .	537
21.2.2 Projected rates for hadronic trigger signatures . . . . .	537
21.2.3 Resource availability of the EF Processing Unit farms . . . . .	537
21.2.4 Hardware Tracking ASIC . . . . .	538

## Contents

---

21.2.5 FPGA resource usage . . . . .	538
21.2.6 Interfaces and Link speeds . . . . .	538
21.2.7 Latency limitations . . . . .	538
21.2.8 FPGA obsolescence . . . . .	539
21.2.9 Commissioning time of the Level-0 Muon Trigger sub-system . . . . .	539
21.2.10 Commissioning time of the Global Trigger sub-system . . . . .	539
21.2.11 Readout's I/O bandwidth margins . . . . .	539
21.2.12 Data Handler processing capabilities for detector-specific functionality	540
21.2.13 Event Size . . . . .	540
21.2.14 Simulation software schedule . . . . .	540
<b>IV Appendix: Glossary</b>	<b>543</b>
<b>The ATLAS Collaboration</b>	<b>557</b>



# Executive Summary

This Technical Design Report documents the strategy and the design of the [ATLAS Trigger and Data Acquisition System \(TDAQ\)](#) system for the High Luminosity upgrade of the Large Hadron Collider ([HL-LHC](#)). The Phase-II upgrade of the [TDAQ](#) system will enable the broad physics programme planned for the [HL-LHC](#), including a detailed exploration of the mechanism of electroweak symmetry breaking through the properties of the Higgs boson, searches for new physics through the study of rare Standard Model processes, searches for new heavy states, and measurements of the properties of any newly discovered particles.

The [HL-LHC](#) is expected to start operations in the middle of 2026, and to reach nominally a peak instantaneous luminosity of  $\mathcal{L} = 5 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ , corresponding to approximately 140 inelastic proton-proton collisions per bunch crossing, which could be maintained for the duration of the [HL-LHC](#) project. In this configuration, approximately  $250 \text{ fb}^{-1}$  will be delivered per year, for a total of  $3000 \text{ fb}^{-1}$  by the end of Run 6, ten times the integrated luminosity of the LHC Runs 1-3 combined. The design of the [HL-LHC](#) accommodates large safety margins with respect to heat deposition and integrated radiation dose. The performance of the [HL-LHC](#) could be accordingly increased to its *ultimate* configuration, where the instantaneous luminosity is levelled at  $\mathcal{L} = 7.5 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$  and the experiments need to cope with pile-up of up to 200 inelastic collisions per bunch crossing. The *ultimate* performance configuration, as defined in the [HL-LHC](#) Technical Design Report, would provide more than  $300\text{-}350 \text{ fb}^{-1}$  per year and a total integrated luminosity up to  $4000 \text{ fb}^{-1}$ . This is the configuration for which the [HL-LHC TDAQ](#) upgrade is designed.

Meeting these requirements poses significant challenges to the ATLAS [TDAQ](#) system to fully exploit the physics potential of the [HL-LHC](#). The ATLAS collaboration described its initial plans and goals for the corresponding Phase-II upgrade of the detector in the Letter of Intent in 2012, and in the Scoping Document in 2015. Two “custom-hardware” trigger levels were proposed in the Scoping Document that allowed for data streaming off-detector either after an initial trigger decision, or in some cases, at the full 40 MHz bunch crossing rate. Since then, the design of the upgraded architecture of the [TDAQ](#) system has developed further, resulting in a baseline architecture with a single-level hardware trigger that features a maximum rate of 1 MHz and  $10 \mu\text{s}$  latency.

The hardware-based Level-0 Trigger system is composed of the Level-0 Calorimeter Trigger ([L0Calo](#)), the Level-0 Muon Trigger ([L0Muon](#)), the [Global Trigger](#) and the Central Trigger sub-systems. In the [L0Calo](#) sub-system, the Phase-I calorimeter feature extraction trigger processors will be complemented by a new [forward Feature EXtractor \(fFEX\)](#) to reconstruct

forward jets and electrons, matching the pseudo-rapidity coverage of the new tracker system. The new [L0Muon](#) sub-system will use upgraded Barrel and Endcap Sector Logic and [New Small Wheel \(NSW\)](#) Trigger Processors, for the reconstruction of muon candidates in the barrel [Resistive Plate Chambers \(RPCs\)](#), in the endcap [Thin Gap Chambers \(TGCs\)](#) and [NSW](#) detectors, respectively. In addition, [Monitored Drift Tube \(MDT\)](#) information will be used in new dedicated processors to improve robustness and efficiency of the Muon Trigger, its  $p_T$  resolution and selectivity. The [Global Trigger](#) will replace and extend the Run 2 and Phase-I Topological Processor, by accessing full-granularity calorimeter information to refine the trigger objects calculated by [L0Calo](#), perform offline-like algorithms, and calculate event-level quantities before applying topological selections. The final trigger decision is made by the [Central Trigger Processor \(CTP\)](#), which can apply flexible prescales and vetoes to the trigger items. The [CTP](#) also drives the Trigger, Timing and Control system network to start the readout process of the detectors.

The result of the Level-0 trigger decision is transmitted to all detectors and trigger processors, upon which the resulting detector and trigger data, respectively, are transmitted to the Data Acquisition system at 1 MHz through the [Readout](#) and the [Dataflow](#) sub-systems. Both sub-systems are based on commodity [PC](#) servers and standard networking infrastructure. The [Readout](#) sub-system includes custom input/output cards to convert the detector front-end protocol data into standard network packets.

The upgraded [EF](#) system provides high-level trigger functionality, and consists of a [CPU](#)-based processing farm complemented by [Hardware-based Tracking for the Trigger \(HTT\)](#) co-processors. The [EF](#) system refines the trigger objects in order to achieve a maximum output event rate of 10 kHz. The [HTT](#) includes regional ([rHTT](#)) and full-scan ([gHTT](#)) track reconstruction capabilities. The [EF](#) trigger decision enables the transfer of data corresponding to selected events from the [Data Acquisition System \(DAQ\)](#) to permanent storage.

Each system and sub-system will be capable of evolving to a dual-level hardware trigger architecture as a mitigation strategy in case pile-up conditions at the [HL-LHC](#) either challenge the readout capabilities of detectors to the limits of the bandwidth available, or in case the rates of hadronic trigger signatures with the needed thresholds exceed our current allocations. The evolved two-level architecture specifies a Level-0 trigger rate up to 2-4 MHz and 10  $\mu$ s latency, followed by a Level-1 trigger rate of 600-800 kHz and latency up to 35  $\mu$ s. Hardware-based track reconstruction is implemented in the Level-1 trigger system by reconfiguring part of the [HTT](#). The reconstructed tracks are combined with calorimeter- and muon-based trigger objects in the [Global Trigger](#), after which the Central Trigger forms the Level-1 decision.

A management plan is in place to deliver the baseline [TDAQ](#) upgraded system fully commissioned by the end of Long Shutdown 3.

# **Part I**

## **General Overview of the TDAQ Phase-II Upgrade Project**



# 1 Introduction

This Technical Design Report describes the upgrade of the ATLAS Trigger and Data Acquisition (TDAQ) system for operation at the [High Luminosity LHC \(HL-LHC\)](#). The overall goal of the TDAQ Phase-II Upgrade Project is to design, build, and install new trigger and data acquisition hardware with its firmware and needed software during the third long shutdown of the [LHC](#) in 2024. In this document, the scientific motivation and the technical implementation of the upgrade are detailed.

The Phase-II upgrade of the ATLAS TDAQ system must satisfy the broad ATLAS physics programme planned for the decade-long [HL-LHC](#), while coping with ultimate [HL-LHC](#) conditions: a peak instantaneous luminosity seven times above the original [LHC](#) design with up to 200 inelastic proton-proton collisions per beam crossing. During this period, ATLAS aims to collect a total dataset of  $4000 \text{ fb}^{-1}$ , allowing for a detailed exploration of the mechanism of electroweak symmetry breaking through the properties of the Higgs boson, searches for new physics through the study of rare Standard Model processes, searches for new heavy states, and measurements of the properties of any newly discovered particles. Both the necessity of a highly efficient selection of events with Higgs bosons in decay modes useful for e.g. self-coupling measurements and of events accessing new physics scenarios in regions of parameter space yet unexplored, requires exceptional trigger and data acquisition performance over the decade-long [HL-LHC](#) run.

This upgrade design takes advantage of and surpasses the Phase-I upgrade of the ATLAS TDAQ system, which will be installed during the second long shutdown of the [LHC](#) starting in 2019. The Phase-I TDAQ upgrade was designed to efficiently select and record events of interest at instantaneous luminosities that are up to twice that of the nominal [LHC](#) design luminosity with up to 80 proton-proton collisions per beam crossing, while maintaining trigger thresholds close to those used in Run 1. The [HL-LHC](#) is expected to start operations in the middle of 2026 and the performance increase to its ultimate configuration, where the instantaneous luminosity is levelled at  $\mathcal{L} = 7.5 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$  and the experiments need to cope with a pile-up up to 200 inelastic collisions per bunch crossing. The Phase-II TDAQ upgrade is designed to cope with this LHC ultimate luminosity.

The upgraded trigger system will take advantage of increased granularity provided by the calorimeters, will improve efficiency for muon-based triggers and perform hardware-based tracking profiting from the extended coverage of the planned [Inner Tracker \(ITk\)](#) [1.1]. A factor of 10 higher trigger rates than the ones expected in Run 3 are needed to retain the events to perform the physics programme in Run 4. The upgrade of the TDAQ system will

## 1.1 LHC Roadmap

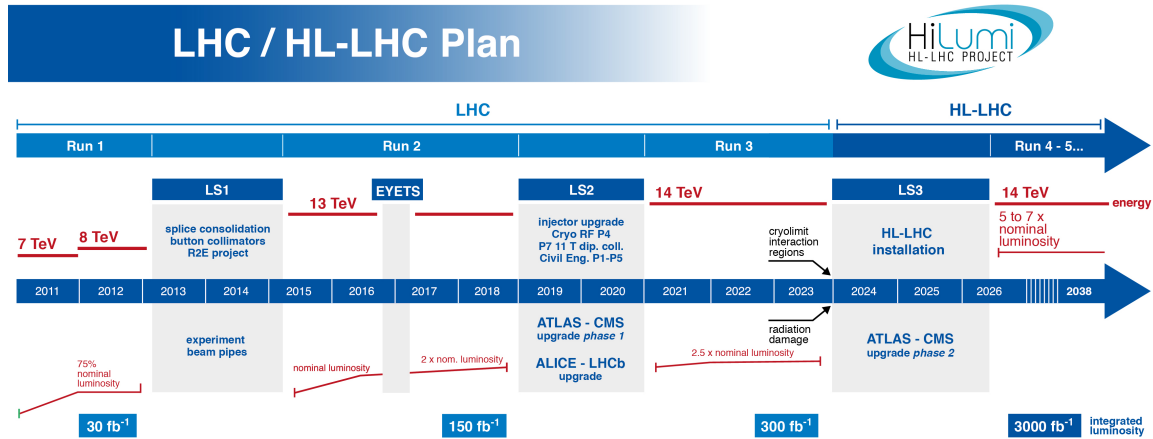


Figure 1.1: Baseline upgrade programme of the LHC accelerator complex over the next two decades. The plan shows the scheduled long shutdowns for the machine and experiment upgrades, and the projected luminosities delivered by the LHC in a three-year run period following each long shutdown [1.2].

provide the required bandwidth and processing capacity to efficiently select events at high luminosity.

## 1.1 LHC Roadmap

The LHC and HL-LHC project schedule is shown in Fig. 1.1. After the consolidation of the machine elements during the Long Shutdown 1, the LHC operated at a centre-of-mass energy of  $\sqrt{s} = 13$  TeV during Run 2, and delivered in the second half of 2017 an instantaneous luminosity in excess of  $\mathcal{L} = 2.0 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ , and an average of up to  $\langle \mu \rangle \simeq 60$  minimum bias collisions per bunch crossing. By the end of Run 2 the LHC is expected to have delivered a total of about  $150 \text{ fb}^{-1}$  to each of the experiments. During the Long Shutdown 2, scheduled for 2019-2020, the luminosity production will be consolidated and the upgraded (Phase-I) experiments will be able to record up to approximately  $300 \text{ fb}^{-1}$  during Run 3 (2021-2023). The HL-LHC project [1.2] is planned to begin collisions in the second half of 2026 and will allow ATLAS [1.1][1.3][1.4] to collect an integrated luminosity of  $3000 \text{ fb}^{-1}$  after ten years of operation. The HL-LHC upgrades of both the machine and the experiments will occur in the Long Shutdown 3 between 2024-2026, allowing full exploitation of the LHC physics programme.

The high luminosity configuration of the LHC will target an integrated luminosity a factor of 10 higher than that expected to be delivered by the end of Run 3. Figure 1.2a shows the possible evolution of the luminosity during the HL-LHC era. In Run 4 the luminosity will progressively increase potentially attaining a levelled instantaneous peak luminosity

Table 1.1: Comparison of the nominal LHC parameters with those of three possible HL-LHC schemes. The levelled luminosity is assumed for  $\mu \simeq 140$ . The levelling time assumes no emittance growth.

Parameter	Nominal LHC	Nominal HL-LHC 25ns		
	[Design Report]	[standard]	[BCMS]	[8b4e]
Beam energy in collision [ TeV]	7	7	7	7
Number of protons per bunch [ $\times 10^{11}$ ]	1.15	2.2	2.2	2.3
$n_b$	2808	2748	2604	1968
Number of collisions in IP1 and IP5	2808	2736	2592	1960
Beam current [A]	0.58	1.09	1.03	0.82
crossing angle [ $\mu$ rad]	285	590	590	554
beam separation [ $\sigma$ ]	9.4	12.5	12.5	12.5
$\beta^*$ [m]	0.55	0.15	0.15	0.15
$\epsilon_n$ [ $\mu$ m]	3.75	2.50	2.50	2.2
$\epsilon_L$ [eVs]	2.5	2.5	2.5	2.5
Levelled luminosity [ $\times 10^{34} \text{cm}^{-2} \text{s}^{-1}$ ]	-	5.32	5.02	5.03
Events / crossing	27	140	140	140
Levelling time [hours]	-	8.3	7.6	9.5

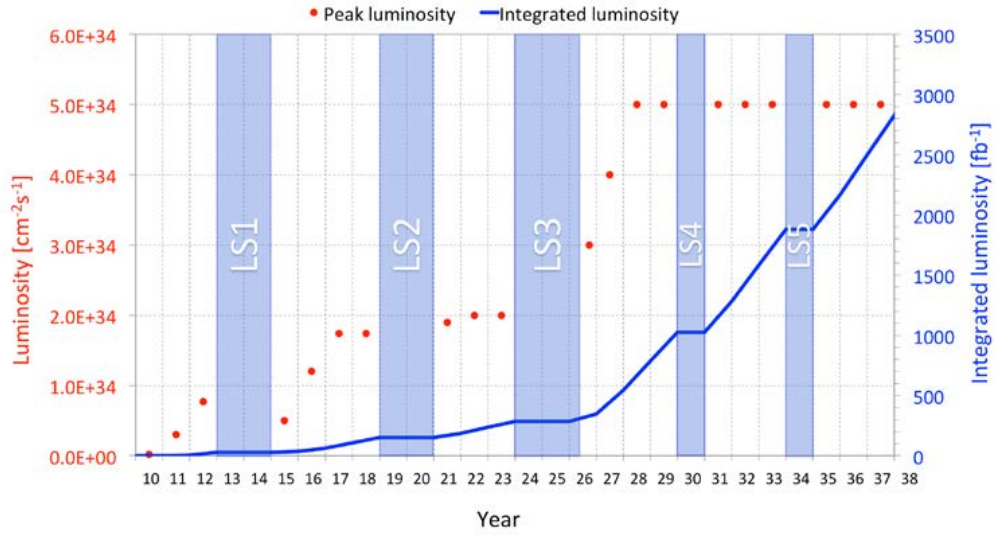
of  $\mathcal{L} = 5 \times 10^{34} \text{cm}^{-2} \text{s}^{-1}$  ( $\langle \mu \rangle \simeq 140$ <sup>1</sup>), which would be maintained for the duration of the HL-LHC lifetime. In this configuration, approximately  $250 \text{fb}^{-1}$  will be delivered to each experiment per year, for a total of  $3000 \text{fb}^{-1}$  by the end of Run 5. The configuration assumes 160 days of physics operations of the LHC complex per year and a physics efficiency of 50%.

The main parameters of the HL-LHC have been re-optimised in 2016 [1.2]. Table 1.1 compares the nominal parameters of the LHC with three possible HL-LHC configuration - schemes, in operation in times in 2017:

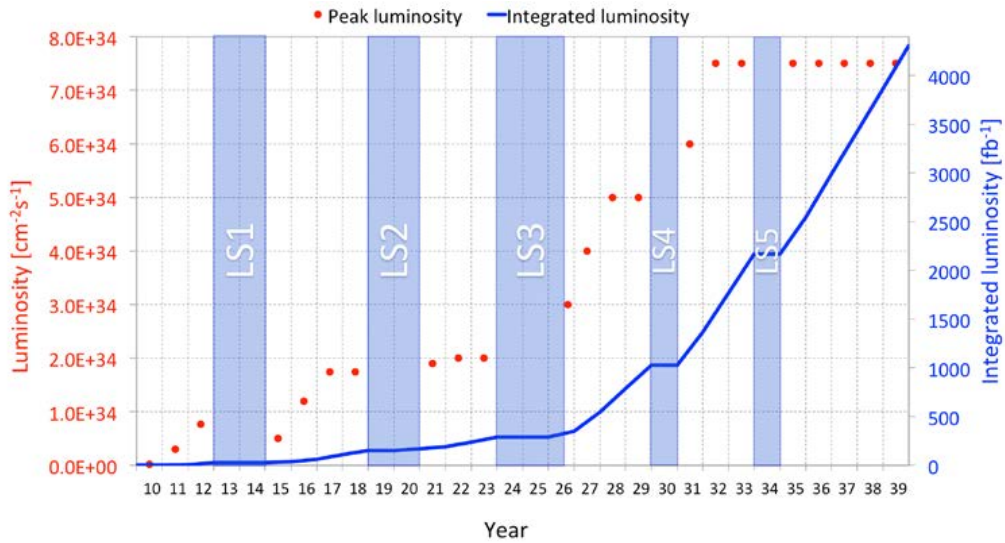
- the standard 25 ns configuration,
- the “Batch Compression Merging and Splitting (BCMS)” scheme: a different configuration with reduced transverse emittance which may be utilised if operations with high beam intensities result in unforeseen emittance blow up, and
- the “8b4e” configuration consisting of 8 colliding bunches followed by 4 empty (collisionless) bunches which may be implemented to mitigate the impact of electron-cloud effects.

<sup>1</sup> The average number of interactions per bunch crossing values are calculated using the formula  $\langle \mu \rangle = \mathcal{L} \times \sigma_{\text{inel}} / n_b \times f_{\text{rev}}$ , where the total inelastic proton-proton cross section is assumed to be  $\sigma_{\text{inel}} = 85 \text{mb}$  at  $\sqrt{s} = 14 \text{TeV}$ , the LHC revolution frequency is  $f_{\text{rev}} = 11.245 \text{kHz}$ , and  $n_b$  is the number of colliding bunches at the ATLAS interaction point.

## 1.1 LHC Roadmap



(a) Nominal *HL-LHC* configuration.



(b) Ultimate *HL-LHC* configuration.

Figure 1.2: Planned instantaneous (●) and integrated (—) luminosity profiles for two *HL-LHC* configurations [1.2].



Table 1.2: Comparison between the planned [HL-LHC](#) nominal and ultimate luminosity parameters.

Configuration	$\mathcal{L}_{\text{inst}}$ [ $10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ ]	$\langle \mu \rangle$	$\int \mathcal{L}$ per year [ $\text{fb}^{-1}$ ]
Baseline	5	140	250
Ultimate	7.5	200	>300

If the performance of the [HL-LHC](#) can be pushed further and the experiments are able to cope with pile-up up to  $\langle \mu \rangle \simeq 200$ <sup>2</sup>, the ultimate [HL-LHC](#) scenario shown in Fig. 1.2b could be realised. Table 1.2 presents a comparison between the two configurations. After the Long Shutdown 4 (2030) the instantaneous levelled luminosity could reach  $\mathcal{L} = 7.5 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ , resulting to more than  $300 \text{ fb}^{-1}$  per year and up to  $4000 \text{ fb}^{-1}$  at the end of the [HL-LHC](#) lifetime.

## 1.2 Physics Drivers for the HL-LHC Upgrades

The Phase-II [TDAQ](#) upgrade must support the broad ATLAS physics programme for the [HL-LHC](#); this programme has been presented and discussed in detail in several documents, here listed in chronological order: (i) the Phase-II Upgrade Letter of Intent [1.4], dating from 2012, (ii) the two reports submitted to the European Committee for Future Accelerators (ECFA) [1.5][1.6], published in 2013 and 2014 respectively, and (iii) the Scoping Document [1.1] released in late 2015. Table 1.3 presents the wide spectrum of physics goals and a representation of analyses that will be carried out by ATLAS to exploit the full potential of the [HL-LHC](#). Also given are the corresponding trigger signatures. These goals include unveiling the paradigm of electroweak symmetry breaking through precision measurements of the properties of the Higgs boson, improved measurements of all relevant Standard Model parameters including the study of rare Standard Model processes, searches for [Beyond the Standard Model \(BSM\)](#) signatures and flavour physics. The trigger has to address also specific challenges of the heavy-ion physics programme. This [Technical Design Report \(TDR\)](#) documents a subset of analyses in Chapter 2, and Table 1.3 includes references to the relevant sections in Chapter 2 for those selected analyses, focusing on the enhancements of the physics object measurements enabled by the upgrades. Unless otherwise specified, the trigger performance studies and the benchmark physics analyses have been carried through assuming the pile-up conditions for the ultimate [HL-LHC](#) configuration,  $\langle \mu \rangle \simeq 200$ .

<sup>2</sup> A benchmark scenario with a  $\langle \mu \rangle$  of approximately 200 is obtained by assuming  $n_b = 2808$  and a peak instantaneous luminosity of  $\mathcal{L} = 7.5 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ . For more details, see the [HL-LHC Technical Design Report](#) [1.2].

Table 1.3: Outline of the main physics drivers relevant for the TDAQ system at the HL-LHC, including examples of final state processes and their corresponding trigger signatures. Only a limited sub-set of those processes are documented in this TDR, reported in the sections referenced in the last column. The abbreviation “diff” stands for differential.

Physics Drivers @ HL-LHC	Processes	Trigger Signatures	TDR Sect.	
Precision measurements of the properties of the Higgs Boson	Couplings to fermions	$H \rightarrow \tau\tau, H \rightarrow \mu\mu, t\bar{t}H, H \rightarrow b\bar{b}$	single/di- $e$ or $\mu$ / di- $\tau$	2.2, 2.4
	Couplings to $W/Z$ , diff. cross-sections	$H \rightarrow \gamma\gamma, H \rightarrow W^+W^- \rightarrow \ell^+\nu\ell^-\nu, H \rightarrow ZZ^{(*)} \rightarrow \ell^+\ell^-\ell^+\ell^-$	$e/\mu, \text{di-}\gamma$	2.3
	Self-coupling	$HH \rightarrow b\bar{b}\tau\tau / b\bar{b}\gamma\gamma / 4b$	di- $\tau/\gamma$ , multi-jets	2.5
	Scalar Higgs boson vs. BSM composite	$H \rightarrow \ell\ell', ZH \rightarrow \ell\ell + (\text{inv})$	$e/\mu$	
Precision Standard Model Measurements	Forward/backward asymmetry	$Z \rightarrow e^+e^-, \mu^+\mu^-$	single $e/\mu$	
	Vector-boson scattering	$WWjj, WZjj$	single $e/\mu$	
	Precision top mass and cross-sections	$t\bar{t}$ production	$e/\mu$ , large $R$ -jets/multi-jets	2.1
Searches for BSM Signatures	Searches for new vector bosons	Vector Boson Fusion (VBF) $Z' \rightarrow \ell\ell$	high- $p_T$ single $e/\mu$	
	Searches for electroweak SUSY	$\chi_1^+\chi_2^0 \rightarrow WH\chi_1^0\chi_1^0$	$E_T^{\text{miss}}$ , single/di- $e, \mu, \tau$	2.2, 2.6
	SUSY top partners	$\tilde{t}_1 \rightarrow t\chi_1^0$	large $R$ -jets/multi-jets+ $E_T^{\text{miss}}$	
	Dark matter	ISR+ $\chi_1^0\chi_1^0$	jets+ $E_T^{\text{miss}}$	
New resonances, SUSY	$Z', \chi_1^0 \rightarrow jjj$	jets, large $R$ -jets, $e/\mu, \gamma, E_T^{\text{miss}}$	2.3, 2.5	
Long-lived particles	$\tilde{g} \rightarrow q\bar{q}\tilde{\chi}_1^0, \tilde{\chi}_1^\pm \rightarrow \pi^\pm\tilde{\chi}_1^0$	high impact parameter, $E_T^{\text{miss}}$		
Flavour Physics	Lepton Flavour Violation	$\tau \rightarrow \mu\mu\mu$	low $p_T$ di- $\mu$	
	Searches for FCNC in top decays	$t \rightarrow u/c + H/Z$	single $e/\mu$	
	Rare $B$ -meson decays	$B \rightarrow \mu\mu, B_s \rightarrow J/\Psi + \Phi$	low $p_T$ di- $\mu$	2.10
Heavy-Ion Physics	Light-by-light scattering	$\gamma\gamma \rightarrow \gamma\gamma$	low $p_T$ di-photons	
	Electroweak production	$W/Z/t$	single $e/\mu$	
	In-medium parton energy loss (jets in PbPb)	mono-jets	jet, minbias	
	Quarkonia production	$J/\Psi, \Upsilon$	low-mass di- $e/\mu$	2.11

Table 1.4: *ATLAS HL-LHC Upgrade plans and reference to the detector system TDRs.*

Detector System	Upgrade scope	CDS Reference
<a href="#">ITk Pixel Detector</a>	Sensors, modules, mechanics, <a href="#">FE</a> electronics	<a href="#">CERN-LHCC-2017-021</a>
<a href="#">ITk Strip Detector</a>	Sensors, modules, mechanics, <a href="#">FE</a> electronics	<a href="#">CERN-LHCC-2017-005</a>
<a href="#">LAr Calorimeter</a>	<a href="#">FE</a> and <a href="#">BE</a> electronics	<a href="#">CERN-LHCC-2017-018</a>
Tile Calorimeter	Mechanics, <a href="#">FE</a> and <a href="#">BE</a> electronics	<a href="#">CERN-LHCC-2017-019</a>
Muon Spectrometer	<a href="#">FE</a> electronics	<a href="#">CERN-LHCC-2017-017</a>
	Inner Barrel <a href="#">MDT</a> chambers	
	Inner Barrel <a href="#">RPC</a> stations	<a href="#">CERN-LHCC-2017-020</a>
<a href="#">TDAQ</a>	On-detector readout and trigger electronics (this document)	

### 1.3 ATLAS Upgrade Strategies for the HL-LHC

The overall upgrade strategy of the ATLAS collaboration for the [HL-LHC](#) upgrades is described in detail in the Phase-II Letter of Intent [1.4] and, more recently, in the Scoping Document [1.1].

In most cases, the design and techniques proposed for the Phase-II upgrades represent an evolution from the new designs and technologies already introduced during the Long Shutdown 1 consolidation programme, and from the Phase-I upgrades now being prepared for installation during Long Shutdown 2. This evolutionary approach leads to a relatively solid understanding of the technical challenges, as well as the uncertainties on the costs and the required resources, even at the present early stage of the Phase-II upgrade activities.

A synoptic view of the upgrade plans, with the references to the [TDRs](#) of each system, is presented in Table 1.4. An internal ATLAS document [1.7] defines the trigger and readout requirements on the interfaces to the front-end electronics systems of each detector in Table 1.4.

#### Inner Tracker

New silicon strip and pixel detectors will be installed during Long Shutdown 3 with a geometrical acceptance up to  $|\eta| = 4.0$ . The final layout of the detectors has evolved with respect to the layout presented in [1.1]. For example, five pixel layers and four strip layers are planned for the [ITk](#) barrel. The full details of the [ITk](#) design are described in the [ITk strip](#) [1.8] and [pixel](#) [1.9] [TDRs](#).

### Calorimeters

The readout electronics of both the [LAr](#) and Tile calorimeters need to be upgraded because of radiation tolerance limits, and because the on-detector front-end electronics cannot operate with the trigger rates and latencies required for the Phase-II luminosities. The current [FCal](#) calorimeter will remain in Phase-II (contrary to the strategy presented in [1.1]). ATLAS has evaluated that the risks of replacing the forward [LAr](#) calorimeter outweighed the possible performance and physics benefits.

### Muon Spectrometer

The scope of the Muon Spectrometer upgrade focuses primarily on the improvement of the performance of the muon trigger chambers. The Level-0 trigger electronics of the [RPC](#) and [TGC](#) chambers will be upgraded. New [RPC](#) detectors will be added in the barrel ( $|\eta| < 1$ ) to significantly increase the solid angle coverage and the redundancy of the Level-0 trigger system. The replacement of the [MDT](#) front-end readout will address the trigger rate and latency requirements of the [TDAQ](#) system in Phase-II and allow the use of [MDT](#) hit information to improve the muon  $p_T$  resolution in the Level-0 trigger.

### Other detectors

One new detector currently expressing interest as an upgrade for Phase-II but not having reached the maturity of providing a [TDR](#) is the [High-Granularity Timing Detector \(HGTD\)](#), a detector intended to distinguish between collisions occurring very close in space but well-separated in time. The current proposed detector design is based on low-gain avalanche detector technology that will cover the pseudorapidity region between 2.4 and 4.0, with a timing resolution of 30 ps for minimum-ionising particles. High-precision timing will improve the track-to-vertex association in the forward region, impacting jet and lepton reconstruction, as well as offering unique capabilities for online and offline luminosity determination. Should this detector be approved, it may impose requirements on the [TDAQ](#) system as described in Section 4.1.4.

### Trigger and Data Acquisition System

Since the ATLAS Phase-II Upgrade Scoping Document<sup>3</sup> [1.1], the design of the architecture of the [TDAQ](#) system has developed, resulting in a single-level hardware trigger in the baseline design. This decision was reached through several approval steps:

---

<sup>3</sup> The TDAQ scenario in the scoping document described two “custom-hardware” trigger levels that allowed for data streaming off-detector either after an initial trigger decision, or in some cases, at the full 40 MHz bunch crossing rate.

- 
- The [TDAQ](#) system prepared an [IDR](#) document that was reviewed in April 2016; the resulting report was approved by the ATLAS Executive Board in June 2016.
  - A sub-committee of the ATLAS Upgrade Steering Committee was appointed by the ATLAS Upgrade Coordinator to address ATLAS-wide aspects of the trigger rates and assess the capabilities and limitations of the detector systems' readout architectures. This sub-committee's final report was discussed and approved by the ATLAS Upgrade Steering Committee in November 2016, recommending a single-hardware-level trigger architecture capable of evolving into a two-level (Level-0 and Level-1) hardware trigger system based on the use of tracking reconstruction in the Level-1 trigger.
  - The ATLAS Executive Board approved the [TDAQ](#) Phase-II Upgrade Project in December 2016; this was endorsed by the ATLAS Collaboration Board in February 2017.

The resulting baseline architecture design is based on a single Level-0 hardware trigger with a detector readout rate of 1 MHz and a maximum latency of 10  $\mu$ s. The Level-0 trigger decision is formed using calorimeter and muon information. The Phase-I calorimeter trigger processors will be maintained during the [HL-LHC](#) operations, and their firmware optimised for the pile-up conditions and the extra latency budget available. The processors will be complemented by additional processors that will implement more sophisticated, offline-like algorithms to provide additional background rejection. The [EF](#) system will further select events based on a commodity processor farm and a custom [HTT](#) to reduce the overall [CPU](#) requirements. The system is designed to accommodate an output rate of 10 kHz to permanent storage, averaged over a fill.

Each system and sub-system is designed to be capable of evolving to a dual-level hardware-based trigger architecture as a mitigation strategy in case pile-up conditions at the [HL-LHC](#) either challenge the readout capabilities of certain detectors (for example of the innermost layers of the [ITk](#)) to the limits of the bandwidth available, or in case the rates of hadronic trigger signatures surpass the current predictions. This two-level hardware trigger design specifies a Level-0 trigger rate up to 4 (2) MHz and 10  $\mu$ s latency, followed by a Level-1 trigger rate of 600 (800) kHz and latency up to 35  $\mu$ s. In this design, hardware-based track reconstruction is implemented in the Level-1 trigger system by reconfiguring part of the [HTT](#); these reconstructed tracks are combined with calorimeter- and muon-based trigger objects to form the Level-1 decision.

## 1.4 Baseline Architecture and Structure of the Upgrade Project

A high-level functional description of the architecture is presented in Fig. 1.3. The three main systems are the Level-0 Trigger System, the Data Acquisition ([DAQ](#)) System, and the Event Filter System<sup>4</sup>.

---

<sup>4</sup> They are at the second level of the [PBS](#) in the project organisation.

## 1.4 Baseline Architecture and Structure of the Upgrade Project

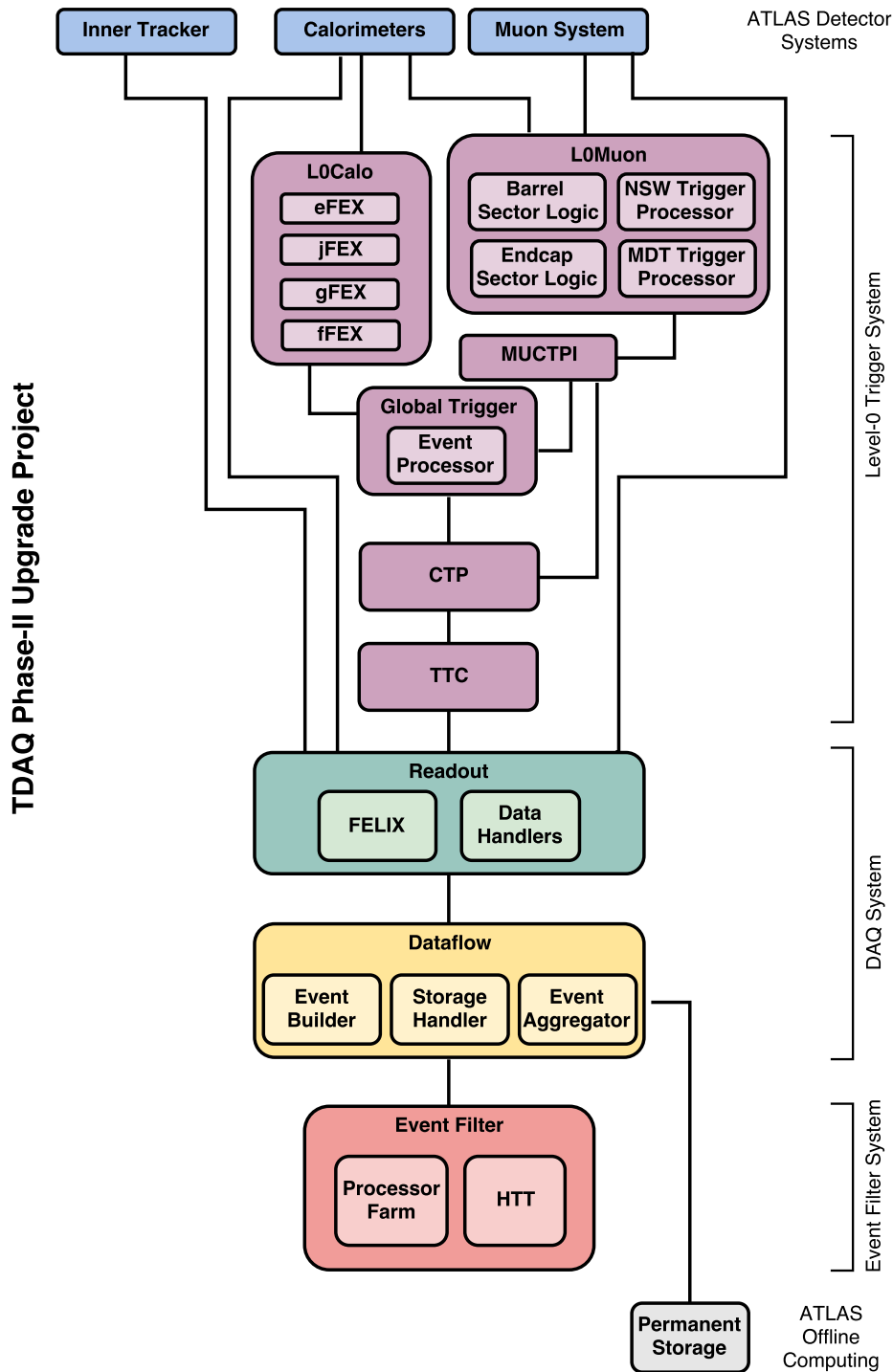


Figure 1.3: Design of the *TDAQ* Phase-II upgrade architecture, highlighting the organisation of the Upgrade Project in three main systems: Level-0 Trigger, *DAQ* (*Readout* and *Dataflow* subsystems), and Event Filter. Direct connections between each Level-0 trigger component and the Readout system are suppressed for simplicity.

---

## Level-0 Trigger System

The hardware-based Level-0 Trigger system is composed of the Level-0 Calorimeter Trigger (**L0Calo**), the Level-0 Muon Trigger (**L0Muon**), the **Global Trigger** and the Central Trigger sub-systems. The **L0Calo** sub-system is composed of the Phase-I **electron Feature EXtractor (eFEX)**, **jet Feature EXtractor (jFEX)**, and **global Feature EXtractor (gFEX)** complemented by a new **fFEX** component to reconstruct forward jets and electrons. The **L0Muon** sub-system comprises the Barrel and Endcap Sector Logic processors (for muon candidates from the barrel **RPC** and **TGC**, respectively), the **NSW** Trigger Processor, and the **MDT** Trigger Processor. The **MUCTPI** component interfaces the **L0Muon** sub-system with the **CTP**. The **Global Trigger** uses full-granularity calorimeter information to perform offline-like algorithms, refines the trigger objects calculated by **L0Calo** and **L0Muon**, calculates event-level quantities, and executes topological algorithms. The **CTP** forms triggers based on combinations of inputs or conditions received from the **Global Trigger** and other sources; it applies pre-scale factors and introduces deadtime when necessary to avoid saturation in the front-end and readout systems; and it ultimately makes the final decision on whether the event is accepted, driving the **Trigger, Timing, and Control system (TTC)** network to start the readout process of the detectors.

## Data Acquisition (DAQ)

The result of the Level-0 trigger decision is transmitted to all detectors, upon which the resulting trigger data and detector data are transmitted at 1 MHz through the **Readout subsystem (Readout)** sub-system, which contains the **Front-End Link eXchange (FELIX)** and Data Handler components, and the **Dataflow subsystem (Dataflow)** sub-system, which contains the Event Builder, Storage Handler, and Event Aggregator components. Together these compose the **DAQ** system. They are based on commodity **Personal Computer (PC)** servers and standard networking infrastructure. The **Readout** system includes custom input/output cards to convert the detector front-end protocol data into standard network packets.

## Event Filter (EF)

The **EF** system consists of a **CPU**-based processing farm and a **HTT** co-processor. The main function of the **EF** system is to refine the trigger objects in order to get down to the final output rate. The **HTT** includes regional (**rHTT**) and global (**gHTT**) track reconstruction capabilities. Events will be rejected as early as possible during their processing. The **EF** trigger decision is communicated to the **Dataflow** system, which transfers accepted events to permanent storage.

The planned physical locations of the Phase-II [TDAQ](#) Upgrade components are described in Section [16.1.1](#). The baseline choice is to install all of the Level-0 Trigger modular electronics, the [DAQ](#) Readout components ([FELIX](#) and the Data Handlers) and the [HTT](#) electronics in the main underground electronics cavern, [USA15](#). The Dataflow and [EF](#) servers will be installed in the surface-level counting room, [SDX1](#).

## 1.5 Outline of this Report

This document is organised into three main parts: a general overview, detailed technical descriptions of the sub-systems, and the project management and organisation.

Part I of this document contains this introduction, a description of the physics motivation for the upgrade (Chapter [2](#)), the limitations of the Run 3 [TDAQ](#) system (Chapter [3](#)) and the subsequent requirements for the Run 4 [TDAQ](#) system (Chapter [4](#)), followed by a description of the proposed [TDAQ](#) system (Chapter [5](#)). Chapter [6](#) contains the expected performance of the Phase-II [TDAQ](#) System.

Part II of this document contains detailed descriptions of each of the components of the proposed system, including hardware specifications and requirements for each component (Chapters [7-13](#)), as well as a discussion of the plan for a potential evolution of the system (Section [14](#)), should it be necessary. A system-by-system guide to this document is presented in Table [1.5](#), along with a list of each system's components and the relevant section of the document that contains their detailed technical descriptions.

Originally the [Detector Control System \(DCS\)](#) was a fully integrated component of the [TDAQ](#) system and reporting to the [TDAQ](#) Management. After installation and commissioning completed in 2008, ATLAS decided to support [DCS](#) operations through Technical Coordination, with the scientific and technical personnel affiliated to the Technical Coordination Project Office. Tight integration between [TDAQ](#) and [DCS](#) groups will be required for the Phase-II upgrades, as [DCS](#) functions will be carried through the [Readout](#) sub-system in a few of the upgrade projects, namely [ITk](#)-strips and [NSW](#). Outlines of the general [DCS](#) plans and of the [TDAQ](#) interfaces are presented in Chapter [15](#) of this report. Nevertheless, it should be highlighted that [DCS](#) is **not** part of the [TDAQ](#) Phase-II Upgrade Project scope, and, consequently, no aspect related to its Project Management will be addressed in this [TDR](#).

The strategy for system integration, installation, and commissioning is described in Chapter [16](#). And finally, Part III of this document focuses on the upgrade project management aspects, covering:

- the organisation of the project and initial project management (Chapter [17](#));
- the product breakdown structure, CORE costs (presented for the baseline [TDAQ](#) architecture), and spending profile (Chapter [18](#));
- the work breakdown structure, planning, and scheduling (Chapter [19](#));



Table 1.5: Organisation of the ATLAS TDAQ system, mapped to the detailed technical descriptions in Part-II of this TDR.

System	Sub-system	Component	Technical Description
Level-0 Trigger	L0Calo	eFEX jFEX gFEX fFEX	Chapter 7
	L0Muon	Barrel Sector Logic Endcap Sector Logic NSW Trigger Processor MDT Trigger Processor	Chapter 8
	MUCTPI	–	Section 10.3
	Global Trigger	–	Chapter 9
	CTP	–	Section 10.1
	TTC	–	Section 10.2
DAQ	Readout	FELIX Data Handlers	Section 11.4
	Dataflow	Event Builder Storage Handler Event Aggregator	Section 11.5
EF	Processor Farm		Chapter 12
	HTT	rHTT gHTT	Chapter 13

- the required human and financial resources (Chapter 20); and
- the risk analysis and mitigation strategies (Chapter 21).

The glossary in Appendix IV defines many of the common acronyms and terms used in this document.

## References

- [1.1] ATLAS Collaboration, *ATLAS Phase-II Upgrade Scoping Document*, Tech. Rep. CERN-LHCC-2015-020. LHCC-G-166, CERN, Geneva, Sep, 2015.  
<https://cds.cern.ch/record/2055248>.
- [1.2] G. Apollinari et al., *High-Luminosity Large Hadron Collider (HL-LHC): Technical Design Report V. 0.1*. CERN Yellow Reports: Monographs. CERN, Geneva, 2017.  
<http://cds.cern.ch/record/2284929>.
- [1.3] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST 3 (2008) S08003.

## 1.5 Outline of this Report

---

- [1.4] ATLAS Collaboration, *Letter of Intent for the Phase-II Upgrade of the ATLAS Experiment*, Tech. Rep. CERN-LHCC-2012-022. LHCC-I-023, CERN, Geneva, Dec, 2012.  
<https://cds.cern.ch/record/1502664>.
- [1.5] ECFA High Luminosity LHC Experiments Workshop: *Physics and Technology Challenges*. 94th Plenary ECFA meeting, ECFA-13-284.  
<https://cds.cern.ch/record/1631032>.
- [1.6] ECFA High Luminosity LHC Experiments Workshop - *Physics and Technology Developments Summary submitted to ECFA2014*. 96th Plenary ECFA meeting, ECFA-15-289. <https://cds.cern.ch/record/1983664>.
- [1.7] F. Lanni et al., *Interfaces with Detector Front-End Systems. Requirement Document for HL-LHC*, Tech. Rep. ATL-DA-ES-0051 v.1, EDMS Id 1563801 v.1, CERN, Geneva, March, 2016. <https://edms.cern.ch/document/1563801>.
- [1.8] ATLAS Collaboration, *Technical Design Report for the ATLAS Inner Tracker Strip Detector*, Tech. Rep. CERN-LHCC-2017-005, CERN, Geneva, Apr, 2017.  
<https://cds.cern.ch/record/2257755>.
- [1.9] ATLAS Collaboration, *Technical Design Report for the ATLAS Inner Tracker Pixel Detector*, Tech. Rep. CERN-LHCC-2017-021, CERN, Geneva, (*in preparation*).

## 2 Physics Motivation

The Phase-II [TDAQ](#) upgrade is designed to enable the [HL-LHC](#) programme which includes a large and rich programme of studies that explores a wide range of topics including rare and precision Higgs boson and Standard Model searches and measurements, searches beyond the Standard Model including [SUSY](#), heavy-flavour physics, and heavy-ion physics.

This chapter describes a subset of the physics measurements that represent the breadth of the planned [HL-LHC](#) physics programme. The representative measurements used to motivate the trigger requirements are selected to be both challenges to the trigger system and central to the scientific goals. Beyond these specific examples, it is important that the system be flexible and the requirements sufficiently general to support new experimental analysis techniques and theoretical ideas in addition to the baseline programme.

The physics requirements for the trigger are a set of  $p_T$  thresholds requirements for the relevant objects ( $e, \mu, \tau, \text{jet}, \dots$ ) and combinations thereof (invariant masses, etc.). Ideal triggers select the same events as the offline object-level reconstruction with good efficiency above the threshold. The effective offline threshold is considered to be the offline reconstructed  $p_T$  where the trigger has reached approximately 90-95% of the plateau efficiency (maximum efficiency for all  $p_T$ ); the exact definition relevant for each trigger varies. In general, systematic uncertainty-limited analyses (e.g., precision studies) do not use data where the trigger is lower than 90-95% of the plateau efficiency because of the additional systematic uncertainty incurred.

For each physics signature described in this chapter, a comparison is made between the thresholds that would be possible in Run 4 conditions with and without the Phase-II upgrade. The no-upgrade scenario, assumes a  $\sim 100$  kHz Level-0 output rate, the Phase-I hardware selections unchanged, and rate allocations comparable to those planned for Phase-I (which has 100 kHz output rate for the first level). For the no-upgrade scenario, the triggers are assumed to be completely limited by the hardware trigger and no additional Event Filter limits are added. Of particular importance for the system design are the single lepton, di-hadronic  $\tau$ , multi-jet, and missing transverse energy triggers.

The [TDAQ](#) upgrade takes advantage of the upgrades of the other ATLAS detector systems to efficiently select physics within or beyond the Standard Model. This is achieved through a cost-effective high-bandwidth readout and the ability to run sophisticated algorithms to distinguish signal from background. The performance of the upgraded system comes from a combination of the high-bandwidth [DAQ](#), which accepts 1 MHz from the Level-0 hardware trigger, and improved object identification delivered at Level-0 by the upgraded

calorimeter, muon, and global trigger systems. In order to support the high rate Level-0, the software-based Event Filter uses hardware-based regional and global tracking systems, [rHTT](#) and [gHTT](#) respectively, for track finding and fitting. Regional tracking is used to do early rejection to reduce the rate below 400 kHz. High-precision global tracking is used in the final selection to reduce the rate to 10 kHz.

The interplay between the Phase-II TDAQ system design (described in detail in Section 5) and representative physics goals, mapped through the corresponding triggers, is summarised in Fig. 2.1. For example, the [Global Trigger](#) enables a low- $p_T$  electron trigger at Level-0. Regional-tracking is then used to reduce the high rate early in the Event Filter processing. The [L0Muon](#) system provides good efficiency at a moderate rate supporting a low threshold single muon trigger. A final example is that the [Global Trigger](#) enables low thresholds for multi-jet and  $E_T^{\text{miss}}$  triggers at Level-0, after which regional tracking and then full-detector tracking are used to reduce the background acceptance rate while preserving the physics acceptance.

The kind of studies used to motivate the trigger thresholds can be classified into three categories: 1) the dependence of the sensitivity on the  $p_T$  threshold for an analysis based on simulated data including background modeling and systematic uncertainty estimates, 2) the dependence of signal acceptance on the trigger threshold, and 3) the impact of the trigger on a Run 2 result. The first category provides the most direct information, however analysis techniques may improve over the long period before the end of the [HL-LHC](#) data-taking. The second category shows the potential gain from the improved trigger acceptance, but does not investigate how much of this gain can be exploited in the full, optimised analysis. The third category is instructive, but does not fully take into account the increased challenges associated with higher instantaneous luminosities or the potential analysis optimisations that can be made with the benefit of two orders of magnitude more data.

The expected performance of the various trigger objects after the Phase-II upgrade is described in Chapter 6, including how trigger acceptance rates depend on thresholds and hardware configuration. Section 6.11 presents the allocation of the [DAQ](#) bandwidth resources, describing the resulting list of trigger thresholds and acceptance rates referred to as the trigger “menu”. The Phase-II trigger menu meets the physics requirements presented in the following sections.

## 2.1 Physics Signatures with Single-Electron and Single-Muon Triggers

Single-electron and single-muon signatures play a central role in [LHC](#) data analyses, since leptons are clear signals of electroweak physics. It is therefore critical to record as many events containing single leptons as possible, including  $W$  and  $Z$  events produced near threshold. The low mass of the Higgs boson, for which precision measurements are a

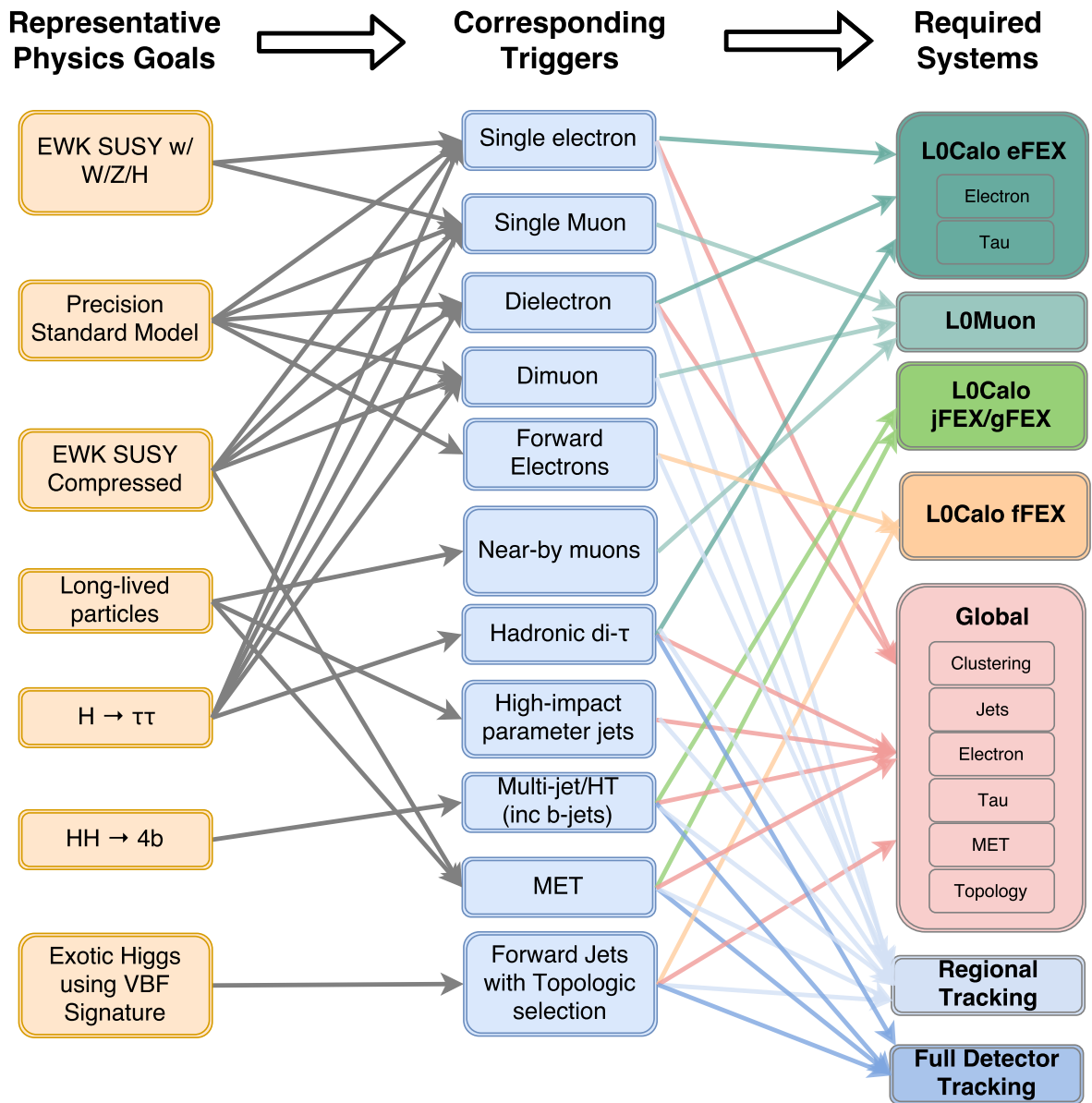


Figure 2.1: Schematic summary of the flow from the representative set of physics goals described in this section (left column) to the hardware systems (right column) needed to achieve these goals. The middle column lists the corresponding triggers required.

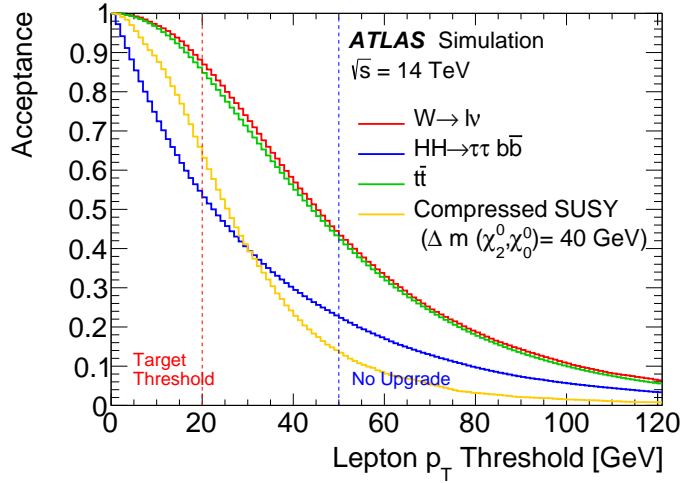


Figure 2.2: The integrated acceptance as a function of the single lepton  $p_T$  threshold for four representative channels:  $W \rightarrow \ell\nu$ ,  $H \rightarrow \tau\tau b\bar{b}$ ,  $t\bar{t}$ , and a compressed spectrum **SUSY** model relevant for “Well-tempered Neutralino” motivated models. The Phase-II TDAQ upgrade would enable lowering the single lepton Level-0 threshold to 20 GeV from 50 GeV, the projected threshold without the upgrade.

primary motivator for the Phase-II physics programme, also requires these single-electron and single-muon thresholds to remain low.

Among the physics processes selected by the single-lepton triggers are  $t\bar{t}$  production, inclusive  $W \rightarrow \ell\nu$ ,  $HH \rightarrow \tau\tau b\bar{b}$  with at least one  $\tau \rightarrow e$  or  $\mu$ , and electroweak **SUSY** signatures with low- $p_T$  leptons. The acceptance for each of these four processes as a function of the lepton  $p_T$  threshold is shown in Fig. 2.2. The **SUSY** model is a “Well-tempered Neutralino” model that is designed to be consistent with the dark matter relic density [2.1]. A threshold of 20 GeV provides good acceptance for  $WH$ ,  $t\bar{t}$ , and  $\tilde{\chi}_1^\pm \tilde{\chi}_2^0 \rightarrow W^\pm \tilde{\chi}_1^0 Z \tilde{\chi}_1^0$  with significant losses if the thresholds are raised to the no-upgrade scenario.

The search for non-resonant  $HH$  production and anomalous Higgs boson self-couplings are key goals of the **HL-LHC** programme. Modification of the Higgs boson self-coupling can lead to changes the cross-section of order unity [2.2]. Specially, because of destructive interference removing the coupling approximately doubles the  $HH$  cross-section. Figure 2.3 shows the sensitivity of the search for  $HH \rightarrow \tau\tau b\bar{b}$ , with one  $\tau \rightarrow e$  or  $\mu$  and one  $\tau$  decaying hadronically, as a function of the offline lepton  $p_T$  requirement which is determined by the trigger threshold. The points in the figure show the estimated sensitivity based on fully simulated signals and backgrounds scaling from the Run 2 result to the **HL-LHC** luminosity and centre-of-mass energy of 14 TeV. The analysis also includes a data-driven estimate of the jets misidentified as  $\tau$  leptons (fake- $\tau$  background), which leads the result to be limited by the required minimum lepton  $p_T$  (27 GeV) corresponding to the Run 2 trigger

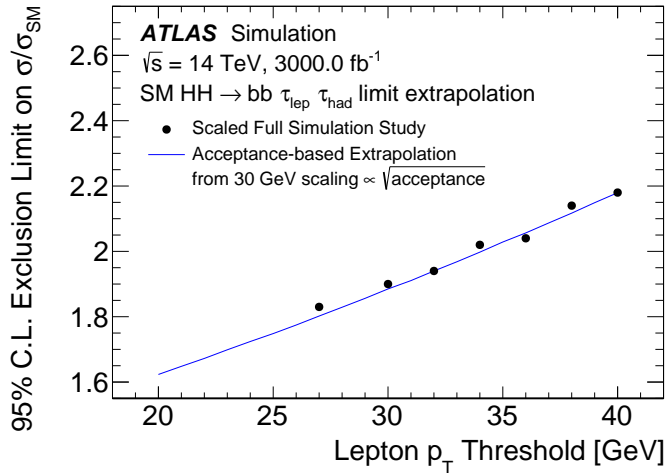


Figure 2.3: Exclusion limit on the ratio of the cross-section to the Standard Model cross-section,  $\sigma/\sigma_{SM}$ , as a function of the leading lepton  $p_T$  requirement for  $HH \rightarrow \tau\tau b\bar{b}$  with one  $\tau \rightarrow e$  or  $\mu$  and one  $\tau$  decaying hadronically. As discussed in the text, modifications of the Higgs self-coupling can modify the cross-section by factors of order unity. The dots are based on scaling signals and backgrounds from the Run 2 result and are thus limited by the lepton  $p_T$  requirement of that analysis to only show above 27 GeV. The acceptance shown in Fig. 2.2 has been used to extrapolate up and down from 30 GeV. Above 30 GeV, the acceptance extrapolation shows good agreement with the Run 2 scaled result.

threshold. Because the fully simulated result cannot be produced for a lepton  $p_T$  threshold below 27 GeV, the acceptance shown in Fig. 2.2 has been used to extrapolate (solid line) up and down from 30 GeV assuming the expected limit just depends on the square-root of the acceptance. Above 30 GeV, the acceptance extrapolation shows good agreement with the Run 2 scaled result. The main background that would increase at lower lepton  $p_T$  is the fake  $e$  or  $\mu$  leptons background, but they are not a major background in the analysis. The acceptance scaling is therefore expected to be a reliable estimate of the gain from reducing the threshold to 20 GeV, giving a  $\approx 70\%$  gain in sensitivity, equivalent to having  $\approx 3\times$  more data than in the no-upgrade scenario with a 50 GeV trigger threshold. The  $HH \rightarrow \tau\tau b\bar{b}$  with one  $\tau \rightarrow e$  or  $\mu$  and one  $\tau$  decaying hadronically channel will be combined with other  $HH$  channels including  $HH \rightarrow \tau\tau b\bar{b}$  with hadronic taus (discussed in Section 2.4),  $HH \rightarrow \gamma\gamma b\bar{b}$  (Section 2.3), and  $HH \rightarrow 4b$  (Section 2.5). Each channel contributes significantly and further possibilities are under investigation.

## 2.2 Physics Signatures with Two Leptons

Even with a low single-lepton  $p_T$  threshold, dilepton triggers are needed to augment the acceptance in certain channels. Figure 2.4 shows the acceptance as a function of the leading and subleading lepton  $p_T$  thresholds for  $\text{VBF } H \rightarrow \tau\tau$  with both  $\tau$  leptons decaying to  $e$  or  $\mu$  leptons and for the same compressed spectrum  $\text{SUSY}$  model, but with a more difficult mass splitting. The  $\text{SUSY}$  model acceptance is relative to a preselection of two leptons with  $p_T > 3$  GeV and  $E_T^{\text{miss}} > 50$  GeV that is used for simulation in the current (Run 2) analysis. A dilepton trigger with a threshold of 10 GeV on each lepton gives a substantial gain in acceptance in the  $\text{SUSY}$  model. The relative acceptance gains compared to a single-lepton trigger for the  $\text{SUSY}$  model is 39% while for  $\text{VBF } H \rightarrow \tau\tau$  it is only about 7%<sup>1</sup>. Including a moderately low dilepton threshold will improve the sensitivity to the broad, interesting class of signatures with small mass splittings similar to the  $\text{SUSY}$  model shown. Figure 2.4 also shows that in the no-upgrade scenario, a dilepton trigger does not have a high acceptance for these channels and will not effectively recover acceptance loss of the high single lepton trigger threshold. Finally, without the upgrade, the muon efficiencies would be significantly reduced, which has a large effect on the dimuon triggers because they are impacted by the square of the efficiency.

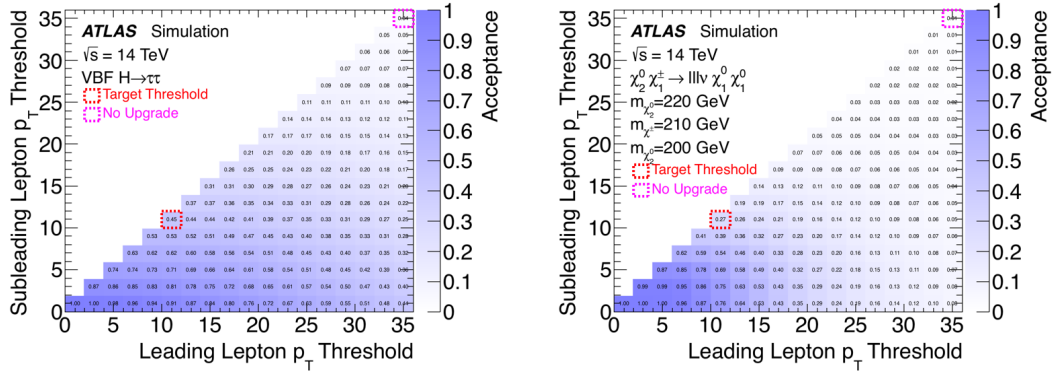
## 2.3 Physics Signatures with Photons

Studies of processes involving  $H \rightarrow \gamma\gamma$  will enter a new regime of precision in the  $\text{HL-LHC}$ . These processes will be a workhorse of precision inclusive and differential cross-section measurements for the various production modes as well as from some beyond the Standard Model processes involving a Higgs boson. Photon triggers, and in particular di-photon triggers, are essential to performing these measurements. Additionally, the  $HH \rightarrow b\bar{b}\gamma\gamma$  process is one of the key contributing channels for the overall  $HH$  sensitivity, which is a key goal of the  $\text{HL-LHC}$  programme. Figure 2.5 shows the acceptance for  $\text{VBF } H \rightarrow \gamma\gamma$  and  $HH \rightarrow b\bar{b}\gamma\gamma$  as a function of the subleading photon  $p_T$  threshold. The target threshold of 25 GeV preserves 22% more acceptance than the no-upgrade scenario with a 35 GeV threshold.

Single-photon triggers also play a significant role through **Initial State Radiation (ISR)**, such as in searches for low-mass di-jet resonances, as seen in Fig. 2.6. In particular, low-mass di-jet resonances are more accessible with photon **ISR** selections than with jet **ISR** selections. Fig. 2.7 shows the di-jet invariant mass distributions for low-mass resonance search in partial Run 2 data [2.3], for the photon **ISR** + di-jet signature (left) and jet **ISR** + di-jet signature

<sup>1</sup> One can get these numbers by comparing the (20 GeV,10 GeV) point to the (10 GeV,10 GeV) point. For example, for the  $\text{SUSY}$  model, 16% of events pass the (20 GeV,10 GeV) selection and 27% pass the (10 GeV,10 GeV) selection, which means that 11% acceptance is added. That can then be compared to the (20 GeV,0 GeV) which corresponds to the single lepton trigger and has an acceptance of 29%, giving a relative gain of  $11\%/28\% = 39\%$ .





(a)  $VBF H \rightarrow \tau\tau$  where both  $\tau$  leptons decaying (b) Compressed Spectra  $SUSY$ . The acceptance at to  $e$  or  $\mu$ . The acceptance at the target is 45% and the target is 27% and the acceptance in the no-upgrade scenario is 4%. upgrade scenario is 1%.

Figure 2.4: Acceptance for  $VBF H \rightarrow \tau\tau$  and compressed spectra  $SUSY$  for dilepton triggers. The  $SUSY$  channel is  $\chi_2^0 \chi_1^\pm \rightarrow ll\nu \chi_1^0 \chi_1^0$  where the masses are  $m_{\chi_2^0} = 220$  GeV,  $m_{\chi_1^\pm} = 210$  GeV,  $m_{\chi_1^0} = 200$  GeV. For the  $SUSY$  example, the acceptance is relative to a preselection of 2 leptons with  $p_T > 3$  GeV and  $E_T^{\text{miss}} > 50$  GeV.

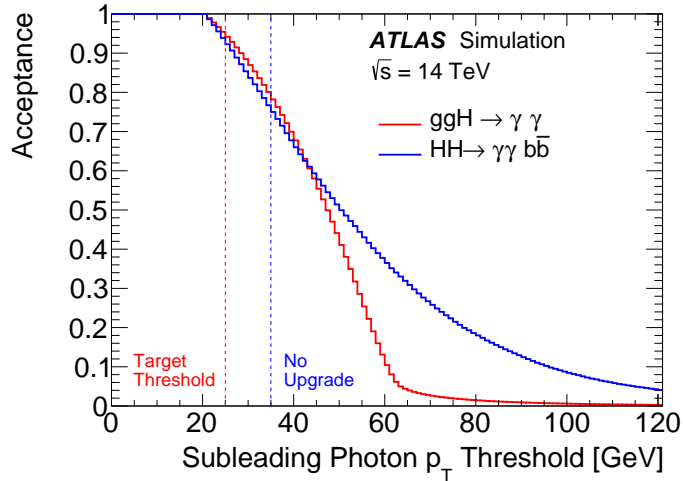


Figure 2.5: Acceptance for  $VBF H \rightarrow \gamma\gamma$  and  $HH \rightarrow b\bar{b}\gamma\gamma$  for di-photon triggers.

## 2.4 Physics Signatures with Hadronically Decaying Tau Leptons

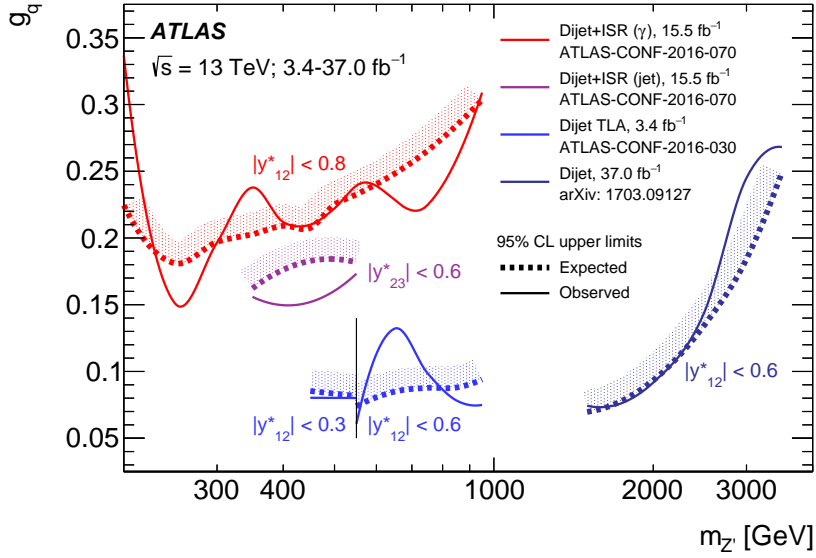


Figure 2.6: Summary plot of ATLAS bounds in the coupling-mediator mass plane from di-jet resonance searches using 2015 and 2016 data. Observed (solid lines) and expected (dotted lines) 95% CL upper limits on the coupling ( $g_q$ ) as a function of the resonance mass ( $m_{Z'}$ ) for a leptophobic  $Z'$  model [2.4]. The darker blue curves show results from a search trigger is associated with a jet from the candidate di-jet resonance [2.5]. The red and purple curves show results from a search triggering on an *ISR* photon or jet, respectively [2.3]. Finally, the lighter blue curves show results from a trigger-level analysis (*Trigger-object Level Analysis (TLA)*) [2.6].

(right). The mass range for the photon is wider than the mass range for jets, both because the trigger thresholds on the *ISR* photon are lower than those on the *ISR* jet and because in the photon *ISR* + di-jet case there is less ambiguity in the jet assignment as in the case when there are three jets where one is *ISR* and the other two are from the resonance. The two searches are complementary in the range where they overlap, due to the increased rate of jet *ISR* with respect to photon *ISR*. The sensitivity and mass range of both searches are driven by the trigger thresholds on photon and jet. More information on the motivation for low-coupling-constant di-jet resonance searches is given in Section 2.5. A low photon trigger threshold will benefit the reach of these analyses both in coupling constant and the minimum resonance mass accessible.

## 2.4 Physics Signatures with Hadronically Decaying Tau Leptons

With 64% of tau leptons decaying to hadrons, a dedicated tau trigger to discriminate between these tau decays and quark- or gluon-initiated jets provides important support to the ATLAS

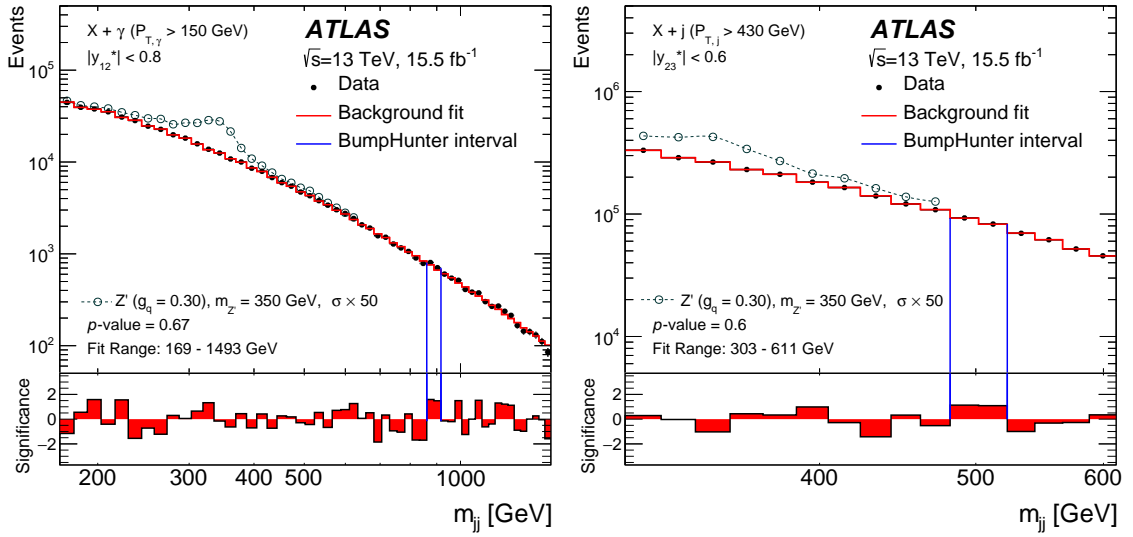


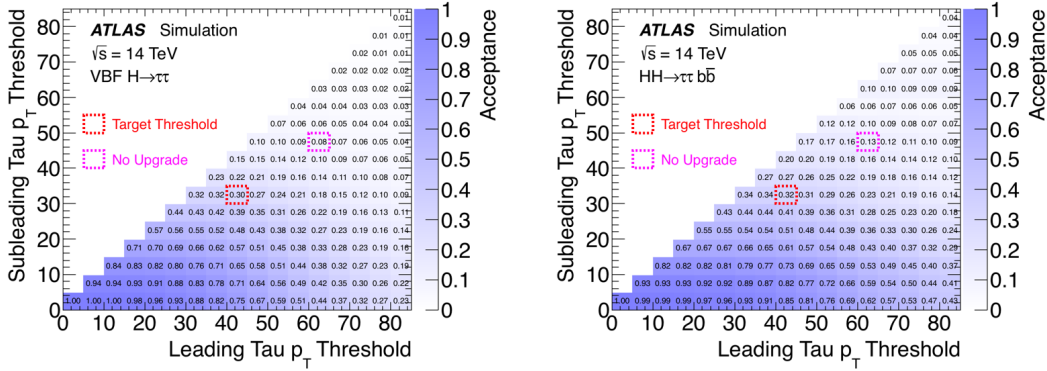
Figure 2.7: The reconstructed di-jet mass distribution (filled points) for events containing (left) a photon with  $p_T > 150$  GeV and two jets with  $p_T > 25$  GeV and (right) a jet with  $p_T > 430$  GeV and two additional jets with  $p_T > 25$  GeV. The open points in both plots show the expected distributions of a  $Z'$  signal. The photon-triggered sample has a much better mass resolution because of the lower trigger threshold of the photon trigger compared to the single-jet trigger.

physics programme. Additionally, these decays carry information about their parent particles, providing the ability to perform spin and CP measurements of heavy resonances that decay to taus. Within the Standard Model, the mass dependence of the Higgs boson couplings means that  $H \rightarrow \tau\tau$  is a process with a substantial rate and will be the most precisely measured leptonic branching fraction of the Higgs boson decay channels. In addition, the  $HH \rightarrow \tau\tau bb$  process will be a significant contribution to the measurement of non-resonant  $HH$  production. Figure 2.8 shows the acceptance for  $\text{VBF } H \rightarrow \tau\tau$  and  $HH \rightarrow \tau\tau bb$  where both tau leptons decay hadronically. Comparison to the no-upgrade scenario shows the substantial gain in acceptance for the achievable target threshold of leading tau lepton  $p_T > 40$  GeV and sub-leading tau lepton  $p_T > 30$  GeV; for  $\text{VBF } H \rightarrow \tau\tau$  the gain is a factor of  $\sim 3.7$  and for  $HH \rightarrow \tau\tau bb$  the gain is a factor of  $\sim 2.5$ . Further acceptance improvement for hadronic tau leptons could be made in the split hardware-level trigger system as described in Section 14.5.

## 2.5 Physics Signatures with Jets

Jet triggers are used in a variety of topics including multi-jet triggers for the  $HH \rightarrow 4b$  signature and  $\text{SUSY}$  particles decaying to two Higgs bosons and  $E_T^{\text{miss}}$ . Single jet triggers are relevant for di-jet resonance searches and searches where an  $\text{ISR}$  jet is used. As compared to the previously mentioned  $\sigma$  photon  $\text{ISR}$ , jets provide higher rates, but also require higher

## 2.5 Physics Signatures with Jets



(a)  $VBF H \rightarrow \tau\tau$  acceptance using  $\tau\tau$  triggers where both  $\tau$ s decay hadronically. The acceptance at the target is 30% and the acceptance in the no-upgrade scenario is 8% (b)  $HH \rightarrow \tau\tau b\bar{b}$  acceptance using  $\tau\tau$  triggers where both  $\tau$ s decay hadronically. The acceptance at the target is 32% and the acceptance in the no-upgrade scenario is 13%

Figure 2.8: Acceptance for  $VBF H \rightarrow \tau\tau$  and  $HH \rightarrow \tau\tau b\bar{b}$  for dihadronic tau ( $\tau\tau$ ) triggers where both  $\tau$ s decay hadronically.

$p_T$  selections. As shown in Fig. 2.6, the two methods are complementary; photon-triggered  $ISR$  events are sensitive at low  $Z'$  masses, while jet-triggered  $ISR$  events are more sensitive at higher  $Z'$  masses.

For the  $HH$  non-resonant process, a dedicated study [2.7] of the sensitivity as a function of the offline jets  $p_T$  threshold has been performed. The results are shown in Fig. 2.9 assuming that systematics are not a strong limitation on the result. This shows that the loss due to the trigger requirements below 50 – 60 GeV moderate (but non-negligible), and above  $\approx 65$  GeV the sensitivity in this important analysis degrades rapidly. The current estimated achievable threshold for the upgraded system is 65 GeV. A no-upgrade scenario would require a threshold of approximately 100 GeV, leading to a loss approximately a factor of two on the  $\sigma(HH \rightarrow 4b)/\sigma(HH \rightarrow 4b)_{SM}$  limit. Results with more pessimistic systematics assumptions show similar trigger impacts. For a more detailed discussion of the analysis and the impact of the system design on it, see Section 6.13. The Standard Model  $HH \rightarrow 4b$  non-resonant process cannot be observed with this channel alone, but the analysis would be sensitive to an enhanced cross-section due to anomalous Higgs self-couplings.

Di-jet resonance searches can be motivated by a variety of physics models. In particular, resonances with relatively small couplings to visible matter have been motivated by recent dark matter models [2.4][2.8]. Figure 2.6 shows a summary of the ATLAS bounds on the coupling  $g_q$  as a function of the  $Z'$  resonance mass. At high mass, above  $\sim 1.5$  TeV (dark blue line), the single jet trigger is used. For di-jet resonances with masses below  $\sim 1$  TeV, triggers are designed to select an  $ISR$  jet (purple line) or photon (red line, see also Section 2.3) present in the process rather than the decay products of the resonance. The jet  $p_T$  threshold affects the lower bound of the di-jet+ $ISR$  search with  $ISR$  jet. This is because

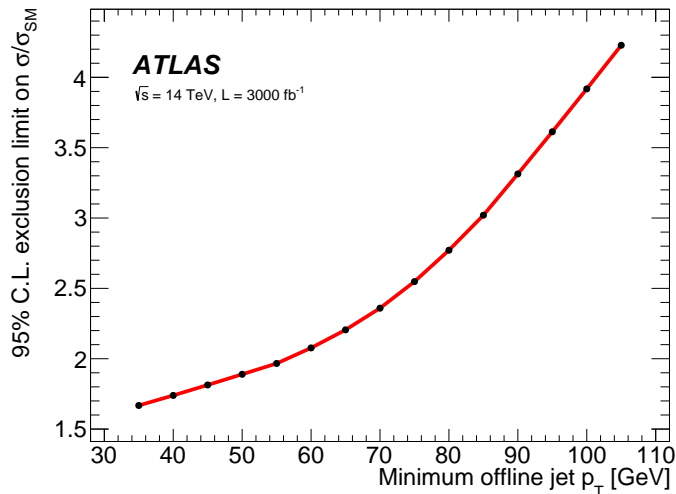


Figure 2.9: Expected 95% C.L. upper limit on the cross-section ratio  $\sigma(HH \rightarrow 4b)/\sigma(HH \rightarrow 4b)_{SM}$  as a function of the minimum  $p_T$  requirement applied to the fourth-leading jet, assuming that systematics are not a strong limitation on the result. As discussed in Section 2.2, modifications of the Higgs self-coupling can modify the cross-section by factors of order unity. Results with systematics show similar trigger impacts. For a more detailed discussion, see Section 6.13.

the trigger selection of a high- $p_T$  ISR jet biases the distribution of the invariant mass of the second and third leading jet for the QCD multijet background, which is problematic for the background estimation used in these searches.

The single-jet  $p_T$  threshold is driven by the Event Filter output limitation. The rate of high- $p_T$  single-jet events that can be accepted by the Level-0 and used for combinations with other signatures at the Event Filter is on the order of 25 kHz for HL-LHC; this corresponds to a  $p_T$  threshold of 180 GeV. Such a rate is much larger than the maximum recording rate of 10 kHz. An alternative is to reduce the amount of recorded data for these events by recording only the reconstructed objects in the Event Filter instead of the full detector data. In order for this to work, the reconstruction in the Event Filter needs to be as close as possible to a well understood offline reconstruction. This has been implemented in Run 2 and the result is shown by the light blue line in Fig. 2.6 (named TLA in the legend). It provides a substantial gain in sensitivity in the  $\sim 450 - 1000$  GeV  $Z'$  mass range (purple line in the figure). This illustrates the value of a low single-jet threshold and reconstruction in the Event Filter that closely follows the offline reconstruction (including having tracking available for pile-up mitigation and calibrations). The mass range of this search is ultimately limited by the Level-0 thresholds, and by the CPU requirements for obtaining tracks associated to trigger jets that can guarantee a good pile-up suppression performance.

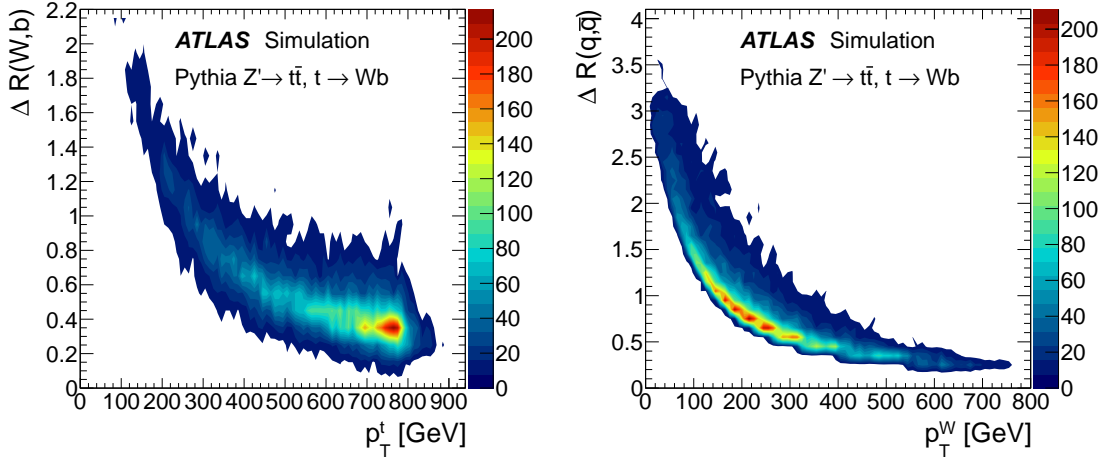


Figure 2.10: Separation between (a) the  $W$  and  $b$ -quark as a function of the top-quark  $p_T$  and (b) the two quarks from the  $W$  as a function of the  $W$   $p_T$  in  $Z' \rightarrow t\bar{t}$  events.

For boosted  $W$ ,  $Z$ , and  $H$  bosons and top quarks, the jets from the final state partons begin to merge into a single large jet which requires a dedicated trigger. Analyses address this by first reconstructing the entire heavy object as one large-radius jet reconstructed with the anti- $k_t$  algorithm with a radius parameter of  $R = 1.0$ . In order to efficiently trigger on these signatures, the trigger must similarly accept events with high- $p_T$  large-radius jets. Figure 2.10 shows the angular separation between the  $W$  and  $b$ -quark as a function of top quark  $p_T$  in  $Z'$  to  $t\bar{t}$  events and the separation between the two quarks from the  $W$  as a function of the  $W$   $p_T$  in same events. They have separations smaller than the standard  $R = 0.4$  jets radius around 300-400 GeV, at which point the multijet trigger will begin to lose efficiency; thus, 300-400 GeV is the appropriate target range for a large- $R$  jet trigger. The target threshold has been set at 300 GeV (see Section 6.8) while in the no-upgrade scenario, the threshold would be  $\approx 460$  GeV.

## 2.6 Physics Signatures with Missing Transverse Energy

Missing transverse momentum ( $E_T^{\text{miss}}$ ) is a critical trigger signature for Standard Model measurements (e.g.,  $ZH \rightarrow \nu\nu b\bar{b}$ ), searches for **SUSY**, and in searches for other new physics that result in events that do not interact with the detector (e.g. dark matter) This also includes processes such as “disappearing tracks” where a charged particle is produced and decays in the detector tracking volume. Figure 2.11 shows the acceptance as a function of the missing energy threshold for  $ZH \rightarrow \nu\nu b\bar{b}$  and a compressed **SUSY** scenario. In the compressed **SUSY** scenario, the momentum of the invisible particles (observed as  $E_T^{\text{miss}}$ ) is primarily due to the boost of the invisible system due to **ISR**. The  $E_T^{\text{miss}}$  distributions will therefore be similar for all models in which the  $E_T^{\text{miss}}$  is due to **ISR** with only a weak depend-

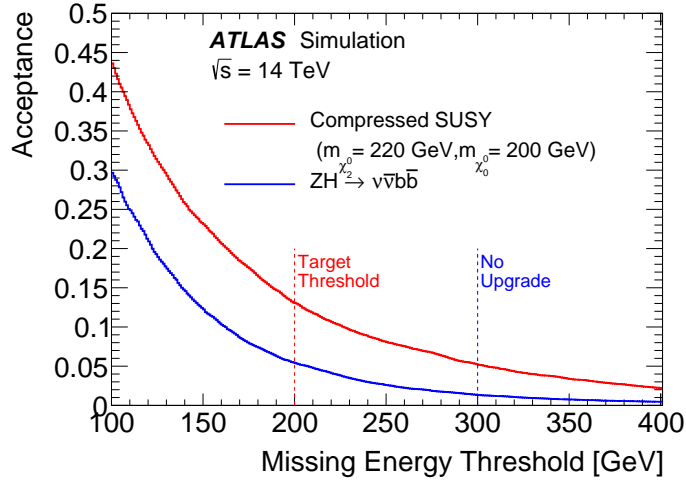


Figure 2.11: The integrated acceptance as a function of the missing transverse energy threshold for representative channels. A representative compressed *SUSY* model and  $ZH \rightarrow \nu b \bar{b}$  are shown.

ence on the mass of the invisible system. The typically achievable  $E_T^{\text{miss}}$  trigger threshold is  $\sim 200$  GeV for the current Run 2 trigger system. Operating the Phase-I system in Phase-II conditions with no-upgrade would result in a threshold of 300 GeV. Each 50 GeV reduction in threshold corresponds to a gain of approximately 50% – 100% in acceptance for the models shown. As discussed in Section 6.8, a target threshold of  $\sim 200$  GeV is achievable in the baseline upgrade scenario.

## 2.7 Physics Signatures with Forward Electrons

The capability for triggering on forward electrons adds acceptance to the trigger system for Phase-II. The acceptance for lepton triggers is generically enhanced, but the forward trigger in particular extends the coverage for high-rapidity electroweak bosons. Central (C) and forward (F) electrons are defined to be those with  $|\eta| < 2.5$  and  $|\eta| > 2.5$ , respectively. Candidate  $Z \rightarrow ee$  events can be classified as CC, CF, and FF. Figure 2.12 shows the distribution of Z boson rapidities accessible by the central trigger (CC and CF) and those only accessible with a forward trigger (FF). This extended reach in boson rapidity translates into an increased reach for determining the proton *Parton Distribution Function* (PDF). In particular, the large and small  $x$  regions are significantly impacted. The impact will be similar to the impact of recent LHCb [2.9] measurements on PDFs [2.10], but with significantly greater statistical power (and a systematic benefit of measuring the central and forward regions in the same experiment).



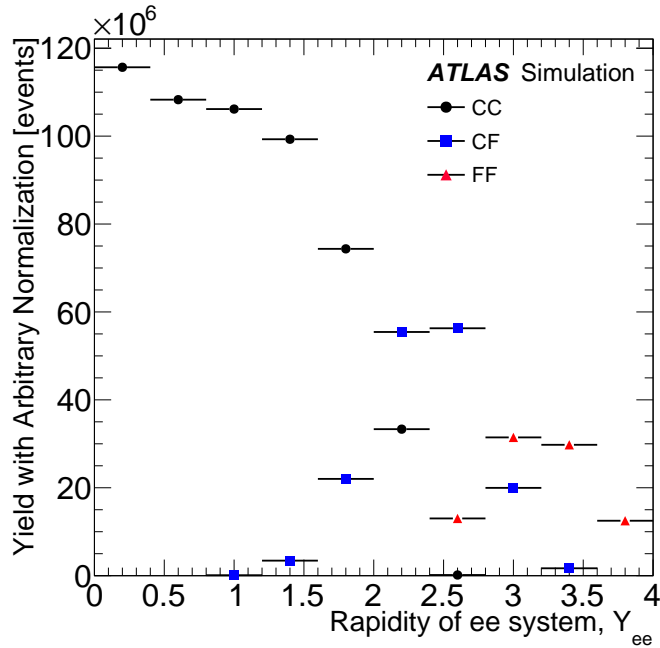


Figure 2.12: Unique  $Z$  boson rapidity ( $y_{m_{ee}}$ ) reach enabled by a forward electron trigger. The labels CC, CF, and FF stand for central-central, central-forward, and forward-forward pseudorapidity electrons, respectively. The  $y$ -axis is the yield with an arbitrary normalisation.

## 2.8 Physics Signatures with Exotic Objects

Some new physics scenarios produce signatures which are not identified by the standard set of reconstructed objects ( $e$ ,  $\mu$ ,  $\tau$ , jet,...). For example, long-lived particles (LLPs) might not have tracks associated with them, or might not point back to the beamline as prompt objects do. In addition, models such as  $\tau \rightarrow \mu\mu\mu$  and muon-jets resulting from Hidden Valley models [2.11] can result in close-by muons that are not identified with high efficiency. Figure 2.13, shows the opening angle between muons in a Hidden Valley model which is much smaller than the current system can resolve (approximately 0.2 in  $\Delta\phi(\mu, \mu)$ ). In the no-upgrade scenario, these can only be recorded by the single muon trigger (estimated to be 25 GeV, with the poor acceptance described in Section 3.2.2). In the upgraded scenario, a dedicated trigger is under development for a dimuon trigger with a  $p_T$  threshold of  $\approx 10$  GeV. The near-by muon conditions would be implemented in the new muon system Sector Logic, NSW Trigger Processor, and MDT trigger processor described in Chapter 8.

Dedicated signature-based triggers will be needed to address these exotic signatures, and such work is ongoing. In general, the system must be flexible enough to allow the development of a wide variety of dedicated triggers and accommodate new ideas introduced between now and the end of HL-LHC data-taking. In addition to supporting such possibil-



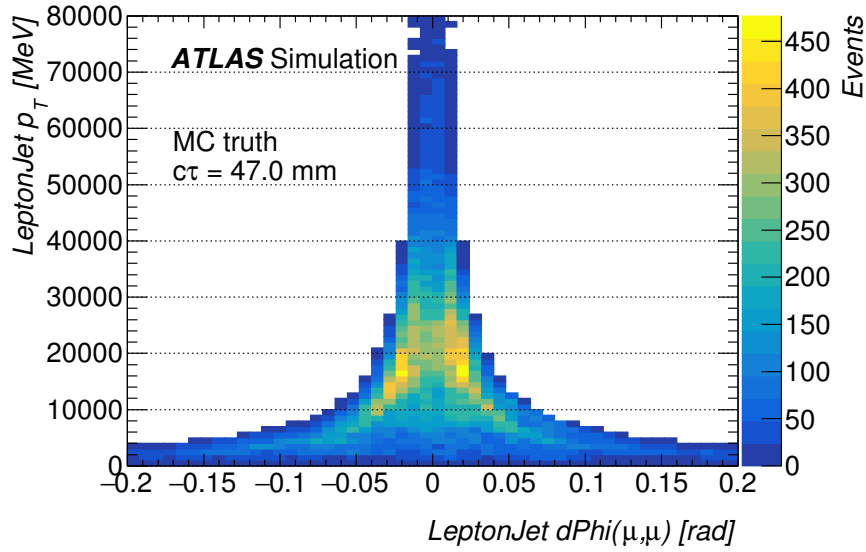


Figure 2.13: Example distribution of the muon momentum ( $y$ -axis) vs angular separation ( $x$ -axis) for muons in muon-jets resulting from Hidden Valley models [2.11]. The Phase-I system cannot find a second muon within  $\approx 0.2$  of another muon, which means that only single muon trigger with  $p_T > 25$  GeV can be used significantly reducing the acceptance.

ities, it is important to maintain inclusive triggers such as the single-object (jet and photon) triggers described above as well as  $E_T^{\text{miss}}$  and VBF inclusive triggers in order to maintain sensitivity to exotic physics signatures.

## 2.9 Physics with an Inclusive Vector Boson Fusion (VBF) Trigger

The VBF process plays an important role in Higgs boson physics and beyond. After the gluon-gluon fusion process it is the next largest Standard Model production cross-section. Using the full granularity in the forward calorimeters in the **ffEX** and the topological capability of the **Global Trigger** allows for the possibility of combining two jets with a large  $\eta$  separation to make an inclusive VBF trigger. This is particularly interesting for Higgs boson decays that are difficult to trigger on, e.g.  $H \rightarrow bb$  in the Standard Model. In addition, there are a wide variety of possible BSM Higgs decay modes summarised in [2.12]. Many of those modes, such as a single Higgs boson decaying to four  $b$ -quarks, are nearly impossible to trigger on directly, because of the low resulting jet energies. Another option is to trigger on the Higgs objects associated with the production mechanism as opposed to its decay products. The  $WH$  and  $ZH$  processes provide single leptons (and motivate a low single lepton threshold), but because they are lower cross-section than VBF and require the boson

to decay leptonically, a  $VBF$  trigger such as the one described in Section 6.10 is a promising option.

## 2.10 $B$ -Physics Signatures

The  $B$ -physics analyses search for new physics effects indirectly, namely through possible deviations from Standard Model predictions of precision measurements or by studying very rare processes. To cope with high pile-up and the relatively low  $p_T$  of the produced  $b$ -hadrons, the ATLAS  $B$ -physics programme focuses on fully reconstructable  $b$ -hadron decays with muons in the final state. In order to maximise signal yield, the muon trigger  $p_T$  thresholds need to be kept low. In Run 2 the lowest unprescaled di-muon trigger for  $B$ -physics requires two muons, each with  $p_T > 6$  GeV. An acceptable rate of this trigger is kept by additional topological selections at Level-1 (which is the first hardware level in Run 2), applying loose cuts on the di-muon system invariant mass and opening angle [2.13].

When no additional measures (as e.g. the Level-1 topological selections) are taken, the lowest unprescaled di-muon trigger in the upgrade menu requires two muons with  $p_T > 10$  GeV. This would lead to a loss of signal yields in flagship  $B$ -physics analyses as seen in Fig. 2.14. A dedicated di-muon trigger with topological selections is under investigation.

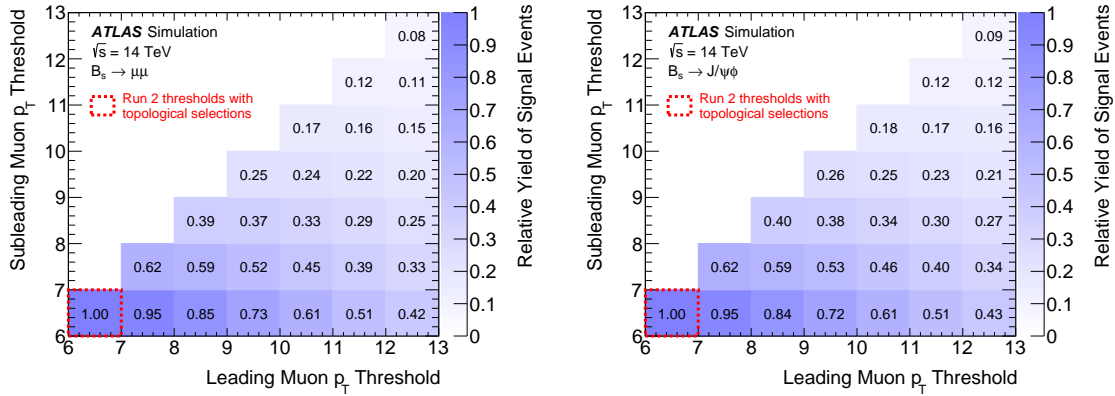


Figure 2.14: The relative yield of signal  $B$ -meson decays (left:  $B_s^0 \rightarrow \mu^+ \mu^-$ , right:  $B_s^0 \rightarrow J/\psi(\mu^+ \mu^-)\phi(K^+ K^-)$ ) as a function of the  $p_T$  thresholds applied on the muon candidate tracks. The yields are normalised to the number of signal events passing di-muon  $p_T$  trigger thresholds of 6 GeV on both the muon candidates. This particular configuration corresponds to the lowest unprescaled di-muon trigger in Run 2 used for  $B$ -physics analyses. Run 1/2 baseline offline cuts are applied to the reconstructed  $B$ -meson decays [2.14][2.15].

Another important aspect is the limited ability of the Run 2 Level-1 muon trigger to distinguish muons too close to one another in the momentum direction. The opening angle between the two muons in the  $b$ -decays depends on the  $b$ -hadron momenta and the invariant mass of the two muons. Figure 2.15 demonstrates this effect, showing signal events fraction under the condition that the two muons are separated either in  $|\Delta\eta(\mu^+\mu^-)| > 0.2$  or  $|\Delta\phi(\mu^+\mu^-)| > 0.2$  rad (typical Level-1 muon trigger granularity). This limitation becomes critical for  $B$  decays with small invariant mass di-muon pairs (namely  $B \rightarrow \mu^+\mu^-X$  and  $B \rightarrow J/\psi(\mu^+\mu^-)X$  decays). A study described in Section 6.9 is underway to improve the trigger acceptance for near-by muons.

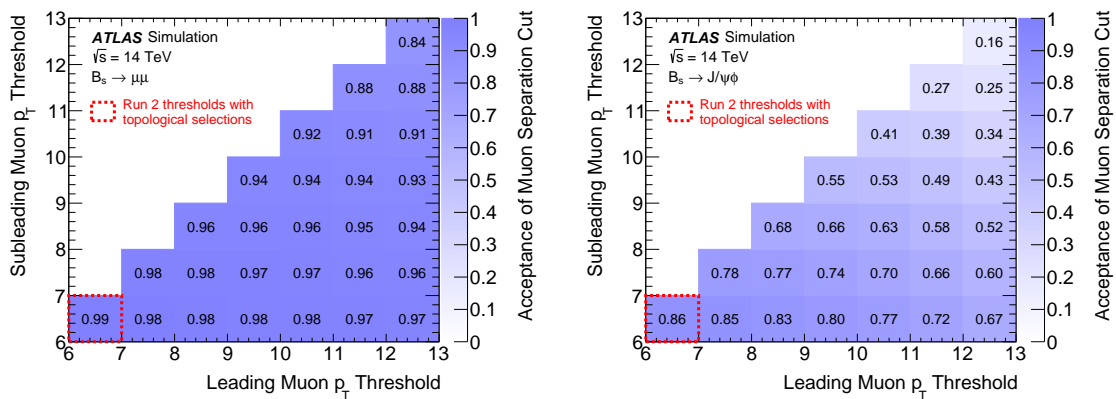


Figure 2.15: The fraction of signal  $B$ -meson decays (left:  $B_s^0 \rightarrow \mu^+\mu^-$ , right:  $B_s^0 \rightarrow J/\psi(\mu^+\mu^-)\phi(K^+K^-)$ ) with muons separated well enough for the Run 2 Level-1 trigger, as a function of the  $p_T$  thresholds applied on the muon candidate tracks. The separation condition used here is that the opening angle of the two muons is either  $|\Delta\eta(\mu^+\mu^-)| > 0.2$  or  $|\Delta\phi(\mu^+\mu^-)| > 0.2$  rad. Run 1/2 baseline offline cuts are applied to the reconstructed  $B$ -meson decays [2.14][2.15].

## 2.11 Physics Signatures for Heavy Ion Collision

The ATLAS HL-LHC programme will also include a study of Heavy-Ion (HI) collisions aimed at understanding the nature of a Quark-Gluon Plasma (QGP) and associated many-body QCD, which have unique triggering requirements. Studies of the HI performance are underway, but not yet complete. Described below is a qualitative summary of the selections used for HI programme. The system design is expected to provide similar performance gains to this programme as the  $pp$  programme, with significant improvements in the tracking-based triggers that are computing-limited in Run 2.

**Signatures with Jets** Jets, di-jets, multi-jets and  $b$ -tagged jets are used to study the interactions of colour charges in the QGP. The dominant issue for jet measurements in HI

collisions is the presence of a large amount of additional energy coming from the [Underlying Event \(UE\)](#). The properties of this energy depend on the impact parameter, or minimum distance, between the two colliding nuclei. Jet reconstruction in [HI](#) events needs to correct for the contributions from the large [UE](#). The proposed system allows for such corrections optimised for [HI](#) in the Level-0 trigger.

**Signatures with Colourless Probes: Electrons, Muons, and Photons** Prompt photons, colourless probes, are an important tool for the study of the [QGP](#). They probe the very initial stages of the collision because due to their colourless nature they are transparent to the subsequent evolution of the matter. The production rate of prompt photons is sensitive to the collision geometry and to modifications of the parton distribution functions of nucleons bound in a nucleus with respect to free nucleon. The complementary information is provided by measurement of electrons originating from decays of  $W$  and  $Z$  bosons. Furthermore, the photons or  $Z$  bosons provide a precise energy reference for the recoiling jet in  $\gamma$ -jet and  $Z$ -jet measurements. These measurements are essential to the understanding of interactions of colour charges with deconfined matter and are statistically limited with current datasets. Two-photon reactions or photo-nuclear reactions in ultra-peripheral heavy-ion collisions can be used to measure processes like light-by-light scattering that is potentially sensitive to [BSM](#) physics or nuclear [PDFs](#), respectively. The sophisticated algorithms used for photon, electron, and muon identification in  $pp$  collision are also well suited to these applications, and this physics benefits similarly from the upgrade.

**Signatures with Muons** In addition to the colourless probes discussed above, muon triggers are used to select events where vector mesons are produced. Particularly  $J/\Psi$  and  $Y$  mesons directly probe the Debye screening of colour charges in the [QGP](#). As the muon triggers are not affected by the [UE](#), no specific treatment of the muon triggers for the [HI](#) programme is needed, however a low-threshold dimuon trigger is of particular interest. The [L0Muon](#) system upgrades and the high [DAQ](#) rate will allow for strong performance.

### Track-based signatures for [HI](#) physics

**Ultra-peripheral Collisions** Vector mesons produced in diffractive ultra-peripheral collisions effectively image the geometric structure of the nucleus. As the different vector mesons have different binding energies and different effective “sizes”, repeating such a measurement with as many vector mesons as possible ( $\rho, J/\Psi, \dots$ ) would allow imaging the target with projectiles of various size. Furthermore, there is an open question of whether the  $J/\Psi$  meson can be produced perturbatively in reactions involving two-gluon exchange. To trigger on these, low activity events, based on Level-0 calorimetry, are processed with full-detector tracking.

---

**High multiplicity Events** An enhancement in the production of particles with small azimuthal-angle separation extending over a wide range of pseudorapidity differences, usually referred as “ridge”, has been observed in p+Pb. The data used in these measurements are selected using several high-multiplicity triggers. These rely on the a Level-0 requirement on the total transverse energy in the calorimeters and on the requirement on the multiplicity of reconstructed charged-particle tracks in the Event Filter. Full-detector hardware tracking will be a powerful tool for facilitating these triggers.

**Low  $p_T$  Heavy Flavours and  $J/\Psi$**  In decays of low- $p_T$   $J/\Psi$  decays one of the two muons may have insufficient momentum to reach the muon spectrometer and, as a result, is not detected. Those muons can be recovered by looking at their tracks in the inner detector, specifically full-detector tracking is again a powerful tool.

## 2.12 Summary of Requirements and Motivation for the Upgrade

The set of benchmark signatures discussed above leads to a list of thresholds and corresponding physics requirements needed to exploit the full potential of the ATLAS HL-LHC programme. Broadly, this includes maintaining low single-lepton, di-lepton, di-hadronic  $\tau$ , multi-jet, and missing transverse energy thresholds, as well as a complementary set of other triggers which complete the full programme. In particular, the ability to make selections not just with object momentum thresholds and multiplicity, but also the angular information (known as topological triggers) allows specific scenarios to be addressed. The following chapters describe the limitations of the Run 3 system in meeting those requirements, followed by the TDAQ Phase-II upgrade plans to address the requirements. Chapter 6 then describes the expected performance of the upgrades, including a example set of trigger object  $p_T$  thresholds (trigger “menu”) that achieve these requirements.

## References

- [2.1] N. Arkani-Hamed, A. Delgado, and G. F. Giudice, *The Well-tempered neutralino*, *Nucl. Phys.* **B741** (2006) 108–130, [arXiv:hep-ph/0601041](#).
- [2.2] R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, P. Torrielli, E. Vryonidou, and M. Zaro, *Higgs pair production at the LHC with NLO and parton-shower effects*, *Phys. Lett.* **B732** (2014) 142–149, [arXiv:1401.7340 \[hep-ph\]](#).
- [2.3] ATLAS Collaboration, *Search for new light resonances decaying to jet pairs and produced in association with a photon or a jet in proton-proton collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector*, ATLAS-CONF-2016-070, 2016.

## 2.12 Summary of Requirements and Motivation for the Upgrade

---

- [2.4] D. Abercrombie et al., *Dark Matter Benchmark Models for Early LHC Run 2 Searches: Report of the ATLAS/CMS Dark Matter Forum*, [arXiv:1507.00966](#) [hep-ex].
- [2.5] ATLAS Collaboration, *Search for new phenomena in dijet events using  $37\text{ fb}^{-1}$  of  $pp$  collision data collected at  $\sqrt{s}=13\text{ TeV}$  with the ATLAS detector*, *Phys. Rev. D* **96** (2017) no. 5, 052004, [arXiv:1703.09127](#) [hep-ex].
- [2.6] ATLAS Collaboration, *Search for light dijet resonances with the ATLAS detector using a Trigger-Level Analysis in LHC  $pp$  collisions at  $\sqrt{s}=13\text{ TeV}$* , ATLAS-CONF-2016-030, 2016.
- [2.7] ATLAS Collaboration, *Projected sensitivity to non-resonant Higgs boson pair production in the  $b\bar{b}b\bar{b}$  final state using proton–proton collisions at HL-LHC with the ATLAS detector*, Tech. Rep. ATL-PHYS-PUB-2016-024, CERN, Geneva, Oct, 2016. <https://cds.cern.ch/record/2221658>.
- [2.8] M. Chala, F. Kahlhoefer, M. McCullough, G. Nardini, and K. Schmidt-Hoberg, *Constraining Dark Sectors with Monojets and Dijets*, *JHEP* **07** (2015) 089, [arXiv:1503.05916](#) [hep-ph].
- [2.9] LHCb Collaboration, A. A. Alves, Jr. et al., *The LHCb Detector at the LHC*, *JINST* **3** (2008) S08005.
- [2.10] NNPDF Collaboration, R. D. Ball et al., *Parton distributions from high-precision collider data*, *Eur. Phys. J. C* **77** (2017) no. 10, 663, [arXiv:1706.00428](#) [hep-ph].
- [2.11] M. J. Strassler and K. M. Zurek, *Echoes of a hidden valley at hadron colliders*, *Phys. Lett. B* **651** (2007) 374–379, [arXiv:hep-ph/0604261](#).
- [2.12] D. Curtin et al., *Exotic decays of the 125 GeV Higgs boson*, *Phys. Rev. D* **90** (2014) no. 7, 075004, [arXiv:1312.4992](#) [hep-ph].
- [2.13] ATLAS Collaboration, *B-Physics topological L1 trigger*, <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/BPhysicsTriggerPublicResults>. 2017-11-16.
- [2.14] ATLAS Collaboration, *Study of the rare decays of  $B_s^0$  and  $B^0$  into muon pairs from data collected during the LHC Run 1 with the ATLAS detector*, *Eur. Phys. J. C* **76** (2016) no. 9, 513, [arXiv:1604.04263](#) [hep-ex].
- [2.15] ATLAS Collaboration, *Measurement of the CP-violating phase  $\phi_s$  and the  $B_s^0$  meson decay width difference with  $B_s^0 \rightarrow J/\psi\phi$  decays in ATLAS*, *JHEP* **08** (2016) 147, [arXiv:1601.03297](#) [hep-ex].

### 3 Challenges and Limitations of the Run 3 TDAQ system

The pile-up conditions dictated by the expected design luminosity of the [HL-LHC](#) indicate a need for a factor of ten increase in trigger rates compared to those expected in Run 3. This increase in trigger rate exposes the following fundamental limitations in the Run 3 [TDAQ](#) system:

- The readout bandwidth is limited by the detector front-end electronics and the available radiation-hard technology at the time of their construction.
- The Level-1 trigger rate in the Run 3 system cannot be increased beyond the 100 kHz rate for which it was designed without an unacceptable increase in deadtime.
- The maximum latency of the Run 3 system is  $2.5 \mu\text{s}$ ; this latency is insufficient to implement more powerful selection algorithms in order to reduce the trigger rate. This limitation and the previous two will be ameliorated by the planned upgrades to the detector front-end systems described in [3.1, 2, 3, 4], and [3.5].
- The Run 3 muon trigger system will have limited muon acceptance and reduced efficiency due to chamber integrated charge limits in the barrel region and high background rates in the endcap region when operated in Run 4 conditions.
- The readout and dataflow components of the Run 3 [TDAQ](#) system do not have the flexibility or scalability required to cope with the factor of more than twenty increase in bandwidth requirement resulting from the increased event size and higher data rates.
- Finally, the extremely high pile-up environment presents a special challenge for charged particle tracking algorithms; the Run 3 hardware-based tracking system will not be able to maintain its high efficiency for track reconstruction at  $\langle \mu \rangle \simeq 200$ .

Thus, the Run 3 [TDAQ](#) system will not be able to cope with these high rates while keeping the same requirements on physics objects such as electrons, muons, tau leptons, and jets. As a result, without an upgrade of the [TDAQ](#) system, the Run 4 ATLAS physics programme will be significantly degraded.

The Phase-I [TDAQ](#) System was designed to maximise the performance of the system within tight constraints, as described above. These limitations are discussed in further detail in this chapter. First, a brief description of the [TDAQ](#) Run 3 system architecture is presented in Section 3.1. In the following three sections, the inherent limitations of each aspect of the Phase-I System (the Level-1 Trigger, [DAQ](#), and [High-Level Trigger \(HLT\)](#)) are described.



Discussions of the hardware components and strategies that will be reused for the Phase-II System are also included. These restrictions steer the definition of the requirements for the TDAQ architecture at HL-LHC, as described in Chapter 4.

## 3.1 Overview of the Phase-I Trigger and Readout Architecture

The ATLAS Run 3 TDAQ system [3.6] is designed to cope with a LHC luminosity of  $\mathcal{L} = 3 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$  and 80 interactions per bunch crossing provided on-detector electronics in the tracking detectors cope with it. The Phase-I upgrade was required to integrate adiabatically new elements of the trigger architecture, within the constraints and the limits of the existing detector systems, including an average trigger acceptance rate of 100 kHz at the output of the Level-1 trigger system, a maximum Level-1 latency of  $2.5 \mu\text{s}$ , and a 1 kHz rate at the output of the HLT.

The three main objectives of improvements in the Run 3 TDAQ system design are to:

- increase the background rejection for electron/photon trigger candidates through the upgrade of the electromagnetic calorimeter trigger electronics [3.7] and the corresponding trigger processors;
- improve muon reconstruction and the rejection of “fake” candidates in the muon endcap spectrometer by installing new precision and high efficiency detectors in the NSW [3.8]; and to
- make use of full-event hardware-based charged particle track reconstruction (Fast TracKer (FTK)) during event processing at the HLT CPU farms, e.g. boosting selectivity for certain signatures, as in events with  $b$ -jets or  $\tau$  leptons, and improving object isolation algorithms [3.9].

Run 3 TDAQ components are required, wherever possible, to be forward compatible with the expected conditions at the HL-LHC. Figure 3.1 is a high-level architectural diagram of the Run 3 TDAQ System. Event data move from the ATLAS on-detector electronics into the front-end buffers at the bunch crossing rate of 40 MHz.

The Level-1 trigger system, built using custom electronics, processes data from the calorimeters and the muon detectors, searching for signatures such as large electromagnetic energy deposits or high- $p_T$  muon tracks. These signatures, as well as global event quantities such as missing transverse energy, are combined to form multiplicities and flags indicating if topological or global event criteria have been satisfied. These multiplicities and flags are then combined in the CTP with reference to up to 512 programmable trigger menu items; the CTP box in Fig. 3.1 includes the local trigger processors and the TTC distribution system. The main new components for Run 3 are highlighted in the architecture diagram: the eFEX, jFEX and gFEX processors, the Level-1 Muon Trigger sector logic for the endcap, a NSW trigger processor for muon reconstruction in the endcap region, the inclusion of Tile Calorimeter information in the muon trigger, an improved MUCTPI capable of providing



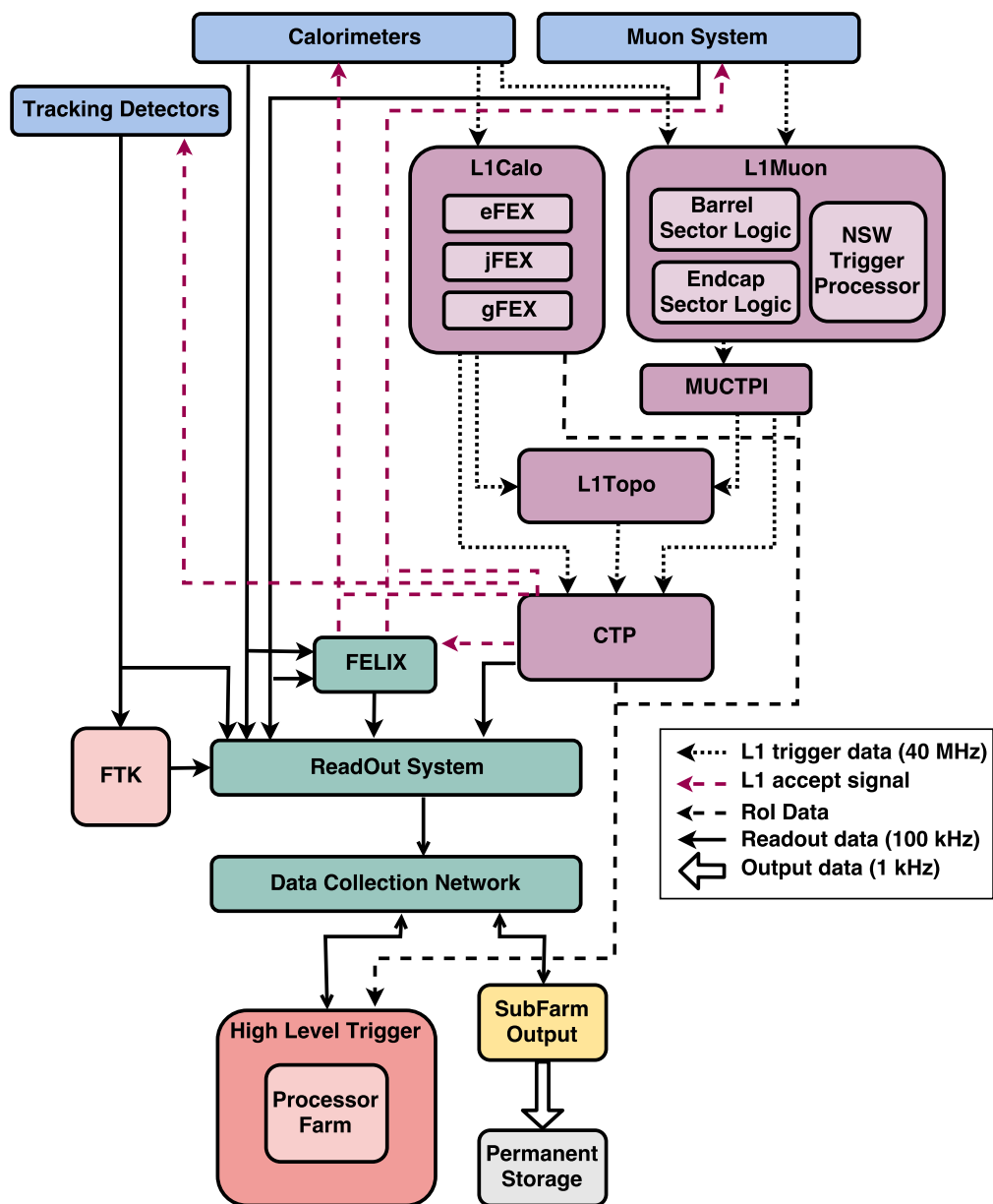


Figure 3.1: Schematic overview of the *TDAQ* system after the Phase-I upgrade, with an average trigger acceptance rate of 100 kHz at the output of the Level-1 trigger system and a maximum Level-1 latency of 2.5  $\mu$ s. The main objective of the system is to filter events and select up to approximately one thousand events per second for recording to permanent storage.

### 3.1 Overview of the Phase-I Trigger and Readout Architecture

---

precision muon inputs to the [Level-1 Topological Processor \(L1Topo\)](#), and an upgraded version of [L1Topo](#).

If the trigger conditions are met, a [Level-1 Trigger Accept \(L1A\)](#) is issued, initiating readout and processing in the [ReadOut Drivers \(RODs\)](#), and subsequent transfer to the [ReadOut System \(ROS\)](#). Readout data is transmitted from each of [L1Calo](#), [L1Muon](#), [MUCTPI](#), and [L1Topo](#); these lines are suppressed from Fig. 3.1 for simplicity. At the same time, [Regions of Interest \(RoIs\)](#) (small data packets describing event feature locations and energies) are built and sent to the [HLT](#) to steer further event processing.

After a [L1A](#) is issued, the [DAQ](#) system is responsible for the transport and assembly of the event data all the way from the sub-detector [RODs](#) to the logging to disk. While the readout path remains unchanged for most systems, the new and upgraded subdetector systems, such as the Muon [NSW](#), [L1Calo](#) and the [LAr](#) calorimeter, deploy a [FELIX](#)-based readout system, the first element of the [DAQ](#) system receiving front-end detector data and routing it using commodity multi-gigabit network to further processing. The data are passed and buffered in the [ROS](#), till requested by the [HLT](#) farm, where they are assembled into events (event building) and ultimately recorded to disk once accepted by the [HLT](#).

The [HLT](#) selection is implemented with software running on a large farm of commercial computer processors. The [HLT](#) executes chains of reconstruction and signature algorithms that analyse the properties of the events, running essentially offline algorithms to make a final decision within approximately one second. At this stage, a rejection factor of 100 is required, resulting in a recording rate of approximately 1 kHz. Wide-ranging software upgrades will be needed for core software, trigger menus and algorithms. Furthermore, the [HLT](#) system exploits the capabilities of the [FTK](#), which is designed to provide charged particle track reconstruction within 100  $\mu\text{s}$  of every event accepted by the Level-1 Trigger.

The trigger rates and trigger latency are critical parameters for the Phase-I [TDAQ](#) system design. These parameters are described in the following, and summarised in Table 3.1. The Level-1 trigger latency is defined as the time from the collision to the output of the corresponding [L1A](#) signal at the output of the [CTP](#). The Level-1 readout latency is defined as the time it takes for the [L1A](#) signal to reach the sub-detector front-ends. As no major changes in the Level-1 readout latency are possible during the Phase-I upgrade, because of constraints from detector front-end electronics that will not be replaced in the Phase-I upgrade, the Level-1 trigger latency must remain within a strict budget of 1.9  $\mu\text{s}$  and may generate at most a sustained 100 kHz rate of accepted events. Thus, the overall Level-1 trigger latency for the new Level-1 trigger components, such as the [L1Calo Feature EXtractors \(FEXs\)](#) and [L1Topo](#), are highly constrained.

While the Level-1 trigger latency is common to all trigger signals and is determined by the slowest one, the Level-1 readout latency depends on the particular sub-detector. Many of the new Run 3 components use high-speed digital optical links that are significant consumers of the latency budget. However, several ATLAS detector subsystems, including

Table 3.1: The main design parameters of the Run 3 TDAQ system, assuming a peak luminosity of  $\mathcal{L} = 3 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$  and 80 interactions per bunch crossing. Specialised calibration data is not taken into account.

Parameter	Run 3 Value
Level-1 trigger latency	$\simeq 1.9 \mu\text{s}$
Maximum Level-1 readout latency	$2.5 \mu\text{s}$
Level-1 trigger rate	100 kHz
Detector system readout deadtime (1.9 $\mu\text{s}$ latency)	$< 2\%$
Number of detector readout data sources	$\sim 100$
Bandwidth of detector ROLs	1.28-9.6 Gb/s
Average readout event size at peak luminosity	$\simeq 2.9 \text{ MB}$
ROLs bandwidth at peak luminosity	$\sim 290 \text{ GB/s}$
Maximum readout fraction of the ROS	50%
Average data collection bandwidth at peak luminosity	$\simeq 25 \text{ GB/s}$
Average output rate at peak luminosity	1.5 kHz
Number of data loggers	$\sim 10$
Average output event size (raw) at peak luminosity	$\simeq 3.5 \text{ MB}$
Average output bandwidth at peak luminosity	$\simeq 3.2 \text{ GB/s}$

large parts of the Inner Detector and the LAr Calorimeter, continue to use the same front-end and readout electronics up to the end of Run 3.

The Level-1 trigger rate was specified to be 100 kHz in Run 2, and is maintained at that rate in Run 3. The detector system readout deadtime is also a Run 2 parameter, which was obtained through the introduction of a complex deadtime algorithm customised to each detector's needs. Overcoming both the maximum trigger rate and deadtime limitations can only be accomplished by new front-end requirements and implementations.

The number of detector readout data sources includes the legacy readout links and the new sources from the FELIX systems; the bandwidth of the detector ReadOut Links (ROLs) is also specified in Table 3.1. The average readout event size includes the FELIX systems of the NSW and LAr Calorimeter, but not all L1Calo data fragments. Finally, the ROS maximum readout fraction is an internal limitation of the ROS, from its internal data paths and processing power.

The bandwidth of data fragments received by the HLT is given by the average data collection bandwidth parameter. The average output event size is larger than the average readout size due to the additional information provided by the HLT. Ten data loggers are estimated to transfer data from the HLT to permanent storage. Finally, the average output bandwidth is given, assuming a compression factor of 60% on the output event fragments.

### 3.2 Features and Limitations of the Run 3 Level-1 Trigger System

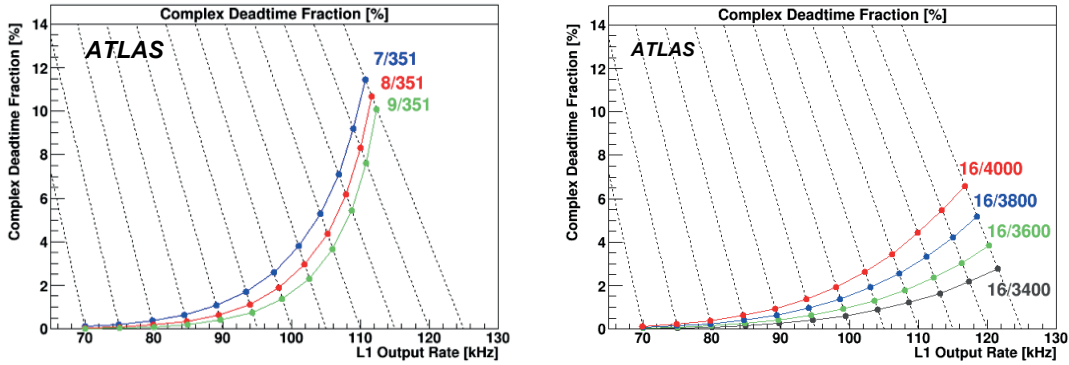


Figure 3.2: The complex deadtime fraction as a function of the Level-1 output rate for the *LAr* calorimeter (left) and *Semiconductor Tracker* (SCT) (right), at the current latency of  $\approx 1910$  ns. The different curves, created with simulation and cross-checked with data, show variations of the “Leaky Bucket” (“Sliding Window”) algorithm parameters implemented for *LAr* (SCT). A very steep increase in deadtime can be observed when the Level-1 output rate increases above 100 kHz.

### 3.2 Features and Limitations of the Run 3 Level-1 Trigger System

The main limitations of the Phase-I Level-1 Trigger system are as follows. The first limitation results from the front-end detector electronics. These electronics systems were designed with the best available radiation-hard technology at the time of their construction, but the 100 kHz trigger rate for which it was designed presents a fundamental limitation to the *DAQ* system (as described in Section 3.3) as well as a fundamental limitation to the Level-1 Trigger System. An appreciable increase in this rate would result in an unacceptable increase in deadtime, or the fraction of the data acquisition time in which no events can be recorded, as shown in Fig. 3.2. The detector electronics also present the limitation on the maximum latency:  $2.5 \mu\text{s}$ . Without a higher rate or longer latency, the Level-1 Trigger algorithms cannot be improved enough to cope with *HL-LHC* conditions.

The planned Phase-II upgrades to the detector electronics allow the design of the Phase-II *TDAQ* system to overcome the remaining limitations from the Phase-I system. These include the limited granularity from the calorimeters, the muon trigger acceptance in the barrel region, and the low background rejection of the muon trigger in the endcap regions.

#### 3.2.1 Level-1 Calorimeter Trigger Limitations

The Phase-I *FEXs* are designed to match the Phase-I calorimeter trigger granularity and, thus, they do not have the available input fibres to handle finer granularity calorimeter information. They are therefore limiting the redefinition of the calorimeter-based trigger

object selection in order to improve the trigger object identification and background rejection at  $\langle\mu\rangle \simeq 200$ . However, the coarser granularity algorithms will be utilised in Phase-II as pre-processors to the fine granularity algorithms anticipated. The general purpose of each Phase-I **FEX** will remain intact:  $e/\gamma$  and  $\tau$  triggers will be handled by the **eFEX** subsystem, single-jet triggers by the **jFEX** subsystem, and large- $R$  (or multi-jet) triggers plus the calculation of global quantities will be handled by the **gFEX** subsystem.

### Electron and Photon Triggers

There are two key elements to an  $e/\gamma$  trigger: efficient cluster formation and strong jet rejection. A cluster algorithm, which is large enough in area to contain the energy of an electromagnetic shower, but small compared with a typical jet in order to reduce the background rate, provides good  $e/\gamma$   $p_T$  resolution and a sharp efficiency curve. In addition, additional discriminating variables, e.g. based on shower shape and activity around the cluster core, are necessary at low  $p_T$  for additional rejection against jets and reduce the overall trigger rates. The clustering algorithm used in the **eFEX** approximates the clustering used in the ATLAS **HLT** and offline, but is based on **super cell** calorimeter granularity that provides information for each calorimeter layer with granularity up to  $\Delta\eta \times \Delta\phi = 0.025 \times 0.1$  segmentation in the front and middle layers. The basic steps of the algorithms developed for Run 3 are:

- Electromagnetic clusters are built around a “seed” - a **super cell** in the middle layer of the electromagnetic calorimeter, where most of the shower energy is deposited and which is more energetic than any of the surrounding **super cells**.
- A cluster is formed by summing this cell with the more energetic of its neighbours in the  $\phi$  direction, then summing both neighbours in the  $\eta$  direction, to form a cluster of  $3 \times 2$  **super cells**.
- The corresponding cells from the front later are added, plus the pair of cells in the presampler and three corresponding to the towers containing the seed and its neighbour in  $\phi$ .

The main part of the cluster slides in steps of 0.025 in  $\eta$ , providing good containment of the showers wherever they are within a tower, and resulting in a turn-on curve with less dependence on the position of the shower within the tower than the Run 2 trigger. Without any additional vetoes, the Run 3 system produces a  $\sim 20\%$  lower rate for the same efficiency of the Run 2 system.

The finer granularity of the **super cell** and the depth information of the calorimeters, enables the **eFEX** firmware to perform a sophisticated rejection of jet backgrounds through the use of shower-shape variables. The most discriminating of these is defined as follows. Given a  $3 \times 2$  group of Super Cells in  $\eta \times \phi$  centred on the highest-energy Super Cell in the middle

### 3.2 Features and Limitations of the Run 3 Level-1 Trigger System

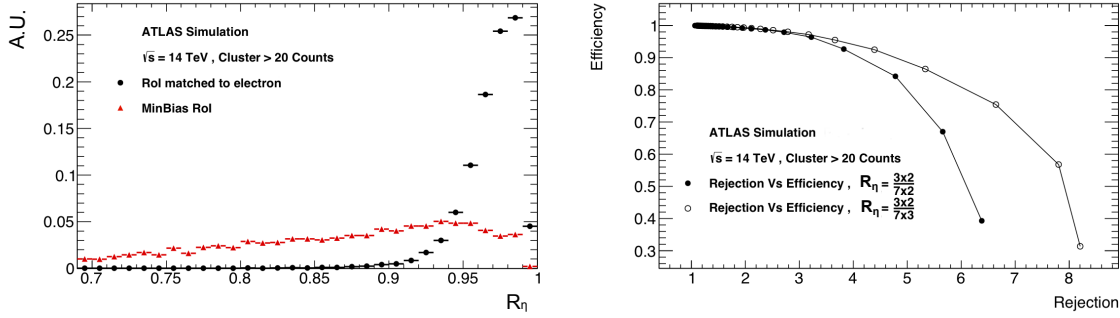


Figure 3.3:  $R_\eta$  distributions for clusters matched to reconstructed electrons and for clusters found in minimum-bias events with Run 3 conditions ( $\langle \mu \rangle = 80$ ). (b) Electron efficiency vs rejection of clusters in minimum bias events for two different  $R_\eta$  variables. All clusters were required to pass a  $E_T > 20$  GeV threshold.

layer,  $R_\eta$  is defined as the transverse energy measured in the  $3 \times 2$  group divided by the transverse energy measured in a  $7 \times 2$  group:

$$R_\eta = \frac{E_{T, \Delta\eta \times \Delta\phi=0.075 \times 0.2}^{(2)}}{E_{T, \Delta\eta \times \Delta\phi=0.175 \times 0.2}^{(2)}}. \quad (3.1)$$

The distribution of  $R_\eta$  for clusters matched to reconstructed electrons compared to clusters found in minimum-bias events is shown in Fig. 3.3. The shower-shape variables implemented in the eFEX in Phase-I are sufficient to provide the background rejection factors required for Run 3 operating conditions, but the super cell granularity is a limiting factor for increasing the background rejection further in Run 4. The eFEXs can only consider a  $0.3 \times 0.3$  algorithm window size; this narrow field of view may limit the use of offline-like algorithms. However, the electron and photon trigger candidates found by the eFEX can be utilised as a starting point for further refinement with full calorimeter granularity in the Phase-II trigger system.

#### Forward Electron Triggers

The availability of tracking information at larger  $\eta$  for the Phase-II ITk introduces many advantages in terms of physics objects reconstruction, and in terms of pile-up mitigation by linking objects to the primary vertex corresponding to the hard-scatter of interest. Since the ITk extends tracking up to  $|\eta| < 4.0$ , the geometrical acceptance of the Phase-I Level-1 Trigger system is a significant limiting factor for the ATLAS discovery potential. The Phase-I eFEX system, dedicated to  $e/\gamma$  identification, covers  $|\eta| < 2.5$ , the coverage of the original ATLAS inner detectors. Without an upgrade, there is no recourse for identifying electrons in the forward regions.

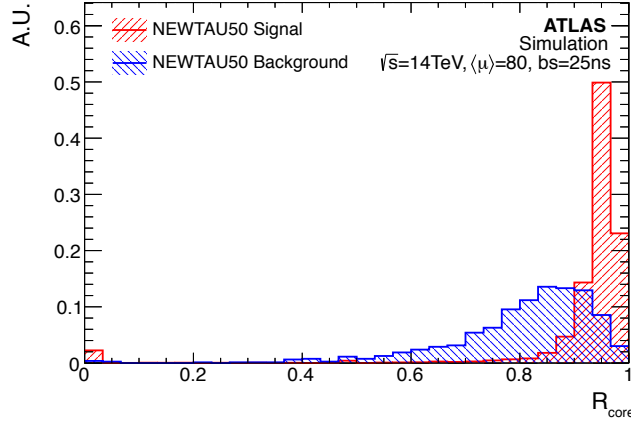


Figure 3.4:  $R_{\text{core}}$  distributions for truth-matched tau clusters and clusters from minimum-bias background events.

### $\tau$ -lepton Triggers

Hadronic  $\tau$  triggering essentially involves the selection of narrow mixed electromagnetic and hadronic clusters and separating these from the jet background. It is particularly challenging since hadronic tau decays include a range of different states, are less distinctive than electrons, and also typically have a lower observable  $E_T$  due to the neutrinos present in the decays. The Run 3 approach used in triggering on  $\tau$  leptons is to use **super cell electromagnetic (EM)** calorimeter granularity information in forming clusters and jet discriminants in the **eFEX**.

The **eFEX** implementation of  $\tau$  identification is based on a cluster calculated in a region of  $5 \times 2$  **super cells** in the middle layer of the **EM** calorimeter ( $0.125 \times 0.2$ ), 3 in the front layer,  $0.1 \times 0.2$  in the presampler, and  $0.2 \times 0.2$  clusters in the back layer and in the hadronic calorimeter. In addition, a  $3 \times 3$  tower cluster ( $0.3 \times 0.3$ , **EM+hadronic**) has been studied using towers built from **super cell**. Jet rejection relies on the narrower width of the  $\tau$  cluster compared to a jet. This can be exploited via isolation, such as the  $R_{\text{core}}$  variable defined as:

$$R_{\text{core}} = \frac{E_{T, \Delta\eta \times \Delta\phi=0.125 \times 0.2}^{(2)}}{E_{T, \Delta\eta \times \Delta\phi=0.225 \times 0.3}^{(2)}}. \quad (3.2)$$

The resulting distribution, comparing truth-matched  $\tau$  clusters to clusters from minimum-bias events, is shown in Fig. 3.4.

This algorithm is effective for identifying narrow showers resulting from a hadronic  $\tau$  lepton decay using **super cell** granularity. The **jFEX** could also be used to search for tau leptons, implementing an algorithm that considers a larger  $\eta - \phi$  area, but without the **super cell** granularity. Neither the narrow window of consideration of the **eFEX** nor the larger window, but coarse granularity of the **jFEX** are optimal for triggering on hadronic taus, and



thus are a limitation of the Run 3 system. However, either option could be used to seed a fine-granularity, large-area hadronic tau trigger identification algorithm in Run 4.

### Single-jet Triggers

In the Phase-I Level-1 Calorimeter Trigger System, the **jFEX** is responsible for identifying isolated  $R = 0.4$  jets for the single-jet trigger. To accomplish this, the inputs are towers of  $\sim 0.1 \times 0.1$  in the region  $|\eta| < 2.5$  and  $0.2 \times 0.2$  between  $2.5 < |\eta| < 3.2$ . In the **FCal** region the  $\phi$  granularity is  $\sim 0.4$ , while the  $\eta$  granularity varies between 0.1 and 0.4. The  $E_T$  in each tower is summed over all of the calorimeter layers.

Iterative algorithms used in the **HLT** and offline jet reconstruction are not suitable in the **jFEX** because of the limited and fixed latency. Level-1 jet trigger algorithms therefore must be executed in parallel within a set of overlapping window environments, and cannot use information from outside those environments. The size of the environment is limited by the available bandwidth and the necessary interconnectivity to share data between **FPGAs**, modules and crates. In the baseline design the environment is  $9 \times 9$  towers, which corresponds to an area slightly larger than a  $\Delta R = 0.4$  cone, though options that would allow this to be increased are being investigated.

The baseline jet-finding technique for the **jFEX** is a simple circular sliding window algorithm, where towers with energy above  $4\sigma$  are considered. The performance of the single-jet trigger algorithm implemented in the **jFEX** is shown in Fig. 3.5. Though there is additional room for improvement between the **jFEX** algorithm and the offline reconstruction, the algorithm does not provide an inherent limitation for isolated single-jet identification. Further improvements in the trigger performance may be obtained by a so-called "Gaussian filter" algorithm, where towers are summed with a weight that decreases as a Gaussian function of the distance from the centre of the cluster. Gaussian jet algorithms may have different responses for quark and gluon jets, etc. The key parameter in determining the performance is the value used for  $\sigma$  (the Gaussian width) large values (0.3 – 0.4) give good inclusive jet performance, while small values ( $\sim 0.1$ ) are good for identifying nearby jets, and perhaps for jet substructure. This option is under investigation. Furthermore, once the tight latency window required in the Phase-I system is relaxed, the **jFEX** may also include an offline-like jet calibration.

### Multi-jet and Boosted Object Triggers

The **gFEX** is intended to enhance the selectivity of the Level-1 trigger and increase sensitivity to key physics channels. The **gFEX** identifies large-radius jets, typical of Lorentz-boosted objects, by means of wide-area jet algorithms refined with subjet information. The architecture of the **gFEX** permits event-by-event local pile-up suppression for these jets using baseline subtraction techniques comparable to those developed for offline analyses.



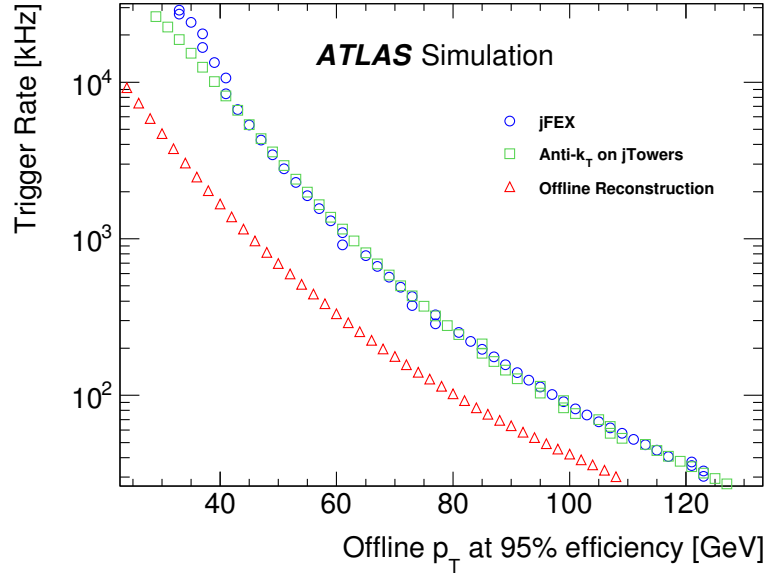


Figure 3.5: Trigger rate for a 95% efficiency on the leading jet vs. offline  $p_T$  thresholds for jets reconstructed in the *jFEX* ( $|\eta| < 3.2$ ). The efficiency is evaluated using  $HH \rightarrow bbbb$  signal events, and the trigger rate is evaluated based on minimum bias background events at  $\langle \mu \rangle \simeq 200$  and correspond to a luminosity of  $\mathcal{L} = 5 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ . The *jFEX* algorithm, the offline anti- $k_t$  algorithm (run over  $\eta \times \phi = 0.1 \times 0.1$  towers), and the full offline reconstruction are compared.

The *gFEX* architecture is also suitable for other global event algorithms such as  $E_T^{\text{miss}}$  and centrality-related observables.

A key feature of the *gFEX* is that the entire calorimeter is available in a single module, which enables the use of algorithms that can scan the entire  $\eta$  range of the calorimeter. However, the granularity available is  $\eta \times \phi = 0.2 \times 0.2$  towers summed over all of the calorimeter layers. One of these full-scan algorithms involves the identification of boosted hadronic topologies that are characteristic of new physics scenarios. At high transverse momentum, the hadronic decay products of energetic bosons and top quarks tend to be highly collimated. In the *gFEX*, large-radius jets can be used to capture the boson or top quark decay products, whereas in Run 2, algorithms identify significant local energy deposits in a limited region of interest. Once these large-radius jets are identified, exploiting the internal structure of these jets allows for differentiation of multi-prong objects from the dominant single-prong jet backgrounds. Thus the *gFEX* will allow the efficient identification of boosted objects while maintaining a suitable rate in the expected LHC environment after the Phase-I upgrade. Figure 3.6a illustrates the large- $R$  jet reconstruction by the *gFEX* in an event with top quark produced at high  $p_T$ . The jet-finding algorithm in the *gFEX* consists of building  $\eta \times \phi = 0.6 \times 0.6$  fully overlapping regions, as illustrated in Fig. 3.6b, in order to maintain high efficiency for multi-jet triggers.

### 3.2 Features and Limitations of the Run 3 Level-1 Trigger System

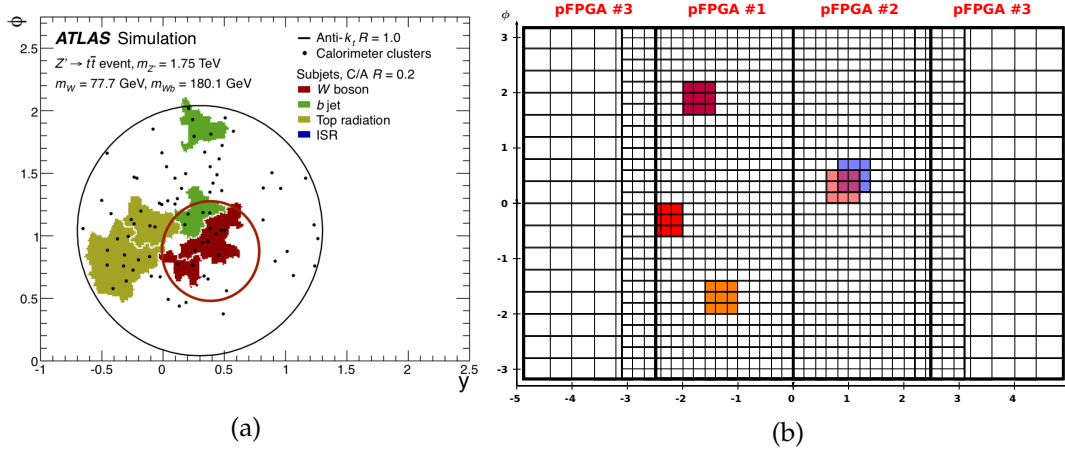


Figure 3.6: (a) Fully-simulated large-radius jet from a top quark produced at high transverse momentum. Subjets of a particular category have the same fill color and their extent represents the subjet active catchment area. Jet constituents are shown as black dots. The red circle represents a  $R = 0.4$  jet, whereas the black circle represents the triggers provided by the *gFEX*. (b) Illustration of the *gFEX* jet-finding algorithm, building multiple jets of size  $\eta \times \phi = 0.6 \times 0.6$ . Note that jets are allowed to overlap.

Models of new physics often result in final states with stable, invisible particles. These include many dark matter (DM) models, supersymmetry (SUSY) and beyond the Standard Model decays of the Higgs boson. Typically, these events are searched for by requiring a large momentum imbalance,  $E_T^{\text{miss}}$ . The  $E_T^{\text{miss}}$  in the detector is reconstructed using inputs of the visible momenta perpendicular to the beam axis. In the Level-1 trigger system, the  $E_T^{\text{miss}}$  is computed using only calorimeter information.

The *jFEX* has a significant inefficiency for event configurations with multiple jets where some of the jets are close together such as might be found in  $t\bar{t}$ , Higgs, or SUSY events. This has been one of the primary motivations to develop the *gFEX* system and it is illustrated in Fig. 3.7.

Finally, because of the granularity of the jet constituents at their inputs, the *jFEX* and *gFEX* systems cannot trigger efficiently on jets with  $p_T < 50$  GeV at HL-LHC. In general, in order to retain similar acceptance to offline for jets and hadronic tau decays higher granularity information is required as the pile-up increases so that cluster identification and separation is not diluted by the effects of pile-up on the fluctuations of the calorimeter energy depositions. Such intrinsic limitations may be overcome by using fine calorimeter granularity and the implementation of offline-like topological clusters and iterative jet reconstruction algorithms (e.g., anti- $k_t$ ). However, the *jFEX* and *gFEX* may be used to pre-process events of interest and restrict the area over which iterative jet-finding is performed.

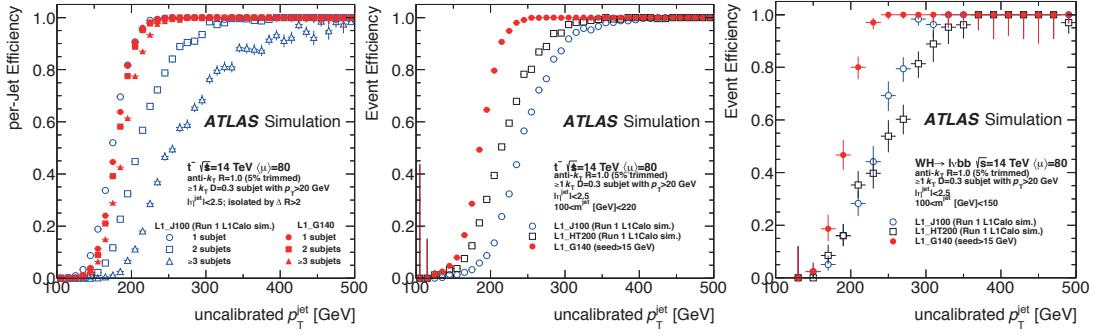


Figure 3.7: Sub-jet reconstruction trigger efficiency turn-on curves for  $t\bar{t}$  events are shown on the left plot as a function of uncalibrated jet  $p_T$ . The middle and right plots show the trigger efficiency curves for  $t\bar{t}$  and  $WH \rightarrow l v b \bar{b}$  events respectively. The red dots correspond to the results for a *gFEX*-like algorithm reconstructing fat jets with  $p_T > 140$  GeV while those corresponding to a *jFEX*-like algorithm reconstructing jets with  $p_T > 100$  GeV are shown by open blue dots.

### Forward Jet Triggers

The *L1Calo* Phase-I system is capable of reconstructing jets in the range  $|\eta| < 4.9$  but the granularities of both the *jFEX* and *gFEX* is extremely coarse in the forward region (as seen in Fig. 3.6b). Thus, no Phase-I Level-1 Trigger component is capable of detailed forward jet reconstruction that is needed to reject pile-up energy depositions in the forward region at the *HL-LHC*.

Furthermore, even the Phase-I *super cell* granularity is rather irregular in the forward region ( $3.2 < |\eta| < 4.9$ ) and particularly coarse in  $\phi$ . Therefore, forward electron identification, jet and missing-energy measurements may degrade significantly due to the excessive pile-up noise. Figure 3.8 illustrates qualitatively the concept by showing the energy flow in the first module of the *FCal* calorimeter by a single jet with  $p_T \sim 50$  GeV ( $E = 2.6$  TeV) at *HL-LHC* conditions either (a) using *super cell* information, or (b) the full granularity of the *FCal* detector.

Finally, Fig. 3.9 (see also Ref. [3.10]) illustrates the extent of the challenge of rejecting pile-up in the forward region. Even with full granularity and offline topological cluster reconstruction at low pile-up ( $\langle \mu \rangle = 30$ ), the topoclusters in the *FCal* tend to merge into a single cluster. A dedicated trigger solution for Phase-II is needed to optimise the reconstruction of forward jet trigger objects.

### 3.2 Features and Limitations of the Run 3 Level-1 Trigger System

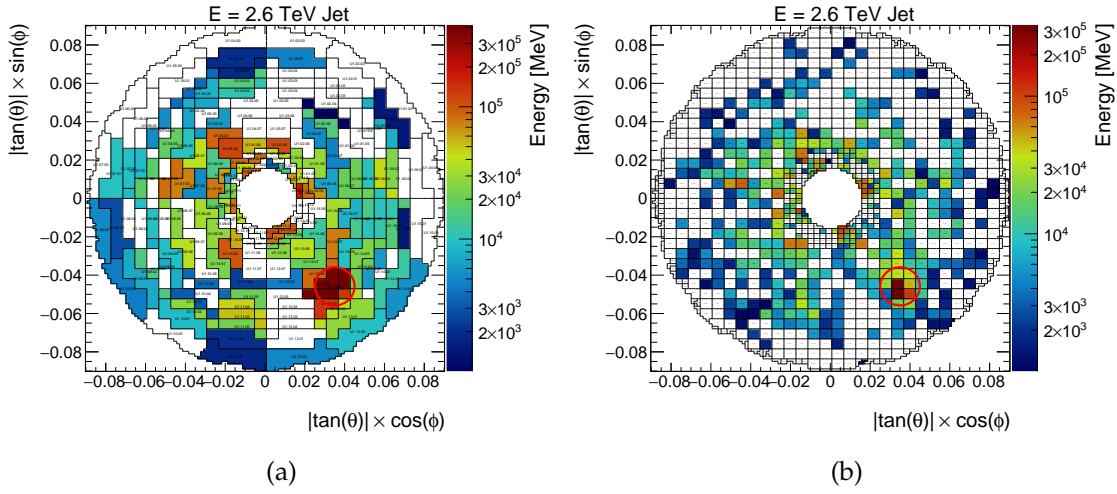


Figure 3.8: (a): Energy flow of a single jet (with  $E = 2.6$  TeV,  $p_T \sim 50$  GeV) in the EM layer of the FCal detector at HL-LHC ( $\mu \simeq 200$ ). The regions indicate the super cell readout granularity developed for the Phase-I system. (b): Energy flow for the same single jet, but the regions indicate the full granularity information of the FCal detector. The red circle in each figure indicates the truth jet position.

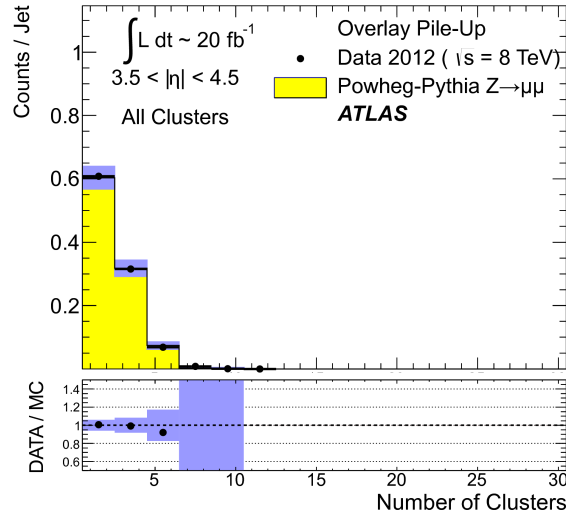


Figure 3.9: Number of topoclusters in the FCal for  $\langle \mu \rangle = 30$ , with pile-up generated from overlaid minimum bias data events.

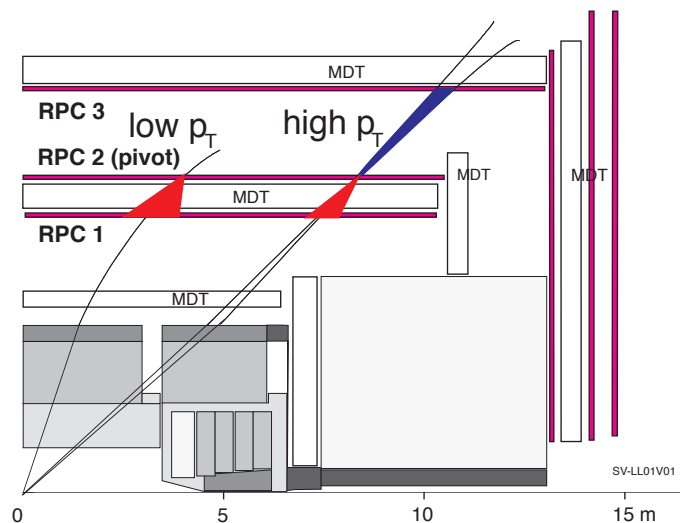


Figure 3.10: The Phase-I Level-1 muon trigger scheme in the barrel region ( $|\eta| < 1.05$ ).

### 3.2.2 Level-1 Muon Trigger Limitations

The ATLAS muon spectrometer has shown that it is able to achieve good muon identification and momentum resolution ( $\sim 10\%$  at 1 TeV). This performance will be maintained through the [HL-LHC](#).

The Phase-I Level-1 muon trigger is sufficiently selective to provide a rate that would take a large but not prohibitive fraction of the 100 kHz Level-1 trigger rate and a relatively small fraction of the proposed Phase-II 1 MHz Level-0 trigger rate. However, the original [RPC](#) and [TGC](#) systems are designed for maximum latencies of  $6.4 \mu\text{s}$  and  $3.2 \mu\text{s}$ , respectively, and for a maximum rate of 100 kHz. These limitations arise from the depth of the readout buffer to store the hit data before the arrival of the trigger accept signals and also from the readout bandwidth. The momentum resolution for low- $p_T$  muon also limits the system performance, since poorly-measured muons below the trigger threshold have an appreciable impact on the trigger rates. Furthermore, there are limitations on the muon acceptance and efficiency, as described below.

#### Level-1 Muon Triggers in the Barrel Region

The Muon trigger in the barrel region ( $|\eta| < 1.05$ ) is based on three trigger stations. Two stations are used for low- $p_T$  (6-10 GeV) muon triggers, while the third station is used in addition for high- $p_T$  (8-40 GeV) triggers as shown in Fig. 3.10.

## 3.2 Features and Limitations of the Run 3 Level-1 Trigger System

---

Each station is composed of two detector planes, and each detector plane is read out in two orthogonal projections,  $\eta$  and  $\phi$ , that will be referred to as the bending and non-bending<sup>1</sup> projections, respectively. The muon trigger  $p_T$  resolution is dominated by the information read out from the detectors in the bending projection. However, the information in the non-bending view helps to reduce the background trigger rate from noise hits in the chambers produced by low-energy photons, neutrons and charged particles, as well as localising the track candidates in space as required for the [HLT](#) trigger. In addition, the trigger chamber information in the non-bending view provides a second coordinate measurement for off-line reconstruction of muons (the precision chambers give information only in the bending projection).

The basic principle of the muon trigger identification algorithm in Phase-I is to require a coincidence of hits in the different chamber layers within a road. The width of the road is related to the  $p_T$  threshold to be applied. Space coincidences are required in both views, within a time window close to the bunch-crossing period (25 ns). The coincidence requirement allows for missing layers due to detector inefficiencies, dead regions, etc. For the low- $p_T$  trigger, hits are required within the road in at least three of the four layers, in each of the two projections. For the high- $p_T$  trigger, an additional requirement is made, demanding hits in at least one of the two layers in each of the two projections of the third station. A system of programmable coincidence logic allows concurrent operation with a total of six thresholds, three associated with the low- $p_T$  trigger and three associated with the high- $p_T$  trigger. Each of the six thresholds is independently programmable.

The muon efficiency for muons with  $p_T > 25$  GeV is only  $\sim 70\%$  in the barrel ( $|\eta| < 1.05$ ), as illustrated in the left plot of [Fig. 3.11](#), based on Run 2 data. This is, by design, limited by the presence of the ATLAS mechanical structure of the air toroid in the barrel. However, it is well below the desired efficiency of 95% that has been achieved with the electron trigger. An additional intrinsic [RPC](#) inefficiency may originate from the reduction of the electric field in the gas volume, aimed to reduce the effects of chamber ageing, the use of new low greenhouse-effect mixture, and the inaccessibility of chambers for repairs. The Muon Phase-II upgrade [[3.3](#)] introduces a Barrel Inner (BI) layer in the central region, presenting an opportunity for a new flexible trigger logic that does not require the strict 3-out-of-3 chamber coincidence present in the Phase-I system.

### Level-1 Muon Triggers in the Endcap Region

The Phase-I Level-1 muon trigger in the endcap region ( $1.05 < |\eta| < 2.4$ ) combines [TGC](#) muon trigger information with that from the [NSW](#), which is composed of [small-strip TGC \(sTGC\)](#) and [Micromega \(MM\)](#) chambers. A schematic representation of the algorithm is shown in [Fig. 3.12](#). The [TGC](#) muon segment is primarily seeded on track segments in the [TGC](#) chambers of the middle muon station, located after the endcap toroid magnet. The

---

<sup>1</sup> However, for central muons there is some bending even in the “non- bending projection”.

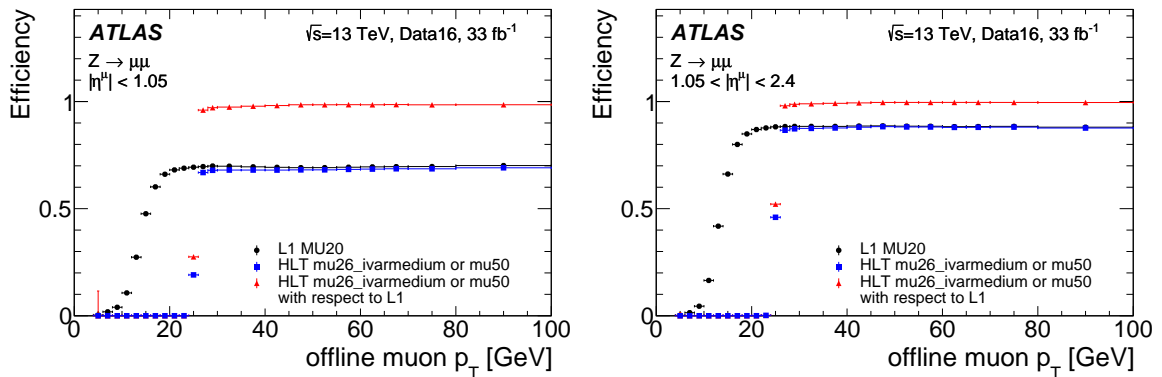


Figure 3.11: Product of geometrical acceptance and efficiency as a function of  $p_T$  as measured in an offline analysis for the present Level-1 and HLT single-muon trigger, for the barrel (left) and the endcap (right) regions. “L1 MU20” labels the Level-1 muon trigger with a  $p_T > 20$  GeV threshold. The blue points show the fractions of muons that were reconstructed by the HLT as isolated muons with  $p_T > 26$  GeV or as non-isolated muons with  $p_T > 50$  GeV. The red points show the fraction of Level-1 triggers that pass the HLT trigger requirements.

endcap muon trigger object is formed by matching the track positions of the TGC and NSW muon candidates; this correlation substantially suppresses the rate of fake triggers due to background particles emerging from the endcap toroids and calorimeters. The muon trigger object  $p_T$  is determined by the angle of the segments with respect to a straight line pointing to the nominal interaction point.

The strategy for separating trigger and precision-measurement muon detectors has been very successful, but carries a limited momentum measurement resolution of muon candidates passing the  $p_T$  nominal threshold. Consequently, the Phase-I system has a low rejection of muons below threshold. The limitation primarily arises from the short latency in the Phase-I system; a longer latency would allow for inclusion of the MDT trigger information, thus improving the momentum resolution and reducing the muon rate.

The detector layout in the transition region between the barrel and the endcap regions of the muon spectrometer, at  $1.0 < |\eta| < 1.3$ , has a complicated structure due to the presence of the barrel toroid magnets. About 70% of the transition region in azimuth is covered by the EIL4 TGC doublets and the EIL4 MDT chambers, corresponding to the large sectors of the spectrometer in-between barrel toroid coils. In small sectors, integrated chambers (BIS78), with RPC and Small-diameter Monitored Drift Tube (sMDT) detectors for trigger and tracking respectively, are installed as part of the ATLAS Phase-I upgrade programme. The trigger logic is different between the systems based on the BIS78 RPC and the EIL4 TGC chambers. In the case of the BIS78 RPC triplet chambers, a majority logic of 2-out-of-3 layers can be applied. In the case of the EIL4 TGC doublets, the only possible choice is a logical OR of the two layers. In either cases the trigger information is combined with the



### 3.2 Features and Limitations of the Run 3 Level-1 Trigger System

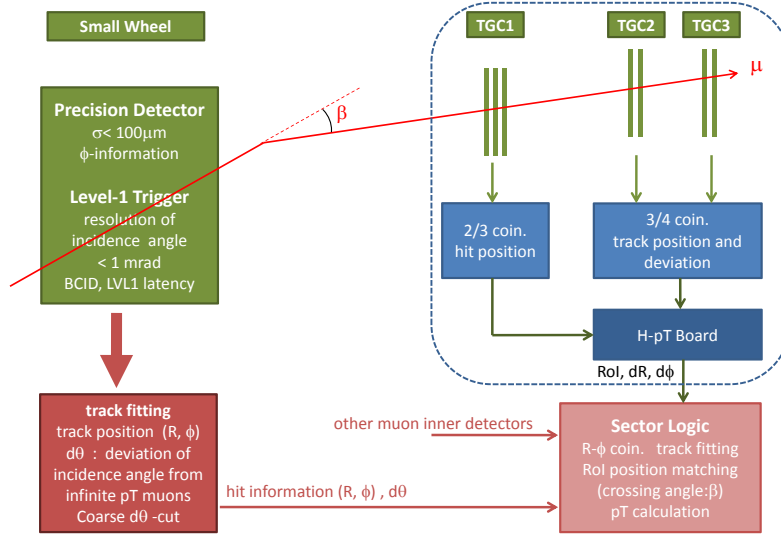


Figure 3.12: Block diagram of the endcap Trigger in Phase-I

endcap TGC information and with the information of the outer layer of the Tile Calorimeter to form an additional coincidence.

As shown in the right plot of Fig. 3.11, based on Run 2 data the muon efficiency for muons with  $p_T > 25$  GeV is  $\approx 85\%$  in the endcap ( $1.05 < |\eta| < 2.4$ ), also below the electron trigger efficiency.

#### Level-1 Di-muon Triggers

Heavy flavour physics (e.g.  $B_s \rightarrow \mu^+ \mu^-$ ) is a particular challenge for the TDAQ system, since it depends critically on low  $p_T$  thresholds in the di-muon trigger. Low di-muon thresholds, in addition to topological selections such as invariant mass and angular criteria, are of paramount importance to maintain sensitivity to  $B$ -meson decays.

The selection of close-by muon candidates in the trigger is a particular challenge; a required separation of  $\Delta\eta \times \Delta\phi = 0.2 \times 0.2$  between two trigger muon candidates results in dramatic inefficiencies for  $B_s \rightarrow \mu^+ \mu^-$  and  $B_s \rightarrow J/\psi\phi$ , as shown in Table 3.2.

Distributions of the  $\Delta\eta$  vs.  $\Delta\phi$  for the two muon candidates in  $B_s \rightarrow \mu^+ \mu^-$  and  $B_s \rightarrow J/\psi\phi$  decays are shown for four muon thresholds respectively in Figs. 3.13 and 3.14. Figure 3.15 shows the fraction of events where the two final state muons are separated by less than  $\Delta\eta \times \Delta\phi = 0.2 \times 0.2$  (Phase-I granularity of the di-muon trigger) as a function of the muon  $p_T$  thresholds.



Table 3.2: Di-muon trigger inefficiencies due to a required separation of  $\Delta\eta \times \Delta\phi = 0.2 \times 0.2$  in the trigger, for several trigger muon  $p_T$  thresholds.

trigger muon threshold		inefficiency	
$p_{T,1}$	$p_{T,2}$	$B_s \rightarrow \mu^+ \mu^-$	$B_s \rightarrow J/\psi\phi$
6 GeV	4 GeV	0.5%	5.7%
6 GeV	6 GeV	1.2%	13.7%
10 GeV	6 GeV	2.3%	23.1%
10 GeV	10 GeV	8.2%	58.8%

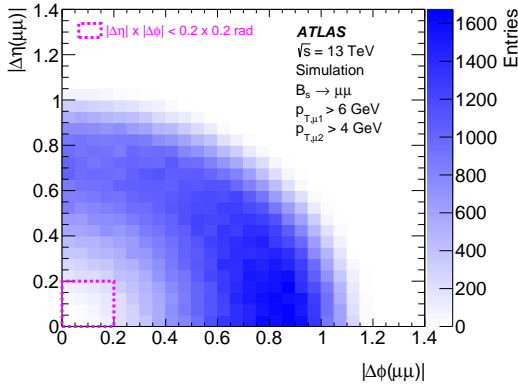
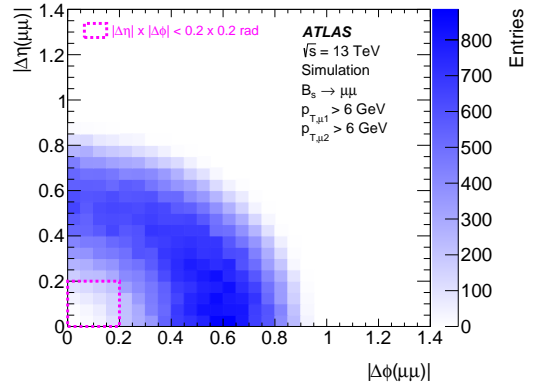
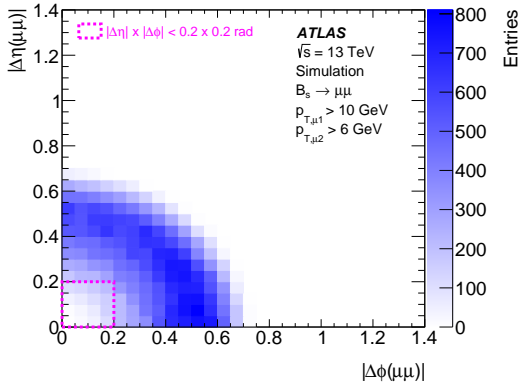
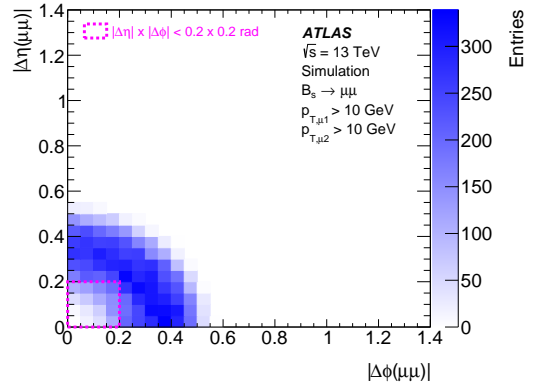
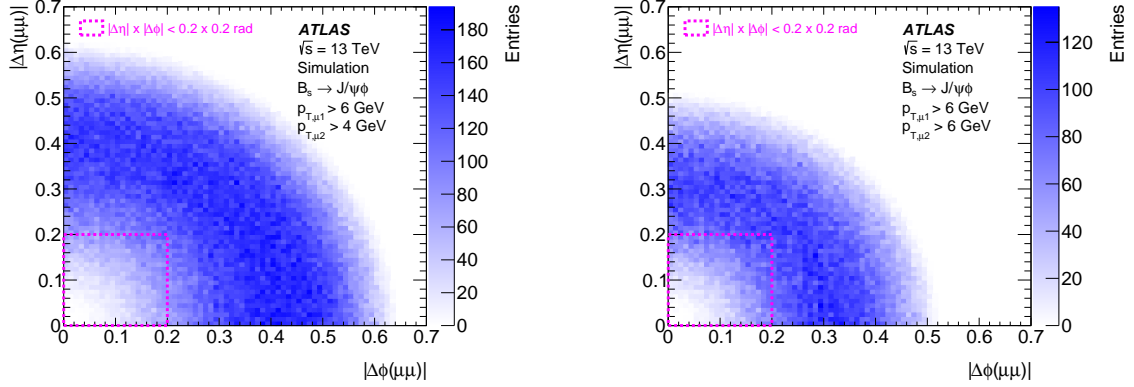

 (a) Trigger muon  $p_{T,1} > 6$  GeV,  $p_{T,2} > 4$  GeV.

 (b) Trigger muon  $p_{T,1} > 6$  GeV,  $p_{T,2} > 6$  GeV.

 (c) Trigger muon  $p_{T,1} > 10$  GeV,  $p_{T,2} > 6$  GeV.

 (d) Trigger muon  $p_{T,1} > 10$  GeV,  $p_{T,2} > 10$  GeV.

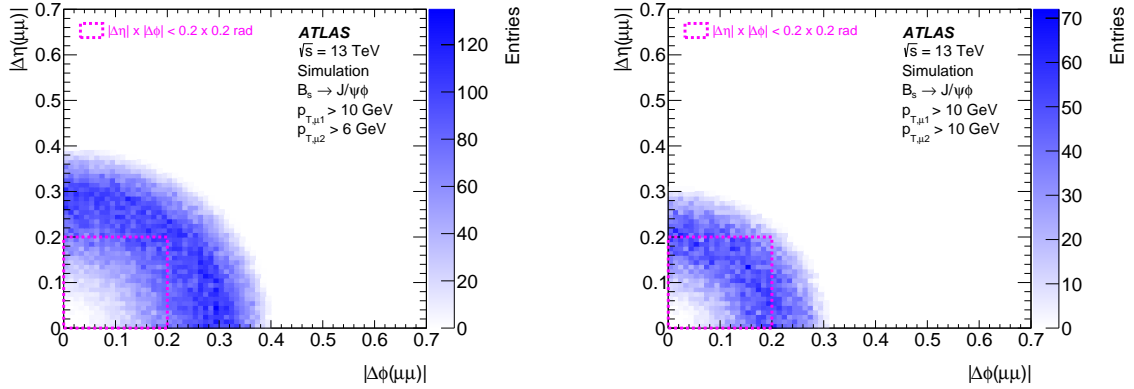
 Figure 3.13: Angular separation of muon candidates for  $B_s \rightarrow \mu^+ \mu^-$  decays as a function of di-muon trigger  $p_T$  thresholds.

### 3.2 Features and Limitations of the Run 3 Level-1 Trigger System



(a) Trigger muon  $p_{T,1} > 6$  GeV,  $p_{T,2} > 4$  GeV.

(b) Trigger muon  $p_{T,1} > 6$  GeV,  $p_{T,2} > 6$  GeV.



(c) Trigger muon  $p_{T,1} > 10$  GeV,  $p_{T,2} > 6$  GeV.

(d) Trigger muon  $p_{T,1} > 10$  GeV,  $p_{T,2} > 10$  GeV.

Figure 3.14: Angular separation of muon candidates for  $B_s \rightarrow J/\psi\phi$  decays as a function of di-muon trigger  $p_T$  thresholds.

### 3.2.3 Level-1 Topological Trigger Limitations

The original **L1Topo** system was designed to perform topological trigger algorithms for Run 2. These algorithms incorporate the geometric and kinematic relationships between trigger objects, such the angular separation between a trigger muon candidate and a jet. Similar functionality will be employed for the Run 3 **L1Topo** system. For Run 2, the **L1Topo** system sent only simple flags to the **CTP**, indicating the results of individual algorithms; multiplicity information for electrons/photons, taus and jets, was provided separately by **L1Calo**. For Run 3, **L1Topo** will provide multiplicities of trigger objects passing the requirements of certain algorithms. In order to make the Run 3 **L1Topo** compatible with the **L1Calo FEX** modules designed for Run 3, this system will accommodate a larger number of inputs and more resourceful **FPGAs**.

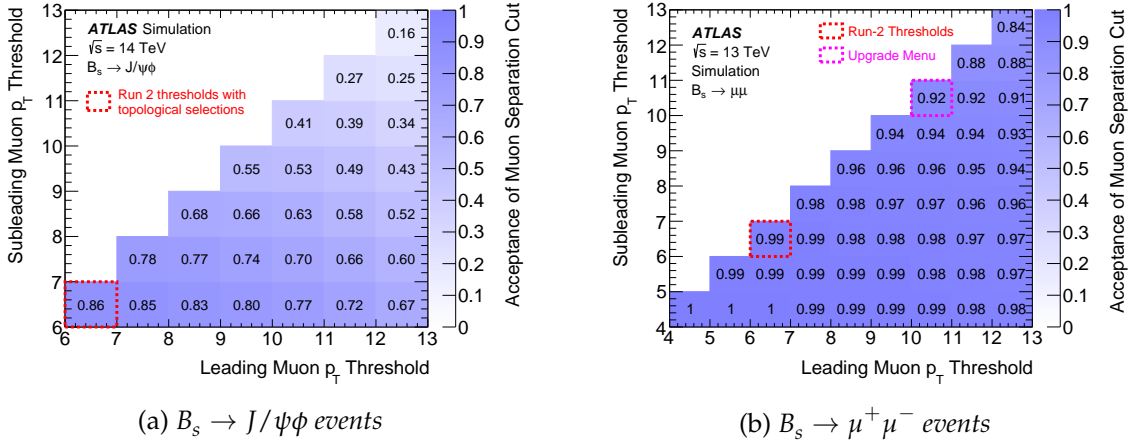


Figure 3.15: Fraction of events where the two final state muons are separated by more than  $\Delta\eta \times \Delta\phi = 0.2 \times 0.2$  as a function of the muon  $p_T$  thresholds.

However, the Run 3 **L1Topo** architecture cannot accommodate the required number of fibre inputs ( $\mathcal{O}(2000)$ ) to be able to take advantage of the full granularity calorimeter information and high-resolution full-detector muon system information that will be provided by the Phase-II detector upgrades. Furthermore, the Run 3 **L1Topo** cannot handle the number of expected trigger objects in the high-pile-up environment in Run 4.

### 3.2.4 Level-1 Trigger Menu up to Run 3

An example of how the Level-1 trigger menus have evolved through Run 1 and Run 2, and how they will evolve to cope with the **LHC** conditions after the Phase-I upgrades (Run 3) is given in Table 3.3. The Run 1 and Run 2 menus have been used in the 2012 and 2015 data taking for the LHC fills with the largest luminosity delivered. The Run 3 menu is expected to evolve as more sophisticated algorithms are being developed, some of which have been documented in the **TDAQ** and in the companion **LAr** Calorimeter Phase-I **TDR** [3.6][3.7]. The thresholds and the rates in the Run 3 column assume an instantaneous luminosity of  $3 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$  and  $\sqrt{s} = 14 \text{ TeV}$ , corresponding to  $\langle \mu \rangle \simeq 80$ .

In general a Level-1 trigger item in Table 3.3 is indicated by a “signature” key, i.e. “EM” for electrons and photons, “MU” for muons, “TAU” for taus, “J” for jets and “XE” for  $E_T^{\text{miss}}$  triggers, followed by the object’s online threshold, and by a generic key indicating whether further algorithms have been applied to improve selectivity as follows: an “H” denotes hadronic isolation, i.e. limiting the energy deposition in the hadronic calorimeter; an “I” indicates electromagnetic isolation (denoted by “I”), i.e. the object has passed selection if the energy deposited in a fixed-size cone in the electromagnetic calorimeter around the object’s is below a given value; “R” indicates the use of a shower shape algorithms in the

## 3.2 Features and Limitations of the Run 3 Level-1 Trigger System

---

electromagnetic and  $\tau$ -triggers (possible only after the Phase-I upgrades); a “V” denotes varying thresholds with  $\eta$  to account for energy loss. Finally, “L”, “M”, “T” denote loose, medium, and tight criteria depending on the values of the isolation thresholds applied.

A comparison among the three menus unveils the general strategy adopted in the past several years and planned for Run 4:

- Maintain thresholds sufficiently low to maximise signal acceptance for all the relevant electroweak signatures with leptons, allocating approximately 50-60% of the total bandwidth to single and di-lepton triggers.
- Control rates increasing with luminosity by deploying more sophisticated algorithms. For example, applying isolation to lepton trigger objects and implementing discriminants based on shower shape in the calorimeters.
- Adapt to the different luminosity conditions and optimise the menu items for hadronic signatures and topological triggers for the remaining  $\sim 40\%$  Level-1 bandwidth.

More specifically for each signature:

- Selection based on shower shapes allows the offline threshold to be maintained around 30 GeV to be compared to the 25 GeV for single electrons in Run 1.
- During Run 2, before the introduction of the [NSW](#), one may be able to maintain a trigger threshold of 20 GeV, but only by significantly cutting bandwidth from other triggers, such as those in the topological processor. After the Phase-I upgrade, there will be additional bandwidth available to design many specialised topological triggers.
- The Phase-I improvements will maintain approximately the offline thresholds of the di-tau triggers to 40 GeV, crucial for example for the reach of the  $H \rightarrow \tau\tau$  analysis, and will contain the threshold increase on the single  $\tau$ -trigger.
- In [Table 3.3](#) it has been conservatively assumed that the jet thresholds and rates will be unchanged, but it is expected that after the Phase-I improvements, the offline  $E_T^{\text{miss}}$  cut can be lowered from 250 GeV to 200 GeV [[3.7](#)].

Up to 256 menu trigger items were available in Run 1 and up to 512 are and will be available in Run 2 and Run 3. While the limitation of 512 items is not problematic, there are other reasons for replacing the [CTP](#) in the Phase-II upgrade, and additional flexibility in the menu would be highly desirable for Run 4 and beyond to benefit from enhancements elsewhere in the system. Only a few representatives of the trigger items are shown in [Table 3.3](#). They correspond to the lowest, i.e. smallest  $p_T$  threshold, unrescaled triggers of the principal single and di-object signatures. The total bandwidth allocated to a signature, as indicated in [Table 3.3](#), is only an approximated value because of the existing correlations and recurring overlaps among different triggers that needs to be taken into account. A single energy deposition can give rise to several triggers that overlap in the same region of  $\phi$  and  $\eta$ , for example, deposition of energy in the electromagnetic layers of the calorimeter can cause overlapping EM, tau and jet triggers. There is no overlap removal in the Level-1 trigger between multi-objects, except for some specialised instances in the topological processor.

### 3.2.4 Level-1 Trigger Menu up to Run 3

Table 3.3: *Level-1 Trigger menus for used configurations during Run 1 and Run 2, and example menu for Run 3. The offline thresholds typically correspond to the point at which the trigger efficiency turn-on curve reaches 90-95% of its plateau value. The items listed in this table assume no overlap removal.*

Run 1			Run 2			Run 3		
Run No.	215414		Run No.	332303				
Date	Nov 2012		Date	Aug 2017				
$\mathcal{L}$ [ $\times 10^{34} \text{cm}^{-2} \text{s}^{-1}$ ]	0.73		$\mathcal{L}$ [ $\times 10^{34} \text{cm}^{-2} \text{s}^{-1}$ ]	1.7		$\mathcal{L}$ [ $\times 10^{34} \text{cm}^{-2} \text{s}^{-1}$ ]	3	
$\langle \mu \rangle$	34		$\langle \mu \rangle$	48		$\langle \mu \rangle$	80	
$n_{\text{bunches}}$	1368		$n_{\text{bunches}}$	2544				
	Offline $p_T$ Threshold [GeV]	Rate [kHz]		Offline $p_T$ Threshold [GeV]	Rate [kHz]		Offline $p_T$ Threshold [GeV]	Rate [kHz]
EM18VH	25	19	EM22VHI <sup>a</sup>	27	27.8	EM25VHR	32	14
EM30	37	7.3				EM80	100	2.5
2EM10VH	2 $\times$ 17	6.5	2EM15VHI	2 $\times$ 18	1.7	2EM12VHR	2 $\times$ 19	5.0
<b>EM total</b>		$\sim 25$			$\sim 30$			$\sim 20$
MU15	25	9.3	MU20	27	15.5	MU20	25	15
2MU10	2 $\times$ 12	0.8	2MU10	2 $\times$ 15	1.8	2MU11	2 $\times$ 13	4.0
2MU6_MU10	8,12	1.9	2MU6_MU10	8,12	5.2			
<b>Muon total</b>		$\sim 20$			$\sim 22$			$\sim 19$
EM10VH_MU6	17,6	2.9	EM15VH_MU10	18,15	2.0	EM10VHR_MU10	17,12	3.0
TAU40	100	5.5	TAU100	170	1.2	TAU80VR	180	3.2
2TAU11I_TAU15	30,40	13.6	TAU60_2TAU40	140,100	2.8	2TAU40VR	2 $\times$ 100	3.9
TAU15_XE35	40,80	2.5	TAU20IM_2TAU12IM_4J12	40,30,40	3.2	2TAU15VR_3J15	2 $\times$ 40,50	8.1
			TAU20IM_2J20_XE45	50,2 $\times$ 60,100	2.4	TAU15VR_XE40_3J15	40,90,50	5.0
<b>Tau total</b>		$\sim 20$			$\sim 10$			$\sim 20$
J75	150	2.1	J100	460	3.3	J100	200	7.0
4J15	4 $\times$ 55	1.8	J100	235 <sup>b</sup>				
			3J50	120	0.5	4J25	4 $\times$ 60	3.3
			4J15	4 $\times$ 45 <sup>c</sup>	5.0	J75_XE40	150,150	8.3
XE40	120	5.2	XE50	200	5.0	XE70	200	13
<b>Jet/<math>E_T^{\text{miss}}</math> total<sup>d</sup></b>		$\sim 10$			$\sim 14$			<b>25</b>
<b>Topological triggers</b>		-			$\sim 5$			$\sim 15$
<b>Total</b>		$\sim 75$			$\sim 75$			$\sim 100$

<sup>a</sup> Isolation is not required for electromagnetic clusters with  $E_T$  above 50 GeV.

<sup>b</sup> Single  $b$ -jet trigger.

<sup>c</sup> Trigger requiring four  $b$ -jets and operating below the efficiency plateau of the Level-1 trigger.

<sup>d</sup>  $E_T^{\text{miss}}$  triggers in the Run 1 menu are assumed to be vetoed on the first 3 bunches to avoid huge bunch train effects giving a 15% inefficiency. Also for Run 1 some of the offline jet thresholds were set at the point where the efficiency reached 99% of its plateau value.

For example, the trigger TAU20IM\_2TAU12IM\_4J12 is designed to trigger events with four jets, and two (medium) isolated taus, where the taus may also satisfy the jet triggers.

### 3.2 Features and Limitations of the Run 3 Level-1 Trigger System

Table 3.4: Level-1  $p_T$  thresholds of few key lepton triggers at the Phase-I luminosity and at the planned peak HL-LHC luminosity of  $\mathcal{L} = 7.5 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$  assuming the Phase-I system with no upgrade. The second column indicates the Run 3 estimated rates. The thresholds indicated in the third column are the ones required to satisfy a maximum of  $\sim 100$  kHz Level-1 trigger rate that the Phase-I system can sustain.

Trigger	Level-1 $p_T$ Threshold [GeV] at $\mathcal{L} = 3 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$	Level-1 Rate [kHz]	Level-1 $p_T$ Threshold [GeV] at $\mathcal{L} = 7.5 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$
isolated single $e$	32	14	50
di- $e$	19	5	35
single $\mu$	25	15	40 (with low efficiency)

Without an upgrade of the TDAQ Phase-I system, the expected trigger rates at a peak luminosity of  $\mathcal{L} = 7.5 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$  are incompatible with the Phase-I readout system. Table 3.4 shows for a few key signatures the Level-1  $p_T$  trigger thresholds in Phase-I and the equivalent in Phase-II if we were to keep the Phase-I readout system with the required maximum Level-1 rate of 100 kHz. It can be observed that  $p_T$  thresholds would need to be increased significantly thus affecting the physics potential of ATLAS. Equivalently, the total trigger rate corresponding to the Phase-I  $p_T$  trigger thresholds would result in a factor of 10 times the allowed rate for the Phase-I readout system.

To maintain the single-electron and single-muon trigger rates at the levels achieved in Runs 1 – 3, leaving sufficient bandwidth for other important triggers in the trigger menu, the thresholds for these key triggers would have to be more than 50 GeV for electrons and more than 40 GeV for muons, as shown in Fig. 3.16. These high thresholds would significantly degrade the potential for the physics programme planned for Run 4, as shown Section 2.1. Furthermore, the improvements expected for tau lepton triggers from the Phase-I Level-1 calorimeter upgrade will be insufficient in Run 4, severely impacting the physics acceptance.

#### 3.2.5 Limitations in the Level-1 CTP and TTC systems

The CTP that was installed at the start of Run 2 will support the Phase-I system through Run 3. However, it is not adapted to the technical requirements of Run 4, such as providing readout at rates of 1 MHz or beyond, and driving the new TTC system.

The TTC system that will be used up to the end of Run 3 was conceived in the 1990s and is constrained by the technology that was available at that time. With a line rate of only 80 Mb/s, its capabilities for transmitting data with each positive trigger decision are limited. For this reason, for example, it relies on local counters in the front-end receiver ASICs for the detection of synchronisation errors. This and other considerations led ATLAS to envisage a much improved system that would be provided to all detector systems in the Phase-II upgrade, that would also be capable of supporting a two-level trigger scheme.

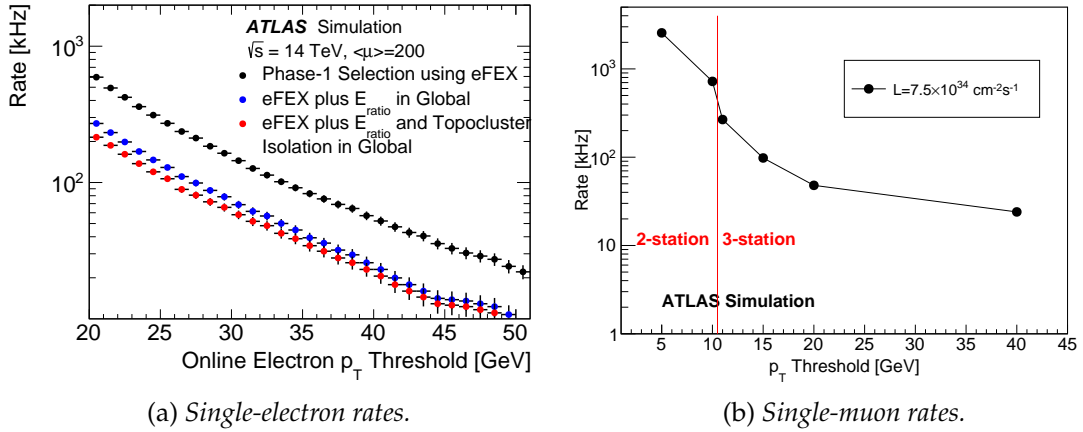


Figure 3.16: *Single-lepton trigger rates at Level-1 versus  $p_T$  threshold expected with the Phase-1 trigger system at a luminosity of  $\mathcal{L} = 7.5 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$  as extrapolated from Run 2 data. Left plot electron rates and right plot muon rates.*

The Run 2 trigger menu uses 406 inputs and 501 items, including triggers for monitoring, control samples, etc. At most about 250 out of the 501 items are enabled concurrently. It is convenient to implement a large trigger menu in which only a subset of the items is actively used at any given time, allowing rapid switching according to the instantaneous luminosity, and simplifying monitoring and offline use of trigger information. Noting that the CTP needs to be replaced for other reasons, it is clearly prudent to foresee support for larger trigger menus from Run 4. This will give flexibility to benefit from enhancements elsewhere in the trigger system, to profit from new ideas for triggering, and to react should new physics require additional trigger items.

### 3.3 Limitations of the Run 3 DAQ System

As discussed earlier in this chapter, the detector electronics readout is one of the main limitations of both the Run 3 trigger and DAQ systems. Furthermore, the readout and Dataflow components of the Run 3 DAQ system do not have the flexibility or scalability required to cope with the factor of more than twenty increase in bandwidth expected in Run 4 conditions. This large increase in bandwidth results from the increased event size and higher data rates, as discussed in this section.

#### 3.3.1 Detector Readout

Most of the present detector readout and trigger elements have hard limits on the maximum first-level trigger latency and rate they can operate, and therefore they have to be replaced



### 3.3 Limitations of the Run 3 DAQ System

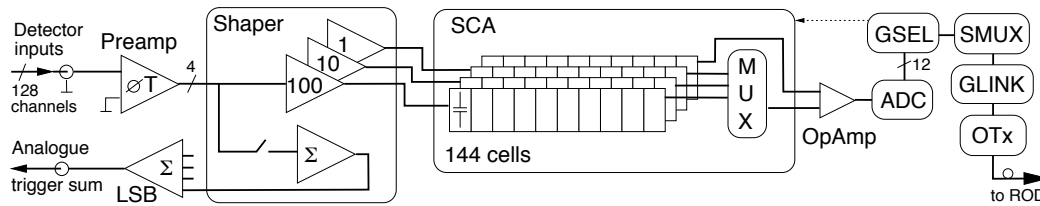


Figure 3.17: Block diagram of the FEB architecture, depicting the data flow for four of the 128 readout channels per FEB. The shaped signals are sampled at the LHC bunch crossing frequency of 40 MHz by Switched Capacitor Array (SCA) analog pipeline chips, which store the signals in analog form during the Level-1 trigger latency. The limitation on the maximum latency is coming from the SCA.

for operation in Phase-II. Only some of the on-detector front-end electronics will remain in limited cases, such as the LAr Calorimeter cold electronics in the Hadronic Endcaps and the RPC front-end electronics for the middle and outer stations. Note that systems designed for Phase-I have been designed with the aim of being used without modification in Phase-II.

**Front-end Pipeline Limitations** The original ATLAS requirements for on-detector front-end electronics stipulated that detector signals were sampled at the LHC bunch crossing frequency of 40 MHz. These original requirements also stated that signals must be stored during the latency of the Level-1 trigger of up to  $2.5 \mu\text{s}$  (100 bunch crossings), and for triggered events, that the readout must be accomplished without significant dead-time for a maximum mean Level-1 trigger rate of 75 kHz. The radiation tolerance issues required that all electronics components selected for use on the detector be subjected to an extensive radiation qualification process. This led to the development of a number of custom ASICs in specialised, radiation-tolerant semiconductor processes, and to a very limited use of commercial components. The specifications of these ASICs were pushed to the limits of the technology available at the time. However, these original specifications are a limiting factor for the detector front-end electronics (with the exceptions listed above).

One example is the LAr calorimeter front-end electronics system, which includes Front End Boards (FEBs) that perform the amplification, shaping, sampling, storage, digitisation, and readout of the calorimeter signals. In Fig. 3.17 the block diagram of the LAr FEB shows the main limitation on the maximum allowed latency due to the analog pipeline, which stores the shaper output samples for a maximum latency of  $144/40 \text{ MHz} = 3.6 \mu\text{s}$ . Similar limitations are built-in for all detector front-ends and they can be overcome only through re-design and by building new readout systems.

**Deadtime Limitations** Given the latency involved in propagating BUSY signals, some form of complex deadtime is mandatory, e.g. leaky bucket or sliding window described



below. This is in addition to the simple deadtime where a minimum time interval between successive [L1As](#) signals is required. The goal foreseen up to Run 3 was to achieve less than 2% deadtime on each detector readout system. Several Level-1 Trigger complex deadtime algorithms have been investigated as a protective measure for shaping the Level-1 deadtime; the two most successful algorithms implemented to date are:

**“Leaky Bucket” algorithm** This algorithm is driven by the data integrity versus latency. Even varying the algorithm parameters, the deadtime depends steeply on the trigger rate. This algorithm is best suited for detectors, like [LAr](#), where constraints are based on whole event buffers (fixed-size event fragments).

**“Sliding window” algorithm** This algorithm is more suited for detectors, like [Pixels](#), where the event fragment size distribution matters, with very rare large event fragments.

Deadtime estimations and measurements from Run 2 are shown for [LAr](#) and [Pixel](#) detectors in [Fig. 3.2](#), where the effectiveness of the protection with different parameter settings is shown. A very steep increase in deadtime can be observed when the Level-1 output rate increases over 100 kHz. This is a hard limitation due to design issues which are explained in the next subsection, which may be overcome only by re-designing the detector readout systems.

**Readout Bandwidth Saturation** [Figure 3.18](#) shows the number of minimum bias interactions per crossing at which the occupancy of the readout link of each detector will exceed a 90% threshold during the 2017 Run 2 data as function of pile-up. Notice that already during the Run 2 operations, the [Transition Radiation Tracker \(TRT\)](#) will start saturating the bandwidth available, and most of the inner detector will start saturating the readout during the Run 3 data taking. The hard limitation is due to the bandwidth limit (S-Link @ 1 Gbps by design) of the readout link between the [ROD](#) and the [ROS](#) systems used by the detectors. Furthermore, the [ROD/ROS](#) technology is not viable for the readout system of all the detectors at the [HL-LHC](#). A different approach based on modern technologies is needed.

Beyond the [ROL](#) occupancy limitation, the [ROS](#) is currently limited to a maximum readout fraction of  $\sim 50\%$ . In order to achieve a higher readout fraction, the firmware of the [Peripheral Component Interconnect Express \(PCIe\)](#) card would need to be rewritten. The next bottleneck to be solved would be the [CPU](#) processing speed, and the networking would ultimately limit the readout bandwidth.

### 3.3.2 Dataflow, Storage and Networking

The estimated event size for each detector system during the Run 3 operations after the Phase-I upgrades is estimated by a linear regression of the Run 2 data as shown in [Fig. 3.19](#). This estimate is summarised in the second column in [Table 3.5](#) while the third column shows the event size as estimated by each detector at the ultimate luminosity during the

### 3.3 Limitations of the Run 3 DAQ System

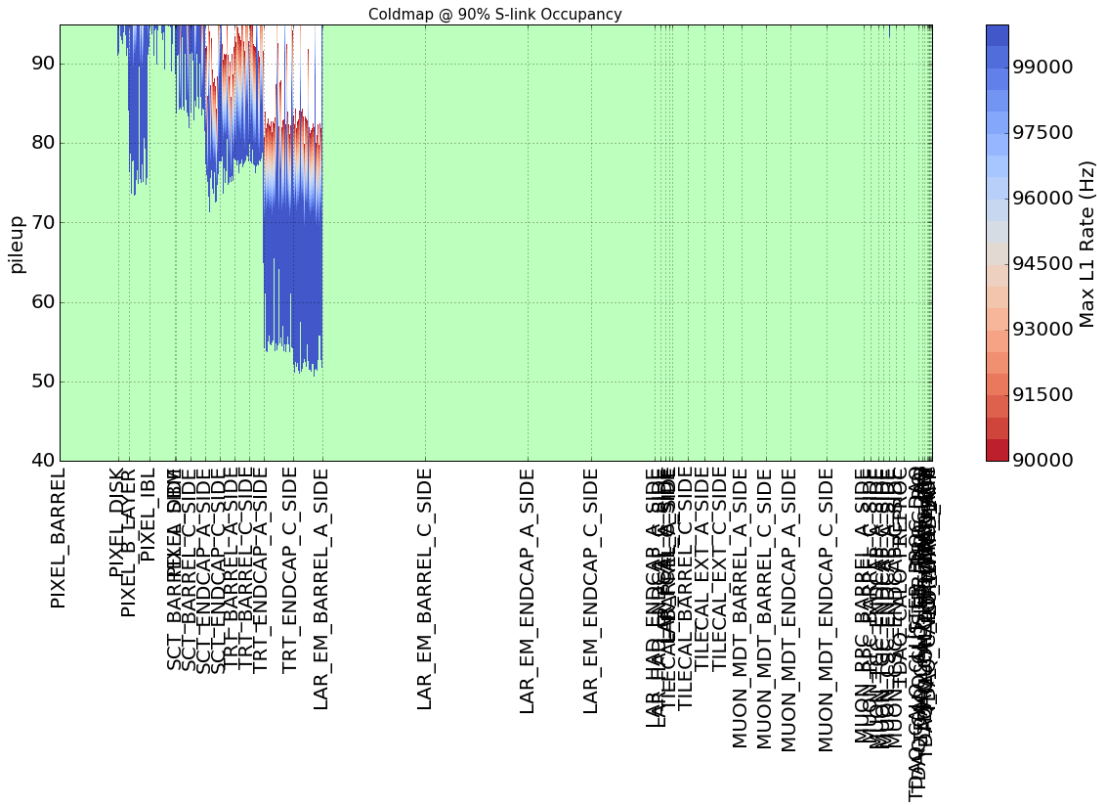


Figure 3.18: Number of interactions per crossing (“pile-up”) at which 90% of the *ROL* occupancy for each ATLAS detector system is reached. The density scale indicates the maximum Level-1 trigger rate.

**HL-LHC** operations. The large increase of the Pixel event size in Phase-II with respect to Phase-I is due to the higher detector occupancy and the significant increase of detector readout channels. A total event size of 5.2 MB is estimated in Run 4. Note that there is ~0.1 MB included in the **TDAQ** size for both Run 3 and Run 4, to accommodate the data from **LAr** trigger electronics, since **super cell** information from **LAr** is needed for the validation of the **L1Calo FEX** trigger logic. Furthermore selected ADC data further allows confirmation of energy assignment to individual bunch crossings.

For the Run 3 system, the network and storage resources have been determined using the trigger/readout parameters described in Table 3.1. There are no intrinsic limitations in the network and storage systems, which fully scale linearly with the parameters of the trigger and readout architecture of Table 3.1 and with the average event size. Therefore, the size of the system would have to be increased by a factor of approximately 2.5 just to account for the increase in event size expected in Run 3 compared to Run 4. This is without considering the additional increase in rate.

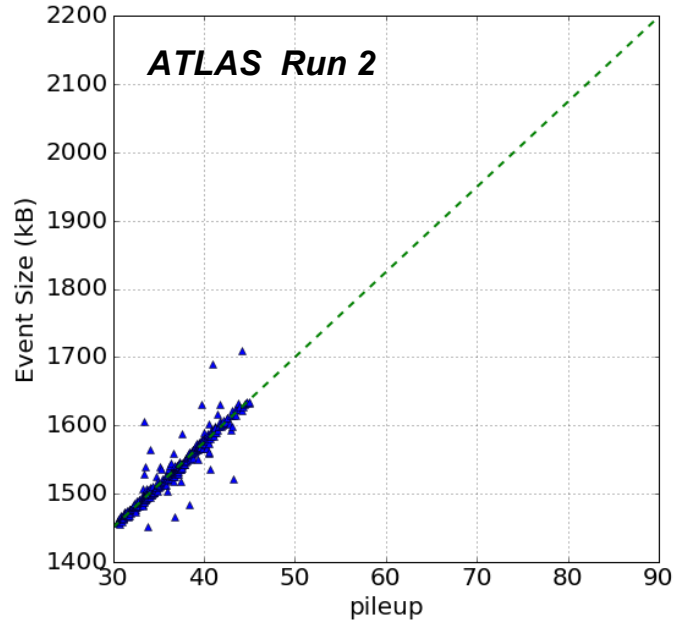


Figure 3.19: Run 2 event size (in kB) as a function of the number of interactions per bunch crossing (“pile-up”). A linear fit is applied in order to extrapolate to the Run 3 conditions after the Phase-I upgrades. Note that the extrapolation does not include the event size of the new detector systems to be introduced at the Phase-I upgrade, which is expected to contribute  $\approx 0.7$  MB in addition, at pileup of 80.

Table 3.5: Average Event Size (before *HLT* selection) as extrapolated from Run 2 data to the Phase-I conditions (including the new detector systems introduced by Phase-I upgrade), and estimated by the detector systems after the Phase-II upgrades. Forward detectors are not listed since the associated event size is negligible.

Detector System	Extrapolated Data Size [MB]	
	Phase-I	Phase-II
Pixel	0.3	2.4
Strip	0.2	0.5
TRT	0.5	-
LAr	0.7	0.7
Tile	0.1	0.2
Muon	0.5	0.8
TDAQ	0.6	0.6
<b>Total</b>	<b>2.9</b>	<b>5.2</b>

## 3.4 Limitations of the Run 3 HLT System

The Run 3 [HLT](#) system will not be able to cope with the increased event input rate and the rise in algorithm execution times resulting from the increased level of pile-up at the [HL-LHC](#). Furthermore, the hardware-based tracking system in Run 3 has been designed to cope with the pile-up conditions and Level-1 trigger rate of the Phase-I [TDAQ](#) system, and cannot perform at the same rate with the same efficiency in the [HL-LHC](#).

### 3.4.1 Estimates of Rejection Factors in the Run 3 HLT

For the [HLT](#) system to provide sufficient background rejection and reasonable rates, it is necessary to implement physics analyses requirements in the [HLT](#). Therefore, the quality of trigger object reconstruction in the [HLT](#) needs to approach that of the offline reconstruction. This statement is true for all objects including leptons, (*b*-tagged) jets, and  $E_T^{\text{miss}}$ . Thus full-event reconstruction in the calorimeters and tracking is required for most events that pass a single-electron selection. Additionally, at [HL-LHC](#) conditions where much of the offline-like selection will be implemented in the hardware-based Level-0 trigger system, background rejection becomes even more challenging.

Table 3.6, extracted from [3.6], shows the rejection at the [HLT](#) for single-lepton and di- $\tau$  triggers during Run 1 and Run 2, and the expected rejection in Run 3 after the Phase-I upgrades. As the pile-up increases, the [HLT](#) selection will be pushed at the limit to maintain reconstruction efficiency and rejection capability: (i) selection parameters will be tuned and optimised with respect to luminosity (e.g., noise thresholds), (ii) object isolation will be tightened for all the trigger signatures, and, (iii) selection algorithms at the [HLT](#) will be made very similar or identical in performance to the offline reconstruction algorithms.

Therefore, an overall rejection of  $\simeq 100$  at [HLT](#) will be a limiting value during the Run 3 data taking, and a target value for the Phase-II upgrades.

### 3.4.2 Fast Tracking

In the expected high pile-up environment, charged particle tracking can greatly assist in identifying trigger objects. However, tracking algorithms are challenging with respect to processing time; as shown in Fig. 3.20, the processing time for the existing [HLT](#) software tracking increases faster than linearly with luminosity, making it prohibitively expensive to run on every event.

A dedicated hardware-based tracking system has been installed in ATLAS in Run 2: the Fast TracKer system ([FTK](#)), that will continue to operate in Run 3. It is a massively parallel hardware system using tracking information from a region  $|\eta| < 2.5$  and designed to cope

Table 3.6: Examples of background rejection factors at the *HLT*. The first column gives the rejection of the *HLT* achieved in Run 1 and Run 2. The second column gives the expected rejection if the Run 1/2 selection criteria were run on the Phase-I Level-1 output. The third column gives the fraction of Run 1/2 events which are selected by the corresponding offline selection. The last column gives the desired rejection factor of the *HLT* in Phase-I. Because of the high purity of the single-object triggers, additional rejection must be obtained using event properties such as jets reconstructed in the *HLT*.

Signature	Run 1/2 selection	Run 1/2 selection	Run 1/2 fraction	Phase-I
	<i>HLT</i> rejection	used at Phase-I <i>HLT</i> rejection	selected offline	(stipulated) <i>HLT</i> rejection
single- $e$	180	60	$\sim 80\%$	60
single- $\tau$	80	25	$\sim 95\%$	25
di- $\emptyset$	500	250	$\sim 70\%$	300

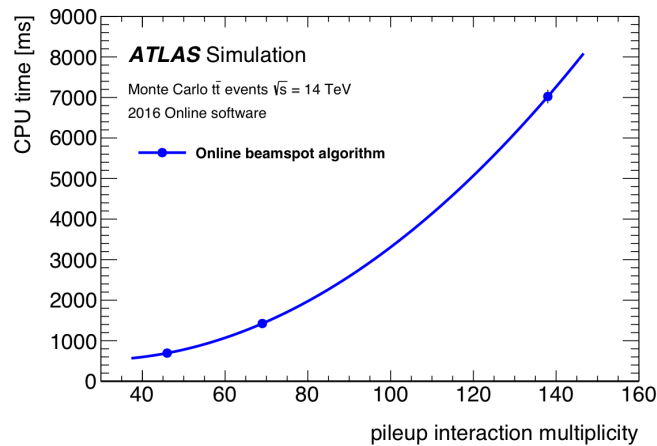


Figure 3.20: The trigger track reconstruction time for the beamspot trigger for simulated 14 TeV  $t\bar{t}$  events with 46, 69 and 138 interactions per bunch crossing with the Run 2 detector simulation, measured on a 2.4 GHz Intel Xeon CPU. The software version used corresponds to the 2016 online trigger system. Statistical uncertainties are shown. A second-order polynomial is fitted to the points.

with an average pile-up of  $\mu \simeq 80$  and at a *L1A* rate of 100 kHz [3.9]. The *FTK* is based on the CDF Silicon Vertex Trigger [3.11][3.12] and much of it is *VME*-based.

The expected Run 4 conditions will lead to higher occupancy in the *ITk*, which would result in an increase in the dataflow (considering the event size and the rate increase) by more than a factor of 20 into the *FTK*. Furthermore, the event complexity due to pile-up would significantly increase the processing power requirements for the pattern-matching stage in the *Associative Memory (AM)* chips and on the processing *FPGAs* that fit the track candidates. Processing events with  $\mu \simeq 140$ -200 would lead to long tails in the event processing time distribution, which intrinsically limit the maximum operating rate of *FTK* well below the 100 kHz *L1A* rate. To compensate for this, the processing time could be reduced, and

### 3.4 Limitations of the Run 3 HLT System

---

Table 3.7: Maximum Level-1 processing rate of the *FTK* system and reconstruction efficiency for tracks with  $p_T$  above 1 GeV during the Phase-I and *HL-LHC* data-taking conditions. To recover 100 kHz processing rate during the *HL-LHC* the reconstruction efficiency required for tracks needs to be loosen to 60%.

Parameter	$\mu = 80$	$\mu = 200$	
Max. Processing Rate [kHz]	100	5	100
Efficiency (for tracks with $p_T > 1$ GeV)	$\sim 90\%$	$\sim 90\%$	$\sim 60\%$

the tail in the distribution suppressed, by relaxing the requirements on the reconstruction efficiency for tracks with  $p_T$  above 1 GeV.

Table 3.7 summarises the *FTK* performance in the configuration optimised for Phase-I data taking and a crude estimation when the same configuration is used in a *HL-LHC*-like environment at 100 kHz Level-1 trigger rate. To maintain an efficiency of  $\sim 90\%$  on tracks with  $p_T$  above 1 GeV, the *FTK* would need to operate on only 5% of the events (equivalent to 5 kHz trigger rate in this case). Alternatively, the reconstruction efficiency would need to be relaxed to 60% to be able to run *FTK* on all the 100 kHz Level-1 triggered events. In order to get a reasonable efficiency, the track  $p_T$  requirement would need to be raised to  $\sim 10$  GeV.

This limitation in the *FTK* performance impacts the trigger selection of hadronically decaying  $\tau$  leptons,  $b$ -tagged jets, single-jet events, multi-jet events, and  $E_T^{\text{miss}}$ , as described below.

#### 3.4.3 Features and Limitations of Event Selection in the Run 3 HLT

The Run 3 *HLT* menu is designed for a peak output rate of 1.5 kHz (an average rate of 1 kHz) with  $\sim 2/3$  of that allocated for leptonic triggers and  $\sim 1/3$  for hadronic triggers. For single-lepton triggers, the production of  $W$  and  $Z$  bosons provides an important constraint on the *HLT* trigger menu. The goal of the Run 3 physics programme is not necessarily to collect as many single  $W$  and  $Z$  bosons as possible, but rather to use them to trigger physics processes such as associated Higgs production ( $WH$  and  $ZH$ ) with minimal bias. This is achieved by postponing additional trigger selections except for those on the single lepton until the full detector information is available at the *HLT*. At the luminosities expected in Run 3, the total rate of electrons and muons from  $W$  boson decays is approximately 1.2 kHz, which by itself would saturate the allowed *HLT* output rate. The rate of full reconstruction for electron and muon events is expected to be 2.5 kHz. About 50% of these electrons and muons are inside the trigger acceptance, but even electron and muon selections with very tight isolation requirements at the *HLT* have considerable background from heavy quark decays and fake jets, raising the rate by a counterbalancing factor of two. Thus additional

rejection can be achieved by including additional event information (such as the presence of jets reconstructed in the HLT).

In the Run 3 HLT, the hadronic  $\tau$  reconstruction strategy has two components. The first component is a simple selection based on the numbers of FTK tracks in a core and an isolation cone, defined with respect to the highest- $p_T$  track in the tau RoI (as described in Ref. [3.9]) to do a fast rejection. The second component is a more sophisticated algorithm involving topological clusters and tracks, mimicking the offline multivariate selection. Additionally, FTK tracks can be used to do full-scan tau finding to recover inefficiencies due to the relatively high tau  $p_T$  threshold at Level-1. Final states involving hadronically decaying  $\tau$  leptons predominantly have at least two trigger objects in the final state, allowing for very large rejection factors, as can be seen in Table 3.6. Therefore, it will generally be unnecessary to add objects beyond those used to seed the triggers at Level-1 in order to obtain low HLT output rates. Similarly, multi-object tau triggers with electrons and muons will have high rejection rates and will not require the identification of additional objects in the HLT.

The identification of jets arising from  $b$ -hadrons is a powerful tool for background rejection that is only available in the HLT. Historically,  $b$ -tagging algorithms were a large consumer of HLT CPU time. To maintain the same relative CPU usage as the rate of pile-up increases, the  $p_T$  threshold of  $b$ -tagged jets must be raised to compensate for the increase in the number of jets present in the events and the more complicated pattern recognition task. By removing the pattern recognition and track-fitting step, the FTK allows the HLT to process many more RoIs for  $b$ -tagging. FTK tracks can be used directly by the tagger, or they can first be refit with the HLT fitting algorithms to improve the track parameter resolutions. In Ref. [3.9] it was shown that simple, pile-up-robust FTK  $b$ -tagging algorithms could be developed with only modest decreases in efficiency with respect to offline taggers.

For generic single-jet triggers and multi-jet triggers, full calorimeter readout allows more accurate determination of jet energies using the full offline jet calibration procedure, including pile-up suppression and correction. This allows the HLT thresholds to be placed very close to the offline ones. The full readout also allows iterative event-level jet-finding algorithms, such as the anti- $k_T$  algorithm, to be used. This requires that the HLT supports the full calorimeter readout rate in excess of 20 kHz, and that sufficient CPU resources are available for the topological clustering of the calorimeter information, as well as calibration close to what is used offline.

The most important of the hadronic triggers in the initial running at the LHC has been the  $E_T^{\text{miss}}$  trigger, which has figured prominently in the Higgs analyses (e.g.  $ZH \rightarrow \nu b \bar{b}$  as discussed in Section 2.6), as well as in the majority of the SUSY and Exotics analyses with hadronic final states. In Run 2 and Run 3, the expected increase in pile-up will inevitably degrade  $E_T^{\text{miss}}$  performance and also lead to a large increase in soft jets from pile-up. Sophisticated pile-up-suppression techniques will be used at the HLT, and FTK tracks can be used to improve the calibration of the reconstructed jets in the trigger. For generic



$E_T^{\text{miss}}$  triggers, and for combined jet and  $E_T^{\text{miss}}$  triggers, the HLT must closely match to the offline algorithm to maximise the overlap of events selected by the two algorithms. The  $E_T^{\text{miss}}$  threshold for a generic  $E_T^{\text{miss}}$  algorithm will have to be approximately a factor of 2 or 2.5 higher than the Level-1 threshold to obtain a factor of 100 rejection. It is anticipated that there will be a large number of specialised HLT algorithms seeded from Level-1  $E_T^{\text{miss}}$ , which are based on jets (with and without  $b$ -tagging applied) that can have  $E_T^{\text{miss}}$  thresholds closer to the Level-1 thresholds. For example, offline  $E_T^{\text{miss}}$ -based algorithms typically apply requirements on the direction of the  $E_T^{\text{miss}}$  relative to hard jets in the event which can be implemented at the HLT.

## References

- [3.1] ATLAS Collaboration, *Technical Design Report for the Phase-II Upgrade of the ATLAS LAr Calorimeter*, Tech. Rep. CERN-LHCC-2017-018. ATLAS-TDR-027, CERN, Geneva, Sep, 2017. <https://cds.cern.ch/record/2285582>.
- [3.2] ATLAS Collaboration, *Technical Design Report for the Phase-II Upgrade of the ATLAS Tile Calorimeter*, Tech. Rep. CERN-LHCC-2017-019; ATLAS-TDR-028, CERN, Geneva, Sep, 2017. <http://cds.cern.ch/record/2285583>.
- [3.3] ATLAS Collaboration, *Technical Design Report for the Phase-II Upgrade of the ATLAS Muon Spectrometer*, Tech. Rep. CERN-LHCC-2017-017. ATLAS-TDR-026, CERN, Geneva, Sep, 2017. <https://cds.cern.ch/record/2285580>.
- [3.4] ATLAS Collaboration, *Technical Design Report for the ATLAS Inner Tracker Pixel Detector*, Tech. Rep. CERN-LHCC-2017-021, CERN, Geneva, (in preparation).
- [3.5] ATLAS Collaboration, *Technical Design Report for the ATLAS Inner Tracker Strip Detector*, Tech. Rep. CERN-LHCC-2017-005, CERN, Geneva, Apr, 2017. <https://cds.cern.ch/record/2257755>.
- [3.6] ATLAS Collaboration, *Technical Design Report for the Phase-I Upgrade of the ATLAS TDAQ System*, Tech. Rep. CERN-LHCC-2013-018. ATLAS-TDR-023, CERN, Geneva, Sep, 2013. <https://cds.cern.ch/record/1602235>.
- [3.7] ATLAS Liquid Argon Calorimeter Group, *ATLAS Liquid Argon Calorimeter Technical Design Report: Trigger Electronics for the LHC Phase-I Upgrade*, Tech. Rep. CERN-LHCC-2013-017, CERN, Geneva, (in preparation). <https://twiki.cern.ch/twiki/bin/view/LAr/LArPhaseITDR>.
- [3.8] ATLAS Collaboration, *ATLAS New Small Wheel Technical Design Report*, Tech. Rep. CERN-LHCC-2013-006. ATLAS-TDR-020, CERN, Geneva, Jun, 2013. <https://cds.cern.ch/record/1552862>.



- [3.9] ATLAS Collaboration, *Fast TracKer (FTK) Technical Design Report*, Tech. Rep. CERN-LHCC-2013-007. ATLAS-TDR-021, CERN, Geneva, Jun, 2013.  
<https://cds.cern.ch/record/1552953>.
- [3.10] ATLAS Collaboration, G. Aad et al., *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, *Eur. Phys. J. C* **77** (2017) 490, [arXiv:1603.02934](https://arxiv.org/abs/1603.02934) [hep-ex].
- [3.11] S. Belforte et al., *The CDF trigger silicon vertex tracker (SVT)*, *IEEE Trans.Nucl.Sci.* (1995) 860.
- [3.12] J. Adelman et al., *The Silicon Vertex Trigger upgrade at CDF*, *NIM* (2007) 361.



## 4 Architectural and Functional Requirements

The upgraded [TDAQ](#) system shall allow for the broad physics programme planned for the [HL-LHC](#), as described in Chapter 2, while maintaining the flexibility needed for standard triggers, exotic object triggers and the potential to respond to changing priorities. As described in [4.1], the Phase-II upgrade programme for the ATLAS detector contains a variety of improvements, such as extended coverage in the forward region for electrons, muons and charged particle tracks. ATLAS has come up with a consistent set of parameters for the design and implementation of the Phase-II upgrade for the subdetector systems and [TDAQ](#). These are documented in an internal [FE](#) Interface Requirement Document [4.2], which defines also a formal approval process and revision control to ensure coherency.

This chapter is organised as follows. The detector improvements foreseen in Phase-II that impose constraints on the Phase-II [TDAQ](#) system are described in Section 4.1. Section 4.2 includes a summary of the Level-0 trigger requirements. The requirements driving the technical design of the Data Acquisition are described in Section 4.3. And finally, Section 4.4 describes the requirements on the [EF](#) system.

### 4.1 Constraints from the Detectors

#### 4.1.1 New Inner Tracker

The new ATLAS Inner Tracker ([ITk](#)) proposed for the Phase-II upgrade is described in detail in two technical design reports: the [ITk](#) strip detector [TDR](#) [4.3] and the [ITk](#) pixel detector [TDR](#) [4.4]. Various characteristics provide constraints to the design of the [TDAQ](#) system in Phase-II and are described in the following.

#### Additional Coverage

The proposed [ITk](#) acceptance (up to  $|\eta| < 4.0$ ) provides an opportunity to extend various triggers to include the forward region through regional or full-detector tracking. A calorimeter-only forward electron trigger is expected to have a high trigger rate due to the coarser calorimeter segmentation and high pile-up in the forward region. Such a trigger would therefore need to apply a high  $p_T$  threshold or require a matching track. Identification of tau leptons and jets arising from  $b$ -hadrons may also be improved over the full

## 4.1 Constraints from the Detectors

acceptance of the *ITk*, as well as track-based pile-up suppression to aid the performance of hadronic triggers.

### ITk Strip Detector

A primary driver of the *TDAQ* system architecture is the limitation on the readout rate of the *ITk* strip detector. The strip detector's on-detector *ASICs*, the core components of the *FE* electronics, are being designed to meet the preliminary trigger requirements for either the baseline or evolved *TDAQ* system options, including high priority requests for data in small regions of the detector needed for the evolved option. As detailed in [4.3], the on-detector data transmission is based on the Hybrid Controller Chip *ASIC* that transmits the data from 10 *FE ASICs* on a 640 Mbit e-link to end-of-stave cards housing one or two *Low Power Gigabit Bidirectional Trigger and Data Link (lpGBT)* transceivers. Table 4.1 shows for each barrel layer and endcap disc of the *ITk* strip detector the expected occupancy at  $\langle \mu \rangle \simeq 200$ , the hit rate, the effective payload in the *lpGBT* readout link for a 1 MHz readout rate when a 8b/10b encoding protocol is used, and the equivalent maximum trigger rate capability. The trigger rate capability is estimated considering that the available e-link bandwidth saturates within a 75% margin and that the time required to read out 99% of the data is less than 50% of the Level-0 latency (to minimise queuing effects in the Front-End *ASICs*). A constraint of a maximum of 1 MHz full-detector readout rate capability results from the *ITk* strip detector design.

Table 4.1: *Expected maximum occupancy and payload per link for the various layers of the ITk-strip detector (from Ref. [4.3]). The last column is the maximum average trigger rate sustainable by the readout within a 75% margin on the available bandwidth of the readout link.*

Layer	Channel Occupancy [%]	Hit Rate [/ $\text{mm}^2/\text{BC}$ ]	Payload/E-link [% BW @ 1 MHz] (8b/10b encoding)	Max. Trigger Rate [MHz]
L0	0.94	$5 \times 10^{-3}$	37.1	1.6
L1	0.55	$3 \times 10^{-3}$	23.4	2.3
L2	0.55	$1.5 \times 10^{-3}$	25.4	2.5
L3	0.35	$0.7 \times 10^{-3}$	15.4	2.9
D0	0.96	$5.6 \times 10^{-3}$	41.6	1.4
D1	0.98	$5.2 \times 10^{-3}$	43.6	1.3
D2	1.00	$5.0 \times 10^{-3}$	45.7	1.2
D3	1.02	$4.9 \times 10^{-3}$	48.4	1.1
D4	1.06	$4.8 \times 10^{-3}$	51.2	1.1
D5	1.16	$4.7 \times 10^{-3}$	55.3	1.0

### ITk Pixel Detector

For the innermost two layers of the [ITk](#) Pixel detector, space and material constraints limit the bandwidth (5 Gb/s) of the copper cables driving the data to the service panels that house the optical transceivers. Thus, the maximum readout rate for complete events is 1 MHz. The total memory is limited by the area of the [ASIC](#) and corresponds to a maximum latency of 12.5  $\mu\text{s}$ .

## 4.1.2 Calorimeter Detectors

The [LAr](#) calorimeter will be improved in Phase-II with an electronics upgrade that will provide optimised [super cells](#) and full granularity data to the trigger system by means of a new pre-processor ([LAr Signal Processor \(LASP\)](#)). A similar upgrade of the [Tile](#) calorimeter readout will use on-detector digitisation and a new backend pre-processor ([Tile PreProcessor System \(TPPr\)](#)). Both the [LAr](#) and [Tile](#) calorimeters expect to implement a 40 MHz readout system for Phase-II. The digitised data will be streamed off-detector at 40 MHz and buffered in modular electronics. This solution is sufficiently fast and with high bandwidth and buffering capability that it places no constraints on the foreseen latencies or rates.

The transmission of high-granularity calorimeter data (all cells above a transverse energy threshold of  $|E_T| > 2\sigma$ ) drives the bandwidth requirement for the upgraded [TDAQ](#) system. The upgraded trigger system must be able to form trigger objects (such as jets) and topological selections based on this high-granularity data. The [LAr](#) and [Tile](#) trigger pre-processors will provide [super cell](#)-derived information to the [FEX](#) modules of the [L0Calo](#) trigger subsystem. The outermost [Tile](#) calorimeter layer can be used to identify muons in the range  $|\eta| < 1.3$  by better identifying particle energy depositions above the [Minimum Ionising Particle \(MIP\)](#) threshold.

### 4.1.3 Muon System

The detailed upgrade plan for the muon system is described in [4.5]. The muon trigger system consists of the Resistive Plate Chambers ([RPCs](#)) in the barrel region ( $|\eta| < 1.05$ ) and Thin Gap Chambers ([TGCs](#)) in the endcap region ( $1.05 \leq |\eta| < 2.4$ ) of the ATLAS detector. After the Phase-I upgrade, the New Small Wheels ([NSWs](#)) will maintain high muon trigger acceptance at high pile-up in the region  $1.3 < |\eta| < 2.4$ . In Run 4, a system of thin-gap [RPCs](#) will cover the entire inner barrel layer and Monitored Drift Tube ([MDT](#)) hits will be incorporated in the hardware-level trigger.

### Resistive Plate and Thin Gap Chambers

A completely new readout system is planned for both the [RPC](#) and [TGC](#) detectors for Phase-II. The digitised data will be shipped off-detector at 40 MHz and buffered directly in the [L0Muon](#) Trigger Processor modular electronics in [USA15](#). The upgrade is designed to be sufficiently fast and with high bandwidth that it places no constraints on latencies or rates. However, as there will be no distinction between trigger data and readout data paths, the [L0Muon](#) trigger processors shall implement the interfaces to the Readout system.

### Monitored Drift Tubes

As the innermost [MDT](#) chambers are replaced, no limitations are placed on the upgraded readout electronics. The long drift time ( $\sim 700$  ns) of the [MDT](#) chambers is close to the average time between two Level-0 triggers; thus, there is no gain from assigning muon hits to a specific event. As for the [RPC](#) and [TGC](#) readout the [MDT](#) Trigger Processor modules shall provide directly the interface to the Readout system as there are no separate data paths. The baseline readout option is to implement Level-0 buffers in the [MDT](#) Trigger Processors, transmitting data packets with all hits belonging to a specific event, and thus allowing data duplication for consecutive triggers. Alternatively, all detector hits from the cavern might be transmitted continuously and a software filter at a higher level in the readout chain shall be implemented with a proper time-tagging reconstruction of the event.

### New Small Wheel

The [NSW](#) readout was designed in 2013 according to what at the time was expected to become the ATLAS Phase-II trigger and readout architecture: a two-level hardware trigger scheme with maximum rates of 500 kHz/200 kHz. To cope with the current design, the [NSW](#) electronics now support the Phase-II [TDAQ](#) specifications, i.e. Level-0 rates of 1 MHz with a maximum latency of  $10 \mu\text{s}$  [4.6]. This requires an increase in the readout bandwidth which will be obtained by increasing the number of data collector chips and optical links, additional to the Phase-I deployment.

The hard limit on the maximum rate is given by the readout rate of the micromega sectors with largest data volume. This is limited to about 1.2–1.3 MHz, without safety factors, in a very limited region of the detector. A safety factor of  $\sim 1.3$  seems appropriate to take into account the new shielding design. Multiple options are available to mitigate this (very local) issue with no or minimal data loss. Three scenarios allowing a readout rate of  $\sim 2$  MHz were identified, with a reduced performance in a small part of the forward detector region: (i) using a reduced readout time window, namely 3 [Bunch Crossings \(BCs\)](#) instead of 5, (ii) suppression of [Time-to-Digital Converter \(TDC\)](#) data, or (iii) suppression of empty events.

#### 4.1.4 Other Detectors

Including additional subdetectors in Phase-II requires that the [TDAQ](#) system architecture provides the necessary infrastructure to allow taking their data. For example, the integration of the [HGTD](#) detector in the ATLAS [TDAQ](#) system would require dedicated hardware processors and storage of the real-time luminosity measurement through a dedicated data path. Additionally, trigger signals would need to be provided to the Level-0 trigger system thus allowing triggers on minimum-bias events. Supplementary readout capacity for driving the [HGTD](#) information will be needed.

## 4.2 Level-0 Trigger Requirements

The Phase-II trigger system shall maintain legacy hardware, where appropriate, accommodate new detectors and exploit the full granularity provided by the detectors.

### Calorimeter Trigger Requirements

Performance requirements on physics objects translate into the following functional and architectural requirements on the trigger processors that use calorimeters information:

- Highly selective single- and multi-lepton triggers are essential to exploit the full physics potential of the [HL-LHC](#). Therefore, a large fraction of the available Level-0 trigger bandwidth shall be allocated to the single-electron trigger of a  $p_T$  threshold as low as possible and necessary for the physics programme.
- Additional discriminating variables (e.g.,  $E_{\text{ratio}}$  and topological cluster-based isolation, as described in Section 6.3) that can identify electron candidates are required in order to provide sufficient background rejection (at  $\langle\mu\rangle \simeq 200$ ) while maintaining a low  $p_T$  threshold and high trigger efficiency.
- The identification of events with low- $p_T$  multi-jets is particularly challenging in a high-pile-up environment. Sufficient processing power is needed to implement the sophisticated tools that are used in offline event reconstruction: pile-up suppression algorithms, offline-like jet reconstruction algorithms to recover low- $p_T$  jets, and jet energy calibration to improve the momentum resolution.
- The Phase-I [FEX](#) modules shall be maintained during the [HL-LHC](#) data taking as part of the [L0Calo](#) trigger system. A re-optimisation of the algorithms within the available memory resources and the upgraded latency budget of the system will be needed to mitigate higher pile-up conditions.
- The [L0Calo](#) trigger system shall be complemented by new processors ([Global Trigger](#)) to extend the functionalities and the resource limits of the [L0Calo FEXs](#) processors. In particular the [Global Trigger](#) shall (i) refine the calculations of the [FEX](#) modules by implementing the algorithms aforementioned that are based on higher granularity data,

## 4.2 Level-0 Trigger Requirements

---

- (ii) apply generalised isolation criteria, and (iii) extend and replace the functionality deployed in the Phase-I topological processors ([L1Topo](#)).
- Several physics analyses will benefit from efficient Level-0 reconstruction of forward objects such as jets or electrons using the full granularity of the [FCal](#) through dedicated forward [FEXs](#) modules.
- The missing transverse momentum ( $E_T^{\text{miss}}$ ) trigger performance at Level-1 will benefit from improvements in jet trigger object reconstruction and full-scan tracking to compute the soft term of the  $E_T^{\text{miss}}$ . In the case of evolution to a dual-level hardware trigger system, tracking information shall be integrated with jet reconstruction in the [Global Event Processors \(GEPs\)](#).

### Muon Trigger Requirements

The momentum resolution of the Level-0 muon trigger is a key attribute in order to select muons with  $p_T > 20$  GeV ( $p_T > 10$  GeV) for single-lepton (multi-lepton) triggers with a rate smaller than approximately 40 kHz. The spatial resolution of the [RPCs](#) and [TGCs](#) limits the momentum resolution. In the Level-0 muon trigger system, a momentum resolution of 5% for 20 GeV muons can be accomplished by the reconstruction of straight muon track segments in the [MDT](#) chambers.

- To comply with the requirements on resolution and efficiency, [MDT](#)'s information shall be used at Level-0.
- The trigger chambers shall send the trigger candidate data to the [MDT](#) Trigger Processor within  $\approx 1.7 \mu\text{s}$ , to allow sufficient processing time to the [MDT](#) processing.
- The maximum seeding frequency per trigger region handled by a [MDT](#) processing element shall be set to a value of 100 kHz.

### Trigger and Readout System Parameters and Requirements

Table [4.2](#) summarises the key parameters of the design of the [TDAQ](#) system in Phase-II, extracted from [\[4.2\]](#). Their description follows.

1. **Level-0 trigger rate:** A total of 1 MHz is the maximum average Level-0 trigger rate, integrated over many full machine cycles, including gaps. The [FE](#) systems and [TDAQ](#) detector readout system shall be designed for sustained operation at this rate.
2. **Minimum interval between two Level-0 Trigger Accept (LOA) signals (so-called Level-0 simple deadtime):** [LOA](#) signals may occur in consecutive bunch crossings.
3. **Consecutive Level-0 triggers:** at most four [LOA](#) in five consecutive [BCs](#) are allowed.
4. **Level-0 burst size:** The Level-0 system shall guarantee no more than eight [LOA](#) in any period of  $0.5 \mu\text{s}$ . This is consistent with a deadtime of 0.1% for a 1 MHz Level-0 rate in a simple model (worst case) in which the Global Trigger processes events in parallel. In addition, the Level-0 system shall guarantee no more than 128 [LOA](#) in any period



- 
- of 90  $\mu\text{s}$  to be consistent with the 0.1% deadtime requirement. This is consistent with an average 1 MHz Level-0 trigger rate and the available buffer size in the front-end ASIC of the strips detector. Note that other limitations can be introduced if needed.
5. **Data transmission to Level-0 Trigger Processors:** The skew in the arrival times of the different calorimeter and muon detector data at the Level-0 Trigger Processors shall be accommodated by dedicated buffers prior to transmission by the relevant subdetector systems or by buffers at the input stage of the relevant Trigger Processors (LOCalo FEXs and Global Trigger modules for the calorimeter, and L0Muon Sector Logic, MDT Trigger Processors for the muon detectors). For the calorimeters, the skew has a maximum possible value of 16 BCs. For the muon detectors the skew is determined by the maximum drift time in the MDT detectors and the fast arrival of the RPC and TGC trigger objects, calculated in Sector Logic modules, into the MDT Trigger Processors.
  6. **Level-0 latency:** The Level-0 system shall be designed to have a fixed maximum latency, as defined in Section 5.2.8, of 10  $\mu\text{s}$ .
  7. **Maximum deadtime value:** Deadtime is the fraction of triggers vetoed by the CTP. A maximum of one percent (1%) deadtime is required for ATLAS as a whole. The 0.1% deadtime per subdetector system is derived by considering all the different ATLAS subdetectors planned to operate during the HL-LHC. A subdetector system shall not veto more than 0.1% of the triggers by the assertion of the BUSY signal and/or by the algorithm implemented in the CTP that specifically vetoes triggers to protect it (a.k.a. complex deadtime). The subdetector systems shall be responsible of defining buffer and/or pipeline dimensions and data transmission speeds in accordance with this specified deadtime.

## Timing and Control Distribution Requirements

Timing, Trigger, and Control (TTC) signals need to be distributed to the common elements (FELIX) of the readout system and other subdetector-specific electronics. Based on the experience from the current implementation, the timing signals must include the 40 MHz beam-synchronous clock derived from the LHC, which is used to clock the subdetector front-end electronics, as well as the ORBIT signal that is synchronous with the LHC turns. Trigger signals must include the LOA signal, along with associated information, such as a LOA counter, a bunch-crossing identifier, type of trigger decision, etc. In addition, control signals and data, sent synchronously or asynchronously with the beam will be needed for specific uses. During physics data-taking, the Central Trigger will act as TTC master, while for calibration and test runs the system needs to be partitioned and controlled by Local Trigger Interface (LTI) modules. The transmission of the downlink TTC signals has to be done with low latency, low jitter, and the quality of the transmitted clock has to be compatible with the very demanding Multi-Gigabit Transceivers (MGT), which are used everywhere along the trigger and readout paths. In the opposite direction, BUSY signals have to be collected from the common readout elements or subdetector-specific electronics

## 4.2 Level-0 Trigger Requirements

Table 4.2: List of the main trigger and readout parameters in the baseline architecture. Latency values are given with respect to the original BC. The values indicated in this table represent the requirements for the interfaces between the detector-specific FE and the TDAQ system. The quoted values include contingency and margins.

Parameter	Phase-II value
Clock frequency	40.08 MHz
Level-0 trigger rate	1 MHz
Minimum interval between two L0A signals	0 BC
Consecutive Level-0 triggers	$\leq 4$ L0A in 5 BC
Level-0 burst size	$\leq 8$ L0A in $0.5 \mu\text{s}$ $\leq 128$ L0A in $90 \mu\text{s}$
Maximum skew between all calorimeter inputs to L0Calo FEXs	16 BC
Level-0 latency	$10 \mu\text{s}$
Calorimeter data reception in L0Calo & L0Muon processors	$1.7 \mu\text{s}$
High Granularity Calorimeter data reception in Level-0	$1.7 \mu\text{s}$
Seeding Muon detector data reception in L0Muon	$1.7 \mu\text{s}$
Precision Muon detector data reception in L0Muon	$2.8 \mu\text{s}$
Deadtime	$<0.1\%$ per detector system

and propagated up to the central trigger. Details of the timing and control signals which will be distributed, along with all possible readout architecture that will co-exist in ATLAS, are described in the Detector FE Interface Requirement Document [4.2].

The on-detector electronics will always require the Clock and ORBIT signals, even if they get L0A signals as well. All event fragments carry Bunch Crossing Identifier (BCID) information that is derived from these signals. Detectors read out on Level-0 can also have a Level-0 identifier. Basically, all detectors send data off detector at the bunch-crossing rate of 40 MHz, except for the ITk strips and pixels due to technical and financial reasons, and the NSW because of constraints from the legacy system. Readout to FELIX from the FE ASICs (ITk and NSW) or from off-detector electronics (all other detectors) is generally done after receiving the L0A signal.

### Hardware Requirements

The hardware requirements for the Level-0 trigger sub-systems are specified as follows:

- **Level-0 Trigger (L0)** Trigger sub-systems shall be specified and implemented assuming ATCA as hardware platform.
- Blade design shall be compatible with PICMG-3.0 Revision-3.0 standard.
- No more than 4 mezzanines compatible with PICMG AMC.0 R2.0 specifications shall be accommodated on a blade.
- CERN-developed **IPMI Management Controller (IPMC)** shall be used for basic blade control and configuration at power up through the ATCA shelf manager.

- 
- Blade configuration and control shall be implemented through a standard Network interface using a [System-on-Chip \(SoC\)](#)-based solution as detailed in Chapter 15. The configuration of some of the Phase-I upgrade subsystems (e.g. [eFEX](#)) is implemented through the IPBUS protocol, which shall be supported during the [HL-LHC](#) lifetime.
  - [DCS](#) interfaces shall be implemented through a standard Network interface using a CERN-developed [IPMC](#) for basic blade monitoring and control at power up through the [ATCA](#) shelf manager, and a [SoC](#)-based solution as mentioned above.
  - Shelf configuration shall not exceed 2 Hub slots and 12 node slots for a total maximum allowed power dissipation of 7.5 kW including the shelf manager, switch and cooling fans at maximum power.
  - [ATCA](#) blade power consumption shall not exceed 400 W.
  - [Rear Transition Module in the HTT \(RTM\)](#) module shall not exceed 50 W in power consumption.

### 4.3 DAQ Requirements

The [DAQ](#) system will receive and deliver data via links with protocol and bandwidth suitable for individual detector subsystems. The capacity should accommodate the significantly increased trigger rate of 1 MHz, and the much larger event size of about 5.2 MB, with enough flexibility to allow detector subsystems to operate at different event rates and to adjust data content with different run conditions.

Overall standardisation of all basic common elements will minimise cost, maintenance and evolution over the 10–20 years of operation in the [HL-LHC](#) period. Effort is required from an early stage to decompose the project into functional elements and their interfaces and to define standard common elements, interfaces and protocols for each logical element.

- The [DAQ](#) system shall be scalable with respect to the trigger rate and the luminosity, capable of handling detector data input of about 5.2 TB/s. The system should take advantage of the technology advances in commodity information technology software and hardware, given the advantages of being cheaper and requiring less maintenance than any custom hardware.
- The Readout and Dataflow subsystems shall sustain high throughput combining the input from detector front-end electronics (up to about 5.2 TB/s) and the output to the Event Filter (up to about 2.6 TB/s) at the required trigger rate, and be able to deal with high rate fluctuations without causing large congestion and data loss.
- The Readout subsystem shall be designed to have a common interface to the detector front-end that accommodates the varieties of protocols and bandwidths of the front-end links of each subdetector system. In particular, it should deliver control and configuration information to the detector front-end via the radiation tolerant [lpGBT](#) link (uplink bandwidth of 9.6 Gb/s and downlink bandwidth of 2.5 Gb/s), and read out data with the [lpGBT](#) protocol or other specific protocols suitable for individual

subdetector systems, e.g., dedicated protocol to be developed for [ITk Pixel](#) and the Full Mode protocol for [LAr](#) trigger electronics installed in USA-15.

- The Readout subsystem interface to the detector front-end shall route raw data for further processing via a standard common network protocol, access to which should make it possible for detector-specific data handling and monitoring to be performed.
- The system shall be capable of executing also detector-specific control and configuration operations, interfacing to the common ATLAS [DCS](#) system.
- It is essential to have sufficient storage to enable the dynamic decoupling of real data taking and the Event filter processing.
- Based on operational experience from Run 1 and Run 2 the system has to be able to operate for up to 48 hours without access to the offline permanent storage. At the average throughput of about 60 GB/s of uncompressed data this requires a storage volume of 10 PB.

## 4.4 Event Filter Requirements

The [EF](#) will be designed to work at a luminosity of  $7.5 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ . The scale of the [EF](#) compute requirements derive from this as they depend on the input rate, which is the accept rate of the Level-0 trigger, and the level of pile-up. There is a close coupling between the scale of the [Dataflow](#) and [EF](#) systems in this respect.

- The [EF](#) system shall be designed to sustain a maximum input rate of 1 MHz.
- The [EF](#) system shall be designed to select events with a maximum output rate of 10 kHz.
- The [EF](#) trigger efficiency and rejection requirements for all signatures are derived from the physics requirements described in Section 2.
- [EF](#) algorithms and selections shall be designed to maximise efficiency with respect to offline reconstruction and minimise systematic bias from the trigger.
- Many of the initial fast rejection techniques presently used in the Run 2 [HLT](#) are implemented in the Level-0 hardware trigger in Phase-II. Therefore, more sophisticated algorithms, identical or as close as possible to the offline reconstruction methods, are needed earlier in the [EF](#) selection. Where necessary innovative approaches shall be developed to reach the required performance and fulfill the time constraints.
- Experience from offline reconstruction shows that the effects of pile-up on reconstruction resolution and quality can be mitigated by using vertices and tracks. This is especially important for hadronic triggers, for which tracks were not used in Run 1 or Run 2 for the purpose of pile-up suppression due to processing or readout constraints. The [EF](#) must be capable of reconstructing vertices and tracks as needed for pile-up suppression.
- Reconstruction algorithm times tend to increase with luminosity, in some cases much worse than linearly, posing a tough challenge to the [EF](#) software. The model that has

---

been developed to estimate the compute requirements at  $\langle\mu\rangle = 200$  assumes that most time-consuming algorithms can achieve linear scaling and speed can be improved over today's measured performance, so these become requirements.

- In particular, tracking in a high pile-up environment is computationally expensive: given the above need for tracking, there is a strong requirement to reduce this cost through both improvements to software tracking and some form of hardware tracking.
- **EF** selection shall follow the principles of early rejection and delaying full event data access, to reduce the event building rate in the **Dataflow** system and the volume of data that the **Storage Handler (Storage Handler)** must hold, to below the need for full event at 400 kHz.
- Event data unpacking in the **EF** is time critical: the data format and unpacking code need to be developed very closely between the detector and **EF** experts.
- Commodity processing technology will continue to evolve over the next 10 years. The **EF** software shall be prepared to make efficient use of many-core processors, which is the current industry direction. Note that the Phase-I upgrade will be an excellent starting point from which to optimise this for Phase-II.

## Hardware Tracking Requirements

There are two complementary tracking requirements for the baseline **TDAQ** system: fast regional tracking and more precise global tracking.

- Regional tracking is required at the full Level-0 rate of 1 MHz to provide tracks for fast initial rejection of single high- $p_T$  lepton and multi-object triggers in the **EF**. The tracks will be used in combination with information from the **Global Trigger**. Tracks are needed for all charged particles with  $p_T > 2$  GeV and  $|\eta| < 4.0$ . Regions will be defined in the **EF** based on objects identified by the Level-0 trigger and, on average, will represent less than 10% of the full event.
- Global tracking for  $b$ -jet identification requires reconstruction of all charged particles with  $p_T > 1$  GeV in an event within  $|\eta| < 4.0$ , with track quality and minimum  $p_T$  similar to offline track reconstruction [4.3].
- Global tracking for soft jets, calculation of the  $E_T^{\text{miss}}$  soft term and pile-up correction and mitigation requires reconstruction of all charged particles with  $p_T > 1$  GeV in an event within  $|\eta| < 4.0$ .
- The global tracking  $d_0$  and  $z_0$  resolutions shall be no more than twice the respective offline resolutions with digital clustering, and the fake rate shall be at the percent level or lower.

The following requirements are related to the interface between the **EF** and **ITk**.

## 4.4 Event Filter Requirements

---

- For regional tracking requests, the **EF** will determine the regions to be requested from the Level-0 data and will map these to parts of the **ITk** data and the relevant **HTT** units.
- The **EF** will obtain and send the relevant **ITk** data as part of the hardware tracking request.
- A **HTT** interface is required to accept requests for regional and full tracking from **EF** processes and relay them synchronously to the relevant boards in the associated **HTT** unit.

## References

- [4.1] ATLAS Collaboration, *ATLAS Phase-II Upgrade Scoping Document*, Tech. Rep. CERN-LHCC-2015-020. LHCC-G-166, CERN, Geneva, Sep, 2015. <https://cds.cern.ch/record/2055248>.
- [4.2] F. Lanni et al., *Interfaces with Detector Front-End Systems. Requirement Document for HL-LHC*, Tech. Rep. ATL-DA-ES-0051 v.1, EDMS Id 1563801 v.1, CERN, Geneva, March, 2016. <https://edms.cern.ch/document/1563801>.
- [4.3] ATLAS Collaboration, *Technical Design Report for the ATLAS Inner Tracker Strip Detector*, Tech. Rep. CERN-LHCC-2017-005, CERN, Geneva, Apr, 2017. <https://cds.cern.ch/record/2257755>.
- [4.4] ATLAS Collaboration, *Technical Design Report for the ATLAS Inner Tracker Pixel Detector*, Tech. Rep. CERN-LHCC-2017-021, CERN, Geneva, (in preparation).
- [4.5] ATLAS Collaboration, *Technical Design Report for the Phase-II Upgrade of the ATLAS Muon Spectrometer*, Tech. Rep. CERN-LHCC-2017-017. ATLAS-TDR-026, CERN, Geneva, Sep, 2017. <https://cds.cern.ch/record/2285580>.
- [4.6] ATLAS Collaboration, *ATLAS New Small Wheel Technical Design Report*, Tech. Rep. CERN-LHCC-2013-006. ATLAS-TDR-020, CERN, Geneva, Jun, 2013. <https://cds.cern.ch/record/1552862>.

## 5 Description of the Baseline System

This Chapter provides a functional overview of the systems within the [TDAQ](#) Phase-II Upgrade Project: the Level-0 Trigger System, the Data Acquisition System ([DAQ](#)), and the Event Filter ([EF](#)) System. An overview of the subsystems within each system is presented here. Detailed descriptions of the system components and hardware design considerations are presented in dedicated chapters in Part II of this document.

### 5.1 Functional overview

The functional overview of the baseline [TDAQ](#) system planned for Phase-II is shown in [Fig. 5.1](#).

The data inputs from the detectors are categorised in the figure as Inner Tracker (which includes the [ITk](#) pixel and strip detectors), Calorimeters (the [LAr](#) and [Tile](#) calorimeters), and Muon System (including [MDT](#), [RPC](#), [TGC](#) and [NSW](#)). Detector-specific off-detector electronics for both readout and trigger data preparation are included as part of the detector systems.

The Level-0 Trigger system is made up of the [L0Calo](#), [L0Muon](#), [Global Trigger](#), and Central Trigger subsystems. [L0Calo](#) and [L0Muon](#) are similar in functionality to their Phase-I predecessors; they use calorimeter and muon system information at 40 MHz, respectively, to apply an initial event selection and to identify features to be examined at the subsequent trigger level. The [L0Calo](#) system is based on the legacy Phase-I [L1Calo](#) system with minor changes, and uses coarse-granularity calorimeter data to identify electron, tau lepton, and jet candidates, and to calculate  $E_T^{\text{miss}}$ . A dedicated “forward” Feature Extractor ([ffEX](#)) is planned to ensure efficient electron identification in the region  $3.2 < |\eta| < 4.0$ . Each component of the [L0Muon](#) subsystem is upgraded compared to the corresponding system in Phase-I, receiving data from all of the muon subsystems and a subset of the [Tile](#) Calorimeter information to identify muon candidates. New features of the [L0Muon](#) system for Phase-II are the inclusion of precision [MDT](#) momentum measurements as well as new [RPC](#) inner stations to improve the muon trigger coverage. The [MUCTPI](#) provides an interface between the barrel and endcap components of the [L0Muon](#) system on one hand, and the [Global Trigger](#) and [CTP](#) on the other. It identifies muon candidates that have been counted twice in the [L0Muon](#) system (overlap removal) and calculates multiplicities for various transverse-momentum thresholds.



## 5.1 Functional overview

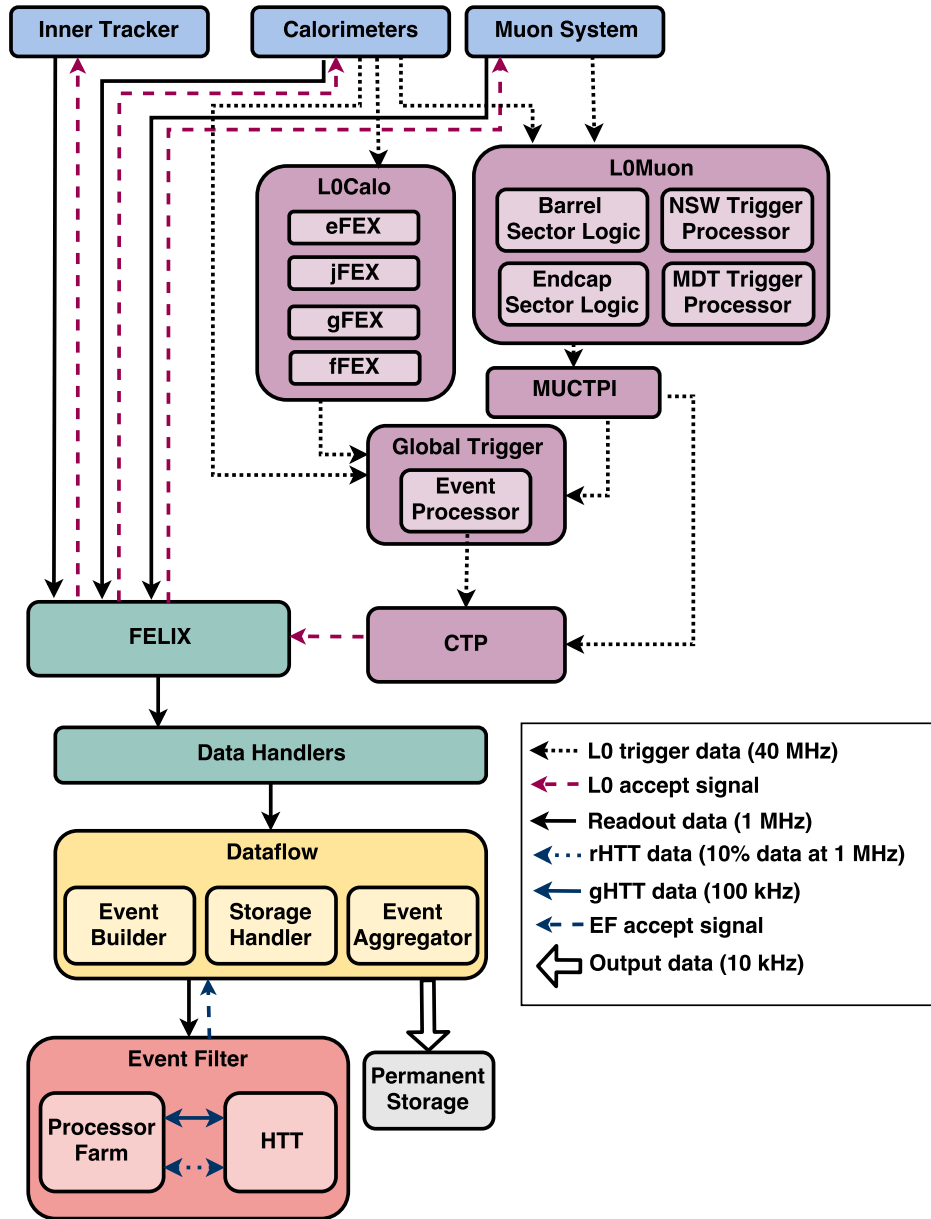


Figure 5.1: Overall baseline design of the *TDAQ* system in Phase-II. The black dotted arrows indicate the Level-0 dataflow from the detector systems to the Level-0 trigger system (composed of the *LOCalo*, *L0Muon*, *MUCTPI*, *Global Trigger*, and *CTP*) at 40 MHz, which must identify physics objects and calculate event-level physics quantities within 10  $\mu$ s. The result of the Level-0 trigger decision (*LOA*) is transmitted to the detectors as indicated by the red dashed arrows. The resulting trigger data and detector data are transmitted through the *DAQ* system (made up of *FELIX*, the *Data Handlers*, and the *Dataflow* subsystem) at 1 MHz, as shown by the black solid arrows. Direct connections between each Level-0 trigger component and the Readout system are suppressed for simplicity. The *EF* system is composed of a processing farm and a hardware-tracking subsystem (*HTT*) that must reduce the event rate to 10 kHz. Events that are selected by the *EF* trigger decision are transferred for permanent storage.



---

The [Global Trigger](#) is a new subsystem of the Level-0 Trigger system, which will perform offline-like algorithms on full-granularity calorimeter data, for example using information from [LAr](#) first-layer strips to better reject  $\pi_0$  in  $e/\gamma$  identification and making topological clusters for refined jet algorithms (anti- $k_t$ ) and pile-up subtraction. It will also identify topological signatures that can include a wide variety of four-vector combinations involving sums, angles and invariant masses. The [L0Calo](#) and [L0Muon/MUCTPI](#) subsystems send their selected objects to the [Global Trigger](#), including spatial locations, reconstructed energy/momentum values and discriminant variables. These objects are then combined with the results of the [Global Trigger](#) calorimeter processing to refine the  $e/\gamma$ , tau, muon and jet selections. The [CTP](#) forms the final Level-0 decision, taking into account the trigger menu configuration, prescale factors, and dead-time requirements. This decision is transmitted as a [LOA](#) signal via [LTIs](#), the [TTC](#) distribution network, and [FELIX](#), to the detector systems.

Following the Level-0 trigger decision, detector data are transmitted over custom point-to-point serial links to the Front-End Link eXchange ([FELIX](#)) subsystem, the first element of the [Readout](#) subsystem within the [DAQ](#) system. [FELIX](#) provides a common interface between the detector-specific custom point-to-point serial-links and the commodity multi-gigabit data network downstream. Along the network, data are received by the [Data Handlers](#), where detector-specific processing, e.g. formatting and/or monitoring, can be implemented before buffering data in the [Dataflow](#) subsystem. The [Readout](#) subsystem is designed to handle 1 MHz event rate, for a total bandwidth of 5.2 TB/s. The [Dataflow](#) subsystem buffers, transports, aggregates, and compresses event data for utilisation in the [EF](#) System, which is effectively shielded from the underlying distributed storage system. Data are buffered here while waiting for the [EF](#) selection results.

A large [EF](#) processor farm is needed to cope with the 1 MHz input rate; commodity-CPU-based event processing will be aided by the new [HTT](#) subsystem, which is designed to provide fast hardware-based track reconstruction. Regional tracking allows a fast initial rejection in the [EF](#) of single high- $p_T$  lepton and multi-object triggers from background processes, to reduce the rate to around 400 kHz. This system is specified to operate at 1 MHz and use up to 10% of the [ITk](#) data, by selecting tracking modules in regions based on the results of the Level-0 trigger system. Software-based reconstruction will follow to achieve further rejection. This will be aided by global tracking at around 100 kHz using the [HTT](#) again, but this time to produce tracks closer to offline quality, suitable for  $b$ -jet tagging,  $E_T^{\text{miss}}$  soft term calculation, soft jets and pile-up suppression. The rates of Level-0 triggers for which regional or global tracking are needed are shown in Section 6.11.

Events selected by the [EF](#) are then transferred to the permanent storage of the ATLAS offline computing system. The raw output event size is expected to be 6 MB, and the total trigger output is expected to be 10 kHz; thus, the total bandwidth out of the system is 60 GB/s.

## 5.2 The Level-0 Trigger System

A diagram illustrating the details of the Level-0 Trigger System is shown in Fig. 5.2. The functionality and hardware implementation of each Level-0 Trigger subsystem ([L0Calo](#), [L0Muon](#), [Global Trigger](#), and [Central Trigger](#)) are summarised briefly in this section.

### 5.2.1 Level-0 Calorimeter Trigger

Table 5.1 contains the list of all Level-0 Calorimeter trigger components in ATLAS showing, per subsystem, the list of Trigger Objects identified, the approximate detector granularity used and the coverage region in  $|\eta|$ .

Table 5.1: *List of Level-0 Calorimeter trigger components showing, per subsystem, the Trigger Objects identified, the detector granularity used and the  $\eta$  coverage. Detector granularities used include either Full detector cells, Super Cells or coarser  $\Delta\eta \times \Delta\phi$  granularity.*

Subsystem	Trigger Object	Approximate Granularity	Coverage $ \eta $
<a href="#">eFEX</a>	$e/\gamma, \tau$	Super Cells (10 in $0.1 \times 0.1$ )	$< 2.5$
<a href="#">jFEX</a>	$\tau, \text{jet}, E_T^{\text{miss}}$	$0.1 \times 0.1$	$< 2.5$
<a href="#">jFEX</a>	$\tau, \text{jet}, E_T^{\text{miss}}$	$0.2 \times 0.2$	$2.5 - 3.2$
<a href="#">jFEX</a>	$\tau, \text{jet}, E_T^{\text{miss}}$	$0.4 \times 0.4$	$3.2 - 4.9$
<a href="#">gFEX</a>	Large-R jet, $E_T^{\text{miss}}$	$0.2 \times 0.2$	$< 4.9$
<a href="#">ffEX</a>	$e/\gamma$	Full detector <a href="#">EMEC</a> , <a href="#">HEC</a> , <a href="#">FCal</a>	$2.5 - 4.9$
<a href="#">ffEX</a>	jet	Full detector <a href="#">FCal</a>	$3.2 - 4.9$

[L0Calo](#) is based on the Phase-I [L1Calo](#) system described in the [TDAQ](#) Phase-I TDR [5.1]. In this system, separate [LAr](#) and [Tile](#) calorimeter data streams with coarse granularity (compared to offline) are sent to [FPGA](#)-based Feature Extractors ([FEXs](#)), which find electron, photon, and tau lepton candidates ([eFEX](#)), tau leptons, jets and  $E_T^{\text{miss}}$  ([jFEX](#)), and large-R jets and  $E_T^{\text{miss}}$  ([gFEX](#)). The [LAr](#) Calorimeter inputs to the system are maximally ten [super cells](#) per region of  $0.1 \times 0.1$  in  $\eta \times \phi$  for  $|\eta| < 2.5$ , and have coarser granularity for larger  $|\eta|$  values. The [Tile](#) Calorimeter inputs to the system are  $0.1 \times 0.1$  in  $\eta \times \phi$ , summed in depth. The output of the [L0Calo](#) subsystem is a set of electron, photon, tau lepton, jet, and  $E_T^{\text{miss}}$  trigger objects that are sent to the [Global Trigger](#). The [ffEX](#) will handle full granularity [LAr](#) in the region  $2.5 < |\eta| < 4.9$  ( $3.2 < |\eta| < 4.9$ ) and reconstruct forward electromagnetic (forward jet) trigger objects.

The interfaces to [L0Calo](#) from other detector systems and from [L1Calo](#) to other components of [TDAQ](#) are the following:

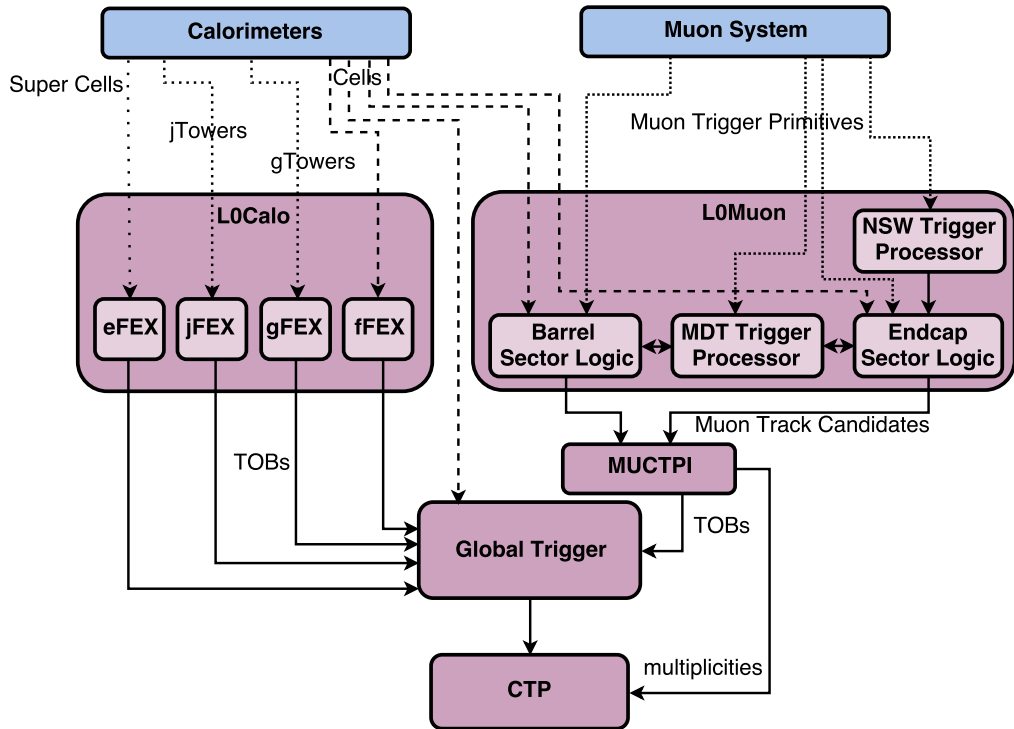


Figure 5.2: Diagram of the Level-0 Trigger architecture. The arrows indicate the type of object that is transmitted between each component of the system. Super cell energies are sent to the *eFEX*, while coarser granularity, based on sums of super cells energies, are sent to the *jFEX* and *gFEX*, respectively, as is done in the Phase-I system. Information from calorimeter cells with  $|\eta| > 2.5$ , without an energy threshold, is sent to the *fFEX*. Tile calorimeter information from the outermost layer is used in coincidence with muon trigger primitives from the *RPC* and *TGC* in the *Barrel* and *Endcap Sector Logic* components, respectively. Hit information from the *NSW* and *MDT* are sent to their corresponding Trigger Processors. The *MDT* Trigger Processor refines the momentum measurement for muon candidates determined by the Sector Logic components; the resulting muon candidates are transmitted to the *MUCTPI* for overlap removal and multiplicity determination. Elementary calorimeter cells above a transverse energy threshold of  $|E_T| > 2\sigma$  are sent to the *Global Trigger* for use in topological clustering and other refined algorithms. *Trigger Objects (TOBs)* are formed by each of the *FEXs*, as well as the *MUCTPI*, which are refined by the *Global Trigger*. The *CTP* combines information from the *Global Trigger* and *MUCTPI*, consisting of trigger algorithm flags and multiplicities of selected objects, to make the final Level-0 trigger decision.

- *LAr* EM, *EMEC*, *HEC* calorimeter pre-processors called *LAr Digital Processing System (LDPS)* and *LASP*, carrying Super Cell and Full granularity detector data to *L0Calo*, respectively.
- *Global Trigger*, to which Trigger Objects are sent for refined processing.

## 5.2 The Level-0 Trigger System

- **CTP**, to be used for commissioning of **L0Calo** before the **Global Trigger** final commissioning has been completed.
- **Readout** System, carrying input, output and intermediate processing data through **FELIX**.
- Timing and Control distribution upstream and downstream information through **FELIX**.
- Control and Configuration via commodity network (Ethernet).
- **DCS** and Monitoring information through **FELIX**.
- **DCS** interface, common to all **ATCA** hardware.

The system uses **ATCA** blades as the standard processing platforms. Dedicated hardware is designed and already deployed in the Phase-I upgrade for **eFEX**, **jFEX** and **gFEX** subsystems, while a new design is foreseen for the **fFEX**. The full Level-0 Calorimeter trigger is deployed in 5 **ATCA** shelves, two for the **eFEX**, one for the **jFEX**, one for the **gFEX** and one for the **fFEX** subsystems.

### 5.2.2 Level-0 Muon Trigger

Table 5.2 contains the list of all the Level-0 Muon trigger components in ATLAS showing, per subsystem, the detector data used and coverage region in  $|\eta|$ .

Table 5.2: List of all Level-0 Muon trigger components showing, per subsystem, the list of Trigger Objects processed and identified. All subsystem use information from multiple detectors: Muon Barrel uses **RPC**, **MDT** and Tile Outer layer, while the Muon Endcap uses **TGC**, **MDT**, **NSW**, **RPC** and Tile outer layer.

Subsystem	Granularity	Coverage $ \eta $
<b>NSW</b> processor	Full <b>NSW</b> detector	1.3 – 2.4
<b>MDT</b> processor	Full <b>MDT</b> detector	< 2.4
Barrel Sector Logic	Full <b>RPC</b> and Tile, <b>MDT</b>	< 1.05
Endcap Sector Logic	Full <b>TGC</b> , Tile, <b>RPC</b> , <b>NSW</b> , <b>MDT</b>	1.05 – 2.4

The entire muon trigger electronics and readout chain will be replaced in Phase-II, except for some parts of the on-detector front-end electronics<sup>1</sup>. All data from the barrel (**RPC**) and endcap (**TGC**) detectors will be transmitted off-detector where the full information will be available for trigger processing. In such a scheme, more refined and flexible algorithms can be used compared to the present ones based on simple coincidences, which are limited by on-detector connectivity. In the **NSW** detector region ( $1.3 < |\eta| < 2.7$ ), both the **MM** and **sTGC** chambers take the dual role of trigger and precision chambers. Muon trigger

<sup>1</sup> The **TDQAQ** responsibility begins at the **USA15** counting room. Upstream components are the responsibility of the Muon detector community.

primitives (track segments) from these detectors will be combined with [TGC](#) hits to provide the endcap trigger candidates in the region  $(1.3 < |\eta| < 2.4)$ <sup>2</sup>.

The Barrel Sector Logic board will take input data at each bunch crossing from the [RPC](#) on-detector electronics and the digital readout from the outermost [Tile](#) Calorimeter cells. The Endcap Sector Logic board will take input data at each bunch crossing from the [TGC](#) on-detector electronics, the [NSW MM](#) and [sTGC](#) chambers via the [NSW](#) Trigger Processor, digital readout from the outermost [Tile](#) Calorimeter cells, and the thin-gap [RPC](#) triplet in the barrel-endcap transition region. Both barrel and endcap muon candidates are formed in [FPGA](#)-based Sector Logic boards; the results are sent to the [Global Trigger](#) via the [MUCTPI](#). After a [LOA](#), the digital readout is sent to the Readout subsystem via [FELIX](#).

The [MDT](#) readout chain will be replaced, including the front-end cards (mezzanine cards) where the hit buffers reside. This replacement, along with the increased [L0](#) latency, also presents the opportunity to use in the [L0](#) trigger the precision coordinates measured by the [MDT](#) to improve the quality of trigger candidates provided by the [RPCs](#) in the barrel and by the [TGCs](#) plus [NSW](#) in the endcaps. The [MDT](#) processing will be seeded by the Barrel and Endcap Sector Logic boards, which will identify potential muon candidates. Pattern recognition and tracking will follow, carried by [MDT](#)-dedicated processors. The refined muon candidate will be sent back to Barrel and Endcap Sector Logic modules for the final selection of candidates to be sent to the [MUCTPI](#) subsystem. Two implementations are being considered for the hit pattern matching functionality of the [MDT](#) Trigger Processors: a baseline design using [FPGAs](#) and an alternative one using [AM](#) chip. The design considerations for these potential implementations are described in more detail in Section 8.6.

The [L0Muon](#) interfaces from other detector systems and to other components of [TDAQ](#) are the following:

- the detector front-end links for all muon detectors, [RPC](#), [TGC](#), [MDT](#), [NSW](#), carrying full granularity data to [L0Muon](#) processors;
- [Tile](#) pre-processor links, carrying muon tag information to the Barrel and Endcap Sector Logic boards;
- the [MUCTPI](#), which is the interface between the [L0Muon](#) system and the [Global Trigger](#) and [CTP](#);
- the [Readout](#) System, transmitting full detector data, input, output and intermediate trigger processing data through [FELIX](#);
- the timing and control distribution (upstream and downstream) through [FELIX](#);
- the control and configuration interface via the commodity network (Ethernet);
- the [DCS](#) and monitoring information via [FELIX](#); and
- the [DCS](#) interface, which is common to all [ATCA](#) hardware.

The system uses [ATCA](#) blades as the standard processing platforms. New dedicated hardware is designed for the Phase-II upgrade. The same [ATCA](#) blade design is foreseen to be

<sup>2</sup> Studies quantifying the implications of extending the trigger coverage for muons in the region  $2.4 < |\eta| < 2.7$  are ongoing. Muon trigger coverage out to  $|\eta|$  of 2.4 is the baseline configuration.

used for Barrel and Endcap Sector Logic, while a dedicated blade design is foreseen for [MDT](#). It is not yet clear whether [NSW](#) will use the existing Phase-I [ATCA](#) blade or a new one: some work is outstanding before the baseline solution is chosen. The full Level-0 Muon trigger is deployed in 15 [ATCA](#) shelves, 2 dedicated to the [NSW](#) processors, 6 dedicated to the [MDT](#) processors, 3 for the Barrel and 4 for the Endcap Sector Logic.

### 5.2.3 MUCTPI

The [MUCTPI](#) aggregates and merges the trigger information from the barrel and endcap muon systems before passing it to the [Global Trigger](#) and the [CTP](#).

The [MUCTPI](#) interfaces only to other components of [TDAQ](#), as follows:

- the Barrel and Endcap Sector Logic modules, from which muon candidates are received for further processing, e.g. overlap removal between the barrel and endcap Sector Logic modules;
- the interface to the [Global Trigger](#), to which muon candidates are transmitted;
- the interface to the [CTP](#), to which muon multiplicity information for various  $p_T$  thresholds is transmitted;
- the interface to the Readout System, transmitting input, output and intermediate trigger processing data through [FELIX](#);
- the timing, trigger and control distribution;
- the control and configuration interface via the commodity network (Ethernet);
- the [DCS](#) and monitoring information via [FELIX](#); and
- the [DCS](#) interface, which is common to all [ATCA](#) hardware.

The system uses [ATCA](#) blades as the standard processing platforms. The current baseline foresees the use of two blades designed for the Phase-I upgrade. The [MUCTPI](#) sub-system is deployed in one [ATCA](#) shelf.

### 5.2.4 Global Trigger

The Level-0 trigger functionality of the [Global Trigger](#) complements the [L0Calo](#) trigger objects with additional high-granularity energy data coming directly from upgraded calorimeter pre-processors (the [LASP](#) and [TPPr](#)). This allows the implementation of offline-like algorithms such as: topological clustering, refined electron and photon identification, tau lepton identification, lepton isolation, calorimeter-based pile-up suppression, sophisticated jet-finding, and dedicated exotic-object selection. Furthermore, the [Global Trigger](#) will replace and extend the functionality of the [L1Topo](#) system, applying topological selections (such as angular requirements) to trigger objects.

The hardware implementation of the **Global Trigger** consists of three primary components: a **Multiplexer Processor (MUX)** layer, a **GEP** layer, and a demultiplexing **Global-to-CTP Interface (CTP Interface)**, all of which have identical hardware composed of **ATCA** blades and **FPGAs** with many multi-gigabit transceivers. The calorimeter detector subsystems, **FEXs**, and **MUCTPI** provide serial data for each bunch crossing to the **MUX** layer. These signals are then time-multiplexed and the signals for a given event are transported to a single **GEP** node that executes the algorithms described above. Finally, the results are sent to the **CTP** through the **CTP Interface**. There are two main advantages of the time-multiplexed architecture. First, the event processor is decoupled from the **LHC** bunch-crossing rate, allowing the use of asynchronous and high-level algorithms that are impossible in the Phase-I hardware trigger. Second, this implementation removes the limitation on the number of **TOBs** from the **FEXs** that are present in the Phase-I hardware trigger. Common hardware modules are implemented across the **Global Trigger** system, minimising the complexity of the firmware and simplifying the system design and long-term maintenance.

The interfaces between **Global Trigger** and other detector systems and components of **TDAQ** are as follows:

- the interfaces from the calorimeter front-end electronics, which send full-granularity cell energies ( $|E_T| > 2\sigma$ ) every **LHC** bunch crossing;
- the interfaces from the **L0Calo FEXs**, which send **TOBs** every **LHC** bunch crossing;
- the interface from the **MUCTPI**, which sends **TOBs** every **LHC** bunch crossing;
- the interface to the **CTP**, which receives a 1024-bit **TIP** (Trigger Inputs) consisting of flags and multiplicity values from the **Global Trigger** algorithms every **LHC** bunch crossing;
- the interfaces to the **FELIX** for the readout of **TIPs**, trigger algorithm results, and error reporting;
- the interface from the **FELIX** containing clock, configuration, per-event **TTC** information, and event accept signals; and
- the **DCS** interface that is common to all **ATCA** hardware.

For every **L0A** the **Global Trigger** will produce a compact data structure mapping the **TIP** onto the **TOBs** which caused those **TIP** bits to be fired; this will be included in the readout data.. All intermediate objects produced by the algorithms will also be read out, but as with the **L0Calo FEX** modules, the full input data will in general not be read out.

## 5.2.5 Central Trigger Processor

The **CTP** makes the final **L0A** decision, aligning and combining digital trigger inputs from the **Global Trigger**, **MUCTPI**, the legacy Phase-I **L1Topo** system, and various forward detectors and subdetector calibration systems, as well as introducing deadtime and applying prescales as required. The **L0A** signal is then transmitted to the subdetectors via the **LTIs**.

The **CTP** interfaces only to other components of **TDAQ**, as follows:



- optical inputs from the [MUCTPI](#) and legacy Phase-I [L1Topo](#) system;
- optical inputs from the [Global Trigger](#), receiving trigger algorithms flags to be processed for the final Level-0 decision;
- electrical trigger signals from various forward detectors and subdetector calibration systems;
- the interface to the Readout System, transmitting input, output and intermediate trigger processing data through [FELIX](#);
- the timing and control distribution to the detectors and all [TDAQ](#) subsystems through the [LTIs](#);
- the [LHC](#) machine interface for the reception of the clock and all [LHC](#)-related signals;
- the control and configuration interface via the commodity network (Ethernet);
- the [DCS](#) and monitoring information through [FELIX](#); and
- the [DCS](#) interface, which is common to all [ATCA](#) hardware.

The system uses [ATCA](#) blades as the standard processing platforms. The current baseline foresees the use of three types of blades designed specifically for the Phase-II upgrade, one dedicated to the machine interface, one input blade for all other inputs and a core processor blade. The full [CTP](#) is deployed in one [ATCA](#) shelf.

### 5.2.6 Trigger, Timing and Control System

The [TTC](#) network transmits the 40 MHz beam synchronous clock with low jitter, the [LOA](#) decision with fixed latency, as well as additional timing information including the synchronous turn signal (bunch counter reset), and other synchronous or asynchronous timing and control signals. In the opposite direction, Busy signals from the subdetector back-end electronics are collected via the [TTC](#) Busy network. The detailed [TTC](#) information is described in the Detector–TDAQ Interface document [5.2].

As shown in Fig. 5.3, the [TTC](#) network has a tree-like structure starting from the [CTP](#), where there is one optical output link per subdetector partition. The [CTP](#) receives the relevant timing signals from the Machine Interface Module. The point-to-point output links connect to [LTI](#) boards close to the subdetector [FELIX](#) components. The [LTIs](#) provide an interface for the [TTC](#) signals between the [CTP](#) and subdetector front-end electronics via [FELIX](#). In addition, the [LTI](#) modules can mimic the [CTP](#) functionality for standalone subdetector running that is used for detector development, commissioning, and calibration. They can also store and send detector commands and data for configuration and calibration.

From the [LTIs](#), [PONs](#) implement a Point-to-Multipoint architecture in which fibre optic splitters are used to enable a single optical fibre to serve multiple end-points. A [PON](#) consists of a top node, called [Optical Line Terminal \(OLT\)](#), one or more destination nodes, called [Optical Network Units \(ONUs\)](#) and the fibres and splitters between them. Time-division multiplexing is implemented to collect the Busy signals from the [FELIX](#) I/O cards.

The [TTC](#) interfaces to detectors and to other components of [TDAQ](#) are as follows:



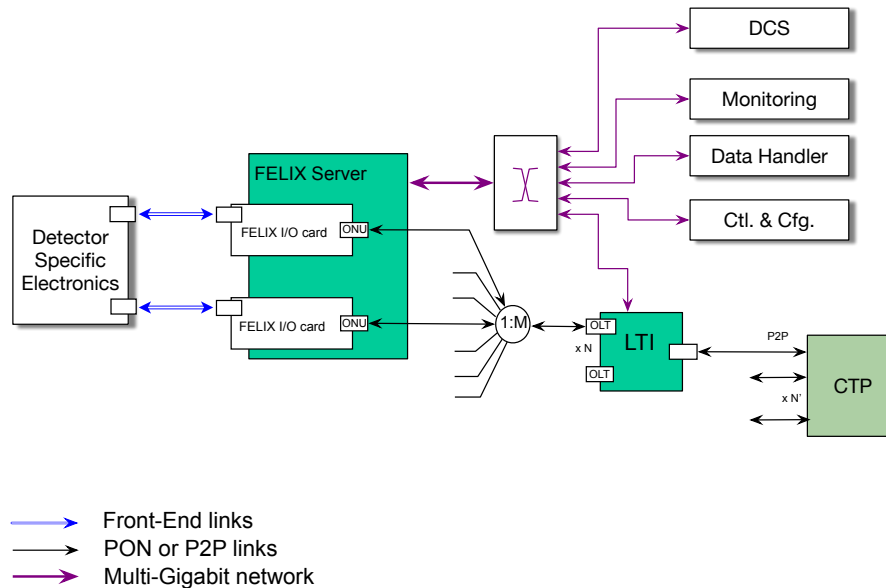


Figure 5.3: The Level-0 Timing and Control paths. Point-to-Point (P2P in the figure) and *PON* links are used in the *TTC* network.

- the *LTI PON* interfaces and optical fibre trees, to *FELIX* or detector-specific clients;
- an interface from the *CTP*, receiving all relevant *TTC* signals;
- electrical trigger signals from the local *TTC* inputs, to be used for commissioning purposes or during tests;
- the control and configuration interface to subsystem specific elements via the commodity network (Ethernet);
- the *LTI* control and configuration interface via the commodity network (Ethernet);
- and
- the *DCS* interface, which is common to all *ATCA* hardware.

The system uses *ATCA* blades as the standard processing platforms. The current baseline foresees the use of one *LTI* blade designed specifically for the Phase-II upgrade. The number of *LTI* blades depends on the detector partitioning; based on experience from Run 2, we currently assume a total of 36 units will be distributed in shelves belonging to detector or *TDAQ* subsystems.

### 5.2.7 Technical implementation and system size

The Level-0 system design aims to maximise the processing away from the detector in order to ease maintenance, accessibility and the possibility of future upgrades. In addition, the number of boards to be designed is minimised as much as possible. Designing and writing firmware will be managed similarly to designing and building hardware. Interfaces on the boundaries of subsystems are defined at an early stage in order to optimise the partitioning and design efforts of the many groups participating in the construction. [ATCA](#) technology is adopted as the baseline for all detector and trigger electronics, and [PCIe](#) cards are used when it is necessary to interface to high performance networks, conforming to industry standards. Point-to-point high speed links and [PON](#) technology are used for all trigger data transmission and broadcasting needs.

The overall size of the Level-0 Trigger System, in terms of the number of boards and number of links corresponding to each hardware component, is summarised in [Table 5.3](#). For these components, the output links to the readout path are not included. This table provides also a list of components which are delivered and commissioned in Phase-I and for which no additional hardware design is needed. For a given component, the number of boards in [Table 5.3](#) only describes the size needed on the ATLAS experiment. [Table 5.4](#) shows the physical shelf space needs of the Level-0 Trigger system and the power requirements which shall be satisfied for the system installation in [USA15](#); the power estimates are the maximum possible values, which include uncertainties and margins; they are 20% more than the current best estimates.

Common solutions are foreseen, where [ATCA](#) blade designs are used in many components of the trigger system. For example, the [L0Muon](#) Barrel and Endcap Sector Logic components will use the same hardware module. The same solution is adopted in the [Global Trigger](#) system, where a common module is used for the time-multiplexed data distribution, for processing units and for interfacing to the [CTP](#). In other cases this approach is not viable. The [NSW](#) and [MDT](#) Trigger Processors require different hardware (for example, a mezzanine is foreseen in order to perform segment finding for the [MDT](#) Trigger Processor), and the [CTP](#) has specific interface needs. The approach of building common blocks was not adopted in the Phase-I system design, since the very tight latency constraints prompted the design of custom processing modules.

Prototypes already exist with a very similar technical complexity with respect to the Phase-II Level-0 components presented in this [TDR](#). The number of Multi-Gigabit Transceivers ([MGT](#)) present in Phase-I designs is comparable or superior to the Phase-II system needs, while the bandwidths of individual transceivers will double, up to  $\simeq 25$  Gb/s. The architecture is already well established in terms of interfaces and module interconnectivity, as well the broad use of large [FPGAs](#) as core processing elements. The [TDAQ](#) Level-0 system is not located in a high-dose rate radiation environment and therefore avoids the challenges of [ASIC](#) designs needed elsewhere in ATLAS.

Table 5.3: The size of the Phase-II hardware trigger system. The first column lists the subsystem and the second column lists the hardware component of that subsystem. Different components using the same common module type are indicated in the third column, and the fourth column indicates which modules are commissioned in Phase-I. The fifth, sixth, and seventh columns list the number of ATCA boards, the number of input Multi-Gigabit Transceiver (MGT) links per board, and the number of output MGT links per board, respectively. The LTI output links use PON technology while all other links are point-to-point high speed links.

Subsystem	Component	Module	Phase-I deliverable	Number of Boards	Input links per board	Output links per board
LOCalo	eFEX	-	Yes	24	144	48
	jFEX	-	Yes	6	240	48
	gFEX	-	Yes	1	312	108
	fFEX	-	-	2	240	48
LOMuon	NSW	-	-	16	148	28
	Endcap SL	SL	-	48	96	60
	Barrel SL	SL	-	32	60	60
	MDT	-	-	64	72	72
Global Trigger	MUX	GCM	-	23	156	108
	GEP	GCM	-	24	108	36
	CTP Interface	GCM	-	1	60	24
MUCTPI	-	-	Yes	2	208	65
CTP	CTPMI	-	-	1	-	-
	CTPIN	-	-	1-2	24	12
	CTPCORE	-	-	1	24	60
	LTI	-	-	36	1	8

The biggest challenge of the Level-0 system resides in the initial design stage in understanding of the resource needs for the implementation of the algorithms within a latency budget envelope.

In some cases, like in the LOMuon system, the muon selection algorithms are well established but the processing strategy will change. The current system is distributed between locations in the cavern and in USA15; however, the Phase-II system design dictates an ATCA-based solution that will be completely housed in USA15. The processing done in one  $\Delta\eta \times \Delta\phi = 1.0 \times 0.4$  current Level-1 Muon Barrel trigger sector is currently distributed across  $\simeq 100$  ASICs ( $\simeq 20 \text{ mm}^2$  area, 180 nm process) and 9 FPGAs (2002 technology). The additional resource needs, required by the new inner RPC layer, will increase this figure. A single FPGA will accommodate the processing equivalent of 150-200 of these ASICs. A preliminary study aimed to establish the processing power of the current biggest FPGAs, has evaluated that with a conservative 50% resource usage, the logic of about 50 ASICs will fit in the largest FPGA available at the time of writing this document, a Xilinx Ultrascale+ device. The challenge in this case is to gain a factor of two by optimising the existing algorithm and another factor of two based on the availability of a new generation of FPGAs.

## 5.2 The Level-0 Trigger System

Table 5.4: A summary of the required rack space in *USA15* for each system component. Maximum Possible Values (MPV) of the power needs are also shown, together with the totals for the Level-0 Trigger system.

Subsystem	Component	Number of Shelves	MPV Power [kW]
L0Calo	eFEX	2	16
	jFEX	1	6
	gFEX	1	3
	fFEX		
L0Muon	NSW	2	9
	Endcap Sector Logic (SL)	4	17
	Barrel SL	3	12
	MDT	6	29
Global Trigger	MUX, CTP Interface	2	13
	GEP	2	16
Central Trigger	CTP	1	3
	MUCTPI	1	3
	LTI	distributed	16
Level-0 Trigger System Total		25	143

In other cases, like in the [Global Trigger](#) system, the resources needed to implement offline-like algorithms is the main technical challenge, while the time-multiplexed architecture has been already proved viable by the CMS experiment.

### 5.2.8 Level-0 Trigger Latency

A bottom-up estimate of the full system latency is an important component in optimising the system architecture. The data coming from all detectors are processed in parallel by [L0Calo](#), [L0Muon](#)/[MUCTPI](#) and in the initial stage of [Global Trigger](#), executing all algorithms that process full-granularity calorimeter data. The [L0Muon](#) system design has the largest latency envelope, since it will perform precise tracking using [MDT](#) long-drift-time data, seeded by [RPC](#) and [TGC](#) muon candidates. A baseline implementation and an option are foreseen in the [MDT](#) trigger; to be most conservative the latency quoted is calculated for the option, since it has the longer latency. The estimated latency for the arrival of the muon information from the [MUCTPI](#) at the input of [Global Trigger](#) is about  $4.5 \mu\text{s}$ .

The latency estimated for the [Global Trigger](#) is given in three envelopes. The first two envelopes are related to the calorimeter and muon-only processing which takes the cumulative latencies respectively of  $\sim 4.4 \mu\text{s}$  and  $\sim 5.4 \mu\text{s}$ . The final step required before taking the final Level-0 decision is the [TIP](#) assembly and transmission to [CTP](#) via the [CTP Interface](#), which takes the cumulative latency to  $5.9 \mu\text{s}$ .

---

The overall **TDAQ** latency budget of  $10\ \mu\text{s}$  is defined at the output of the **FELIX** optical links. Considering the longest possible fibre length at the input of **FELIX**, the final **TDAQ** Level-0 latency estimation is  $6.9\ \mu\text{s}$ .

Table 5.5 shows a detailed breakdown of the internal processing, link serialisation, deserialisation and multiplexing steps, and signal propagation times along cables and optical fibres for each subsystem. The estimates quoted so far have been calculated as Current Best Estimates (CBE). Starting from Current Best Estimates (CBE) and estimating uncertainties, we introduce the definition of Maximum Possible Values (MPV). All unknown uncertainties contribute at the same level as the known uncertainties, so the MPV values are calculated by summing to CBE known uncertainties twice.

The CBE latency is  $6.9\ \mu\text{s}$ , which leads to an MPV latency of  $9.1\ \mu\text{s}$ , well within the specified  $10\ \mu\text{s}$  limit. Nevertheless every effort must be made to minimise the latency of each component of the system, since the final readout buffers latency will have an additional contribution given by the optical fibres connecting **FELIX** to the detector front-end electronics; this additional contribution is not part of the **TDAQ** envelope but needs to be considered in the design of the detector front-ends.

## 5.3 Data Acquisition System

With the increased Level-0 rate and the larger event-size, the data throughput in Run 4 is expected to increase and the **DAQ** capabilities need to be enhanced accordingly. Moreover, the upgraded **DAQ** needs to accomplish the requirements coming from the planned changes in the subdetectors, like the new **ITk** tracker, the modified calorimeter and the new muon-detector readouts, together with other potential new systems. The **DAQ** system includes Detector Readout subsystem, Dataflow subsystem, the network connection and the common online software framework. Figure 5.4 shows these main functional blocks as part of the Phase-II **DAQ** system.

### 5.3.1 Readout

The detector Readout subsystem receives event data from detector **FE** links and facilitates detector-specific processing, such as formatting and monitoring, before final transfer to the dataflow system. It also relays **TTC** signals and Control and Configuration information to on-detector electronics and relays **DCS** information between the on-detector electronics and the **DCS**. It is comprised of the **FELIX** and the **Data Handler**.

### 5.3 Data Acquisition System

Table 5.5: *Phase-II Level-0 Trigger Latency. Detailed estimations and cumulative values (in  $\mu\text{s}$ ) are shown for the best estimated values (BEV) and for maximum estimated values (MEV). TP stands for Trigger Processor.*

Subsystem	Item	CBE	Uncertainty	$\int$ CBE	$\int$ MPV
L0Calo	LAr signals at FEX inputs	1.100	0.15		
	Tile signals at FEX inputs	1.425	0.14		
	FEX Processing	0.500	0.08		
	FEX to to MUX (5 BC + 10m)	0.175	0.03	2.1	2.6
L0Muon	TGC signal at Endcap SL inputs	0.888	0.03		
	RPC signals at Barrel SL inputs	1.110	0.03		
	Barrel SL final pre-processing	0.390	0.05		
	Barrel SL to MDT TP (5 BC + 10m)	0.175	0.05	1.7	1.9
	NSW signals at SL inputs	1.425	0.14		
	Tile signals at SL inputs	1.425	0.14		
	Endcap SL final pre-processing	0.075	0.05		
	Endcap SL to MDT TP (5 BC + 10m)	0.175	0.03	1.7	2.1
	Last MDT info at USA15	2.358	0.10		
	MDT Trigger (FPGA-based)	1.277	0.10		
	MDT Trigger (AM-based)	1.578	0.10		
	MDT TP to SL (5BC + 10m)	0.175	0.03	3.8	4.3
	SL final processing	0.025	0.03		
	SL to MUCTPI (7BC+10m)	0.175	0.03	4.0	4.6
MUCTPI	MUCTPI processing	0.250	0.10		
	MUCTPI to MUX (5BC + 15m)	0.200	0.03	4.5	5.3
Global Trigger Muon	Muon MUX processing (serdes+fibres)	0.650	0.30		
	Global Trigger pipelined processing	0.100	0.05		
	Global Trigger Muon Triggers	0.200	0.03	5.4	7.0
Global Trigger Calo	LAr signals at MUX inputs	1.100	0.15		
	Tile signals at MUX inputs	1.425	0.14		
	Calo MUX	1.400	0.05		
	Global Trigger pipelined processing	0.100	0.05		
	Global Trigger non-jet triggers	0.200	0.03	3.1	3.6
	Jet processing (from Global Trigger pipelined)	1.250	0.50		
	Global Trigger non-muon triggers	0.200	0.03	4.4	6.1
	Global Trigger Global				
Central	TIP assembly	0.050	0.03		
	Serial Transmission	0.250	0.03		
	Global to CTP (5BC + 5m)	0.150	0.03	5.9	7.6
Central	CTP Processing	0.275	0.10		
	Link to LTI (5BC + 30 m)	0.275	0.03		
	LTI Internal processing	0.100	0.03		
	Link to FELIX (5BC + 10m)	0.175	0.05	6.7	8.8
FELIX	FELIX Processing	0.200	0.05		
TDAQ	From Collision to FELIX out			6.9	9.1

**FELIX** FELIX is the component of the Readout subsystem that will implement the interfaces to the detector-specific electronics via custom, point-to-point serial-links, e.g. the Versatile link<sup>3</sup>, and acts as an interface to the Data Handler and DCS via a commodity

<sup>3</sup> Versatile Link Plus for HL-LHC: <https://espace.cern.ch/project-Versatile-Link-Plus/SitePages/Home.aspx>

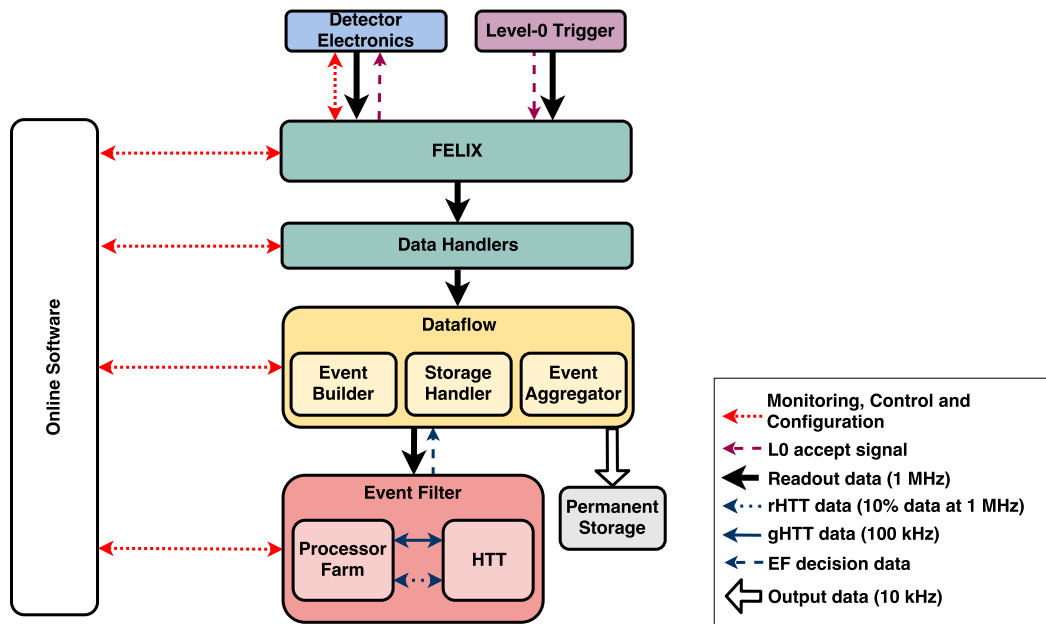


Figure 5.4: The main functional blocks of the *TDAQ* architecture for Phase-II, with focus on the *DAQ* aspects. Shown in green are the Readout components and yellow the Dataflow components, with interfaces to other systems indicated by white for Software dependence, *DCS* System in blue, Event Filter in Red and Permanent Storage in gray.

multi-gigabit network. *TTC* information will be received from a *PON* and relayed to the detector *FE* electronics as well as to the *Data Handler*.

The *lpGBT* protocol is being explored by subdetector systems and *L0* trigger system for the bi-directional point-to-point serial links connecting the *FELIX* and their electronics. These systems will either directly incorporate a *GBTX* chip on the electronics components or implement *lpGBT* protocol with *FPGAs*. For the downlinks (links to the detector electronics), most systems plan to use *lpGBT* to send *TTC* information, collect Busy information, and to perform configuration and control. For the uplinks (links from the detector electronics) some systems like *ITk* foresee to have different protocols to transmit detector data and *DCS* information. To accommodate the downlink transmission *FELIX* must also receive packets from the commodity multi-gigabit network and route them to the serial links. This places a requirement on *FELIX* reliability and uptime beyond that of data taking operations. *FELIX* therefore functions as a router between custom serial links and commodity multi-gigabit networks. It is largely detector-agnostic and encapsulates common functionality, but minimal detector-specific functionality may be required in *FELIX* to decode or process the received data beyond what is required to determine its destination.

*FELIX* is being developed and will be deployed in the Phase-I upgrades of the *NSW*, *LAr* calorimeter trigger electronics and the *L1Calo* and *L1Muon*. These upgrades only require a subset of the functionality described above and the design and prototyping for Phase-I is



described in the [FELIX](#) Phase-I preliminary design review [5.3]. For Phase-II the design of [FELIX](#) will be expanded to address the Phase-II specific requirements, namely: new protocol needed; configuration and control of calibration procedures; support of a [PON](#) for reception of [TTC](#) information and Busy propagation.

Similar to the Phase-I development, [FELIX](#) will be implemented with PC server holding custom [FPGA](#) I/O cards, which provide [lpGBT](#) and other needed protocols as well as the [PON](#) interface, the [ONU](#). It is expected that each I/O card will be able to support up to forty-eight bi-directional links and two such mezzanines will be hosted in a commodity-computing server. It is required that the total average input data rate does not exceed 70% of the total input bandwidth availability of the input links of the [FELIX](#) card. Based on this model of implementation the size of the Phase-II [FELIX](#) system is summarised in Table 5.6. In prototyping for Phase-I the custom card is implemented as a third generation [PCIe](#) card. The Phase-II prototype will build on the Phase-I prototyping and evaluate other emerging technologies, such as the fourth generation [PCIe](#).

**Data Handler** The Data Handler receives data from [FELIX](#) via a commodity multi-gigabit network. It allows for detector-specific processing, e.g. formatting and/or monitoring, of the data prior to storing them in the Dataflow). To meet the requirements of detector-specific trigger-aware monitoring, automated recovery from error conditions during data taking and fragment book-keeping based on [LOA](#), the Data Handler will also receive the Level-0 trigger information.

The Data Handler infrastructure is expected to operate on commodity [PCs](#) and the system size is shown in Table 11.1. The size of the Data Handler system is estimated based on the Phase-I system composition of the [FELIX](#) and swROD (1:1), plus additional processing power foreseen for Phase-II detector-specific data processing. Any detector-specific data processing still needed as well as all aspects of configuration, control, and monitoring will be implemented by the customisation of the back-end software services, the raw data processing that is implemented by firmware in the current [RODs](#) will be implemented as a customisation of the detector Data Handler.

Since [FELIX](#), the Data Handling infrastructure and the [EF](#) farm will all be connected by networks, a change in the current event building paradigm will be possible. In the simplest scenario, [FELIX](#) could use a round-robin policy to concentrate data from different Level-0 events into different Data Handling elements. This would provide a first building step in which fragments are composed into larger 'macro' fragments. At the second step, downstream processing would simply join macro fragments into a full event and collect data from a subset of the Data Handling elements for each event. Such a scheme would split the event builder load onto two different systems. However, this may require changes to the detector-specific data-sanity algorithms since not all triggers will be seen by a given Data Handling server. Furthermore, implementing two-step event building may require the synchronisation of the event building steps across the different Readout paths.



Table 5.6: Summary of Phase-II Detector Readout Link and Bandwidth Requirements. Downlink refers to data travelling toward the front-end electronics, and uplink to data travelling from the front-end toward the rest of the DAQ system. Detectors with existing FELIX installations from Phase-I will be updated with new hardware as required.

Detector	Number of FELIX boards	Number of Links	Bandwidth (Gb/s)
ITk Pixel downlink	224	1285	2.5
ITk Pixel uplink		10596	5
ITk Strips downlink	76	1552	2.5
ITk Strips uplink		1824	10
LAr LASP downlink	36	100	2.5
LAr LASP uplink		770	10
LAr LDPB downlink	8	31	2.5
LAr LDPB uplink		155	10
L0Calo downlink	8	16	2.5
L0Calo uplink		120	10
NSW downlink	96	864	5
NSW uplink		1440	5
NSW Trigger Processor	4	64	5 or 10
Global Trigger downlinks	4	48	2.5
Global Trigger uplinks		96	10
Tile	8	160	10
TGC	8	192	10
MDT	64	1536	10
RPC	4	64	10
CTP	1	12	10
MUCTPI	1	4	10
LUCID uplink	1	24	5
Zero Degree Calorimeters (ZDC) uplink	1	24	5
AFP uplink	1	12	5

5.3.2 Dataflow

The Dataflow subsystem buffers, transports, aggregates and compresses event data. It is responsible for the transport of data from the output of the Detector Readout to CERN permanent storage. The main functional elements of the Dataflow system are the Event Builder, the Storage Handler, and the Event Aggregator. The Event Builder is the logical interface to the Readout System. It builds event records and manages the storage volume of the Storage Handler system. The Storage Handler is a high throughput large volume storage system which buffers event data before and during processing by the Event Filter. For events accepted by the Event Filter, the Event Aggregator collects, formats and transfers the output to CERN permanent storage.

The Dataflow subsystem receives events from the Readout subsystem’s Data Handlers at 1 MHz and buffers them in the Storage Handler at a total writing (i.e. input) throughput of 5.2 TB/s. The Storage Handler provides the Event Filter access to Region of Interest information at 1 MHz and full event access at 400 kHz for a total reading (i.e. output) throughput of about 2.6 TB/s. The Event Aggregator receives the full event record at the output rate of 10 kHz for a total throughput of 60 GB/s out to CERN permanent storage. Figure 5.5 shows a high-level description of the communication between the Dataflow components.

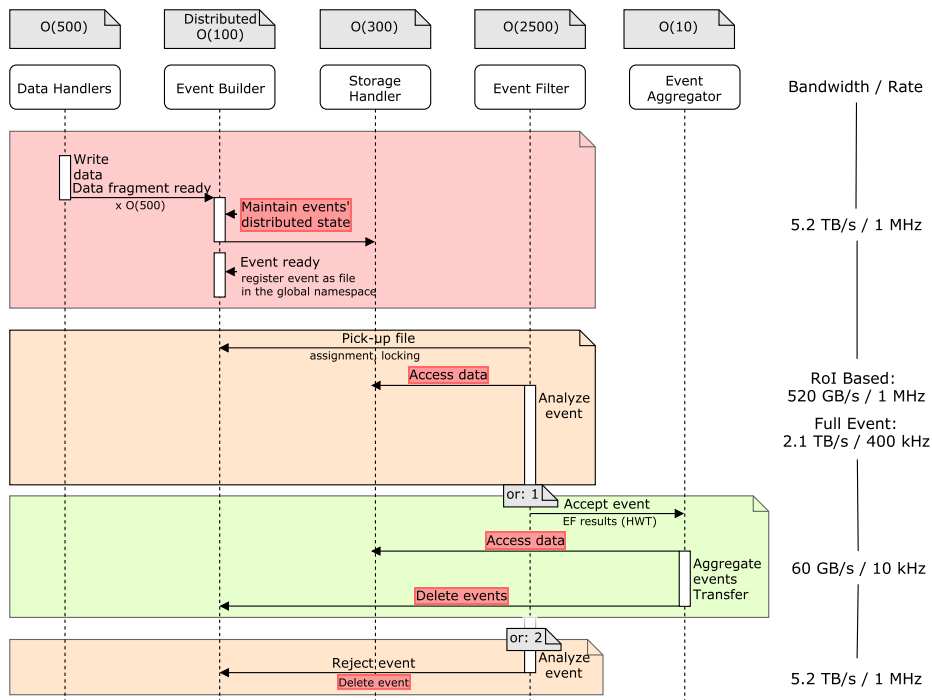


Figure 5.5: Logical communications between different components of the Dataflow system.

The Dataflow subsystem will be implemented with commodity servers and storage units. Compared to Run 1 & Run 2, the plan is to deploy centralised storage buffers large enough

to decouple real-time data taking and [EF](#) processing. Additionally, the storage buffer will provide a new abstraction layer between the Dataflow and the [EF](#): location information of events or data files on a permanent storage system will be exchanged, offloading the data movements to the storage layer.

**Event Builder** The role of the Event Builder, as in Run 1 and Run 2, is to physically or logically join the fragments or macro-fragments of an event in a single place. The baseline scheme for Phase-II is to use Regions of Interest ([RoI](#)) and Level-0 trigger information to reduce the event rate at Level-0 (1 MHz) to 400 kHz. The events passing this first rejection step are then passed to the Event Filter for further analysis, making use of all data available for each event. Depending on the Storage Handler implementation and characteristics, event building could be physical or logical. In the former case, the event data are indeed gathered in a single storage unit. This can be achieved via a specific computing element, fetching the complete event data before storage, or by a collective effort of multiple elements interacting via a specialised concurrent data format and storage technology. Alternatively, in case of logical event building, the event data are not contiguously stored together. Instead only the metadata information about the fragment locations are collected.

The Event Builder is the logical interface of the Dataflow subsystem to the Readout subsystem and the [EF](#) system. The Event Builder will be implemented as a software interface to the Storage Handler. The Data Handlers will write data to the logical storage volume via this interface. Across the Event Builder system events must be tracked at the L0-accept rate of 1 MHz. In the case of an operational issue causing the Storage Handler volume to fill up the Event Builder interface shall throttle Data Handler traffic to reduce the rate of data being written, thus asserting back pressure to the Readout subsystem.

**Storage Handler** The Storage Handler buffers data received from the Readout subsystem to decouple the Readout and Event Filter. The large storage volume needed to achieve the required 7.8 TB/s throughput allows for increased event processing latency. The Readout subsystem accesses the Storage Handler through the Event Builder software layer. The Storage Handler presents the event data to the Event Filter and the Event Aggregator in the form of files within a single name-space implemented with a [Distributed FileSystem \(DFS\)](#).

The storage space can be provided within a performance-tiered cluster using a variety of different media types with different throughput and capacities including: hard-disk drives (HDD), solid-state drives (SSD), non-volatile memory express devices (NVMe) and random-access memory (RAM). While the use of different media types allows the system to meet the high-throughput requirements, the assumption is that to meet the throughput requirement SSD will be the dominant technology for the long-time buffer space during the [EF](#) processing. The storage system provides 2.6 TB/s of read access to the Event Filter.

Therefore the long-time buffer storage needs to provide total 7.8 TB/s of reading and writing throughput. The system is sized according to this assumption. Over the past ten years HDD capacity grew by a factor of 10 and the trend is expected to continue through to 2025. The same improvement in capacity can be reasonably expected for SSD technologies, the highest capacity SSD available is  $\sim 2$  TB currently (thus 20 TB in ten years). Today's SSD technologies provide 2 GB/s of aggregated throughput and this throughput scales with capacity up to the extent that the connectivity can support. With PCIe Gen4 connected SSDs, 5 GB/s can be assumed as an achievable baseline for future SSD throughput. So the throughput needed for the long-time buffer space can be achieved with  $\sim 1800$  SSDs. Such a system would provide up to 36 PB of total storage capacity and represents over an hour of event buffering before and during Event Filter processing.

The Storage Handler system comprises its own storage units and devices by default but may also take opportunistic advantage of other systems' storage media to aggregate storage space where possible and pertinent.

**Event Aggregator** The Event Aggregator integrates a physical storage system on top of which software functionality is implemented. It will take advantage of the Storage Handler and only implement the functionality needed to format and expose the data stored there to the offline infrastructure. It receives events accepted by the Event Filter, performs compression (if needed) and then prepares output files to be transferred to permanent storage. Finally the Event Aggregator will be responsible for any necessary communication with the Tier-0 processing centre if anything beyond data transfer is required.

To meet the requirements of CERN-IT the Event Aggregator provides a buffer area capable of storing up to 48 hours of accepted event data. This decouples online operation of the data acquisition system from offline activities and enables the system to cope with disruptions or malfunctions of the physical connection to the permanent storage outside of the ATLAS experimental area.

Although [Hard Disk Drives \(HDDs\)](#) could provide the stable storage volume needed [HDDs](#) throughput is not expected to increase considerably in the next decade and throughput would be a major design constraint. In order to buffer 48 hours of data at 10 kHz with an event size of 6 MB, it will require the Event Aggregator to have a storage volume of 10 PB. This capacity would require 500 [Solid State Disk Drive \(SSD\)](#)s which are naturally included as part of the Storage System needed for the Storage Handler. The required output throughput of 60 GB/s is negligible compared to the total 7.8 TB/s of the Storage Handler. A common storage system for both the input and output buffer more efficiently use resources and provides greater system flexibility.

### 5.3.3 Network

The amount of control traffic during data taking is not expected to grow significantly in the Phase-II and a dedicated network is not planned for control and monitoring traffic. Therefore the network upgrade for Phase-II is mainly needed for the Readout network and the Dataflow network.

**Architecture** The Readout network need high throughput to connect the [FELIX](#) and the Data Handlers with additional requirements for [DCS](#), control and monitoring infrastructure interconnection. It will organise the system into slices and connect all the nodes in a given slice to the same high-throughput switch. Each slice will contain a certain number of [FELIX](#) and a set of Data Handlers they need to communicate with. All slices will be connected together with a pair of routers on a higher hierarchical level allowing all-to-all communication for [DCS](#), control and monitoring purposes. The Dataflow network provides connectivity between all the Dataflow components and the [EF](#) farm. It will aggregate Data Handlers with pizza-box switches and connect the Storage Handler units to the same network switch. A set of Data Handlers will then always write to the same set of Storage Handler units, maximising the throughput and reducing the latency. The switches will have high-throughput uplinks to the network core routers to allow all-to-all communication between the Storage Handler and the rest of the system. For connectivity the [EF](#), [EF](#) servers are stored in racks and connected to top-of-rack switches. Every top-of-rack switch will be connected to the core network with enough uplink capacity to ensure the required throughput. For the Event Aggregator to send accepted events to the permanent storage, high-throughput long-range links will be needed between the core routers at Point 1 and the CERN data centre.

**Implementation and Technology** A possible network implementation is illustrated in [Fig. 5.6](#). The router cluster can provide a basic level of network redundancy as needed simply by connecting pairs of uplinks to different devices. Its implementation could have different solutions, a pair of chassis routers or a distributed leaf-spine switch topology. Failures of Data Handlers can be dealt with transparently by re-assigning data to neighbouring handlers.

Such an implementation is suitable for Ethernet technology. There are also several other network technologies with promising future performance projections in terms of bandwidth and latency that can be considered for the Phase-II networks, such as InfiniBand and OmniPath. R&D work will continue to identify the best suitable technology with technology tracking, technology evaluation and network slice construction.

## 5.4 Event Filter System

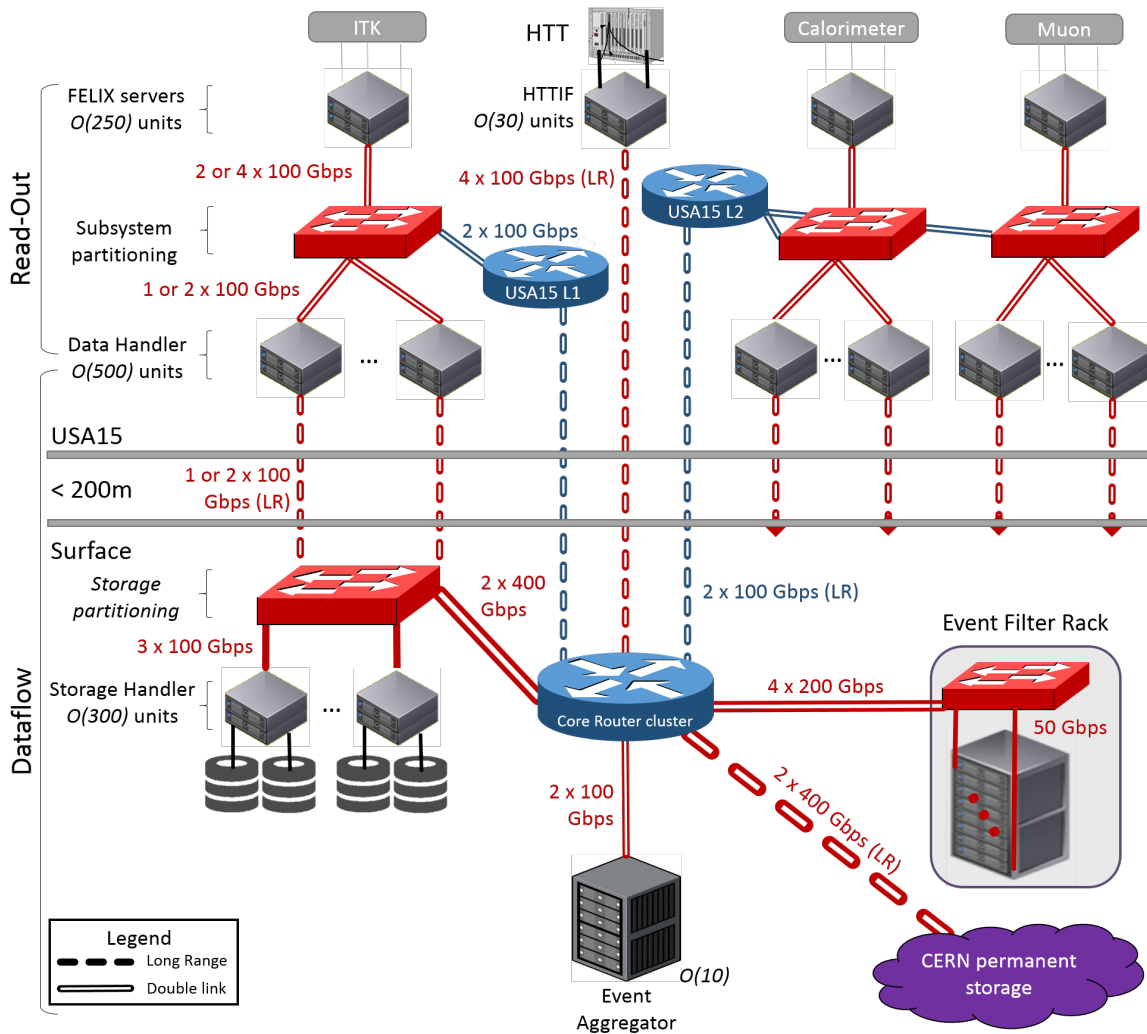


Figure 5.6: Possible network implementation using Ethernet technology.

## 5.4 Event Filter System

The Event Filter (EF) system takes as input the detector data from events accepted by the preceding hardware trigger at 1 MHz (see 6.11). It must select the most useful events according to the trigger menu at a rate of 10 kHz (see 6.11) and reject the rest. Fast rejection is important to achieve high throughput and minimise Storage Handler space requirements. An initial selection based on combining information from the Level-0 hardware trigger with the regional hardware tracking (rHTT) is used to reduce the event rate down to around 400 kHz, as shown in Table 6.4. This is followed by reconstruction of events using the EF processors to run algorithms that are similar to or in common with the offline reconstruction, and use of global hardware tracking (gHTT) as also shown in Table 6.4.

### 5.4.1 Use of Tracking in the Event Filter

High luminosity, and consequently high pile-up, conditions give rise to specific challenges for object and event reconstruction algorithms: higher occupancy of tracking detectors makes pattern recognition slower and more susceptible to fakes, and energy resolution in calorimeters is reduced. This negatively affects, for example:

- the distinction between reconstructed particles from a hard scatter and the many superimposed soft events,
- the separation of electrons from background jets,
- calculation of global event quantities like  $E_T^{\text{miss}}$ ,
- jet energy resolution,
- and separation of  $b$ -jets from light jets.

This in turn leads to higher trigger rates and/or worse efficiency. To achieve the goal of maintaining a menu with thresholds similar to Run 1, the EF selections need to be robust against pile-up so that rates scale approximately linearly with luminosity. Use of tracking to identify a primary vertex and associate reconstructed objects with it is a well-established approach to address this problem.

Efficient charged particle reconstruction in HL-LHC conditions relies on the tracking detector upgrade (ITk) [5.4][5.5] and software improvements. Figure 5.7 shows the CPU usage per event in units of HEP-SPEC06<sup>4</sup> (HS06) times seconds for the inclined ITk layout. Even with the preliminary tuning of the ITk reconstruction software, the total CPU usage spent on tracking at  $\langle \mu \rangle = 200$  is less than five times that spent on tracking in the Run 2 system at  $\langle \mu \rangle = 20$ . Considering the worse-than-linear expectation for track reconstruction time using Run 2 software described in Section 3.4.2, a substantial improvement in precision track reconstruction is expected.

Therefore, the use of information from the Inner Tracker ITk as early as possible in the trigger selection is considered a key ingredient in the ATLAS Trigger strategy for HL-LHC. The proposed use of tracking for specific trigger signatures is shown in Table 6.4.

Several technologies may be considered for reconstructing tracks in the EF:

- i a Hardware-based Tracking system for the Trigger (HTT) based on FPGAs and custom-designed Associative Memory (AM) ASICs,
- ii commodity CPU-based servers,
- iii systems based on accelerators (e.g. GPGPUs),
- iv future architectures based on devices integrating machine learning capabilities.

Some of these options have been studied extensively in the past few years and some are detailed in Part II of this document.

<sup>4</sup> <http://w3.hepik.org/benchmarking.html>



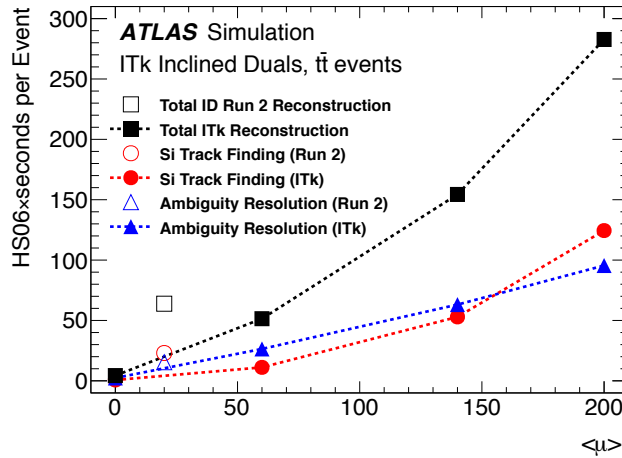


Figure 5.7: The CPU usage required to reconstruct a  $t\bar{t}$  event in the ITk as a function of the average pileup, for the inclined ITk layout [5.4]. The total CPU usage required to reconstruct the charged particles in the ITk (filled squares) is compared to the corresponding CPU requirements for reconstructing the current Run 2 detector (open square) at an average pileup of 20 events per bunch crossing. The CPU usage required for the track-finding (circles) and for the ambiguity resolution (triangles) separately are also shown.

The baseline design is an HTT system based on AM ASICs for pattern recognition and FPGAs for track reconstruction and fitting to meet the high trigger rate and throughput requirements at the HL-LHC, as described in Section 5.4.5.

This decision is motivated by the following: extensive experience with the technology within ATLAS, its short latency, its lower power budget and less demanding space requirements, its cost effectiveness and the independence of its cost from the commodity computing market, and the capability to evolve the HTT system for use in the hardware-based Level-1 trigger should ATLAS need to change to a dual L0/Level-1 Trigger (L1) (see Chapter 14). The evolution of the above technologies will be followed, so that the baseline choice may be reconsidered in case of a major technological breakthrough.

Software tracking will still be done in the EF when the ultimate precision is needed, otherwise tracks will be requested from the HTT and simply unpacked and used in EF selections. More advanced software reconstruction techniques for calorimeter and muon detectors will also be needed to achieve performance robust against pile-up and fast enough to run on-line. In general these will be drawn from or developed with the offline reconstruction, as a high correlation in EF and offline reconstruction efficiencies is desirable.



## 5.4.2 Event Filter System Overview

The EF system will have interfaces with the HTT and Dataflow systems, principally the Storage Handler and Event Builder. Figure 5.8 shows the interactions among these systems. The EF communicates with the HTT via the HTT Interface (HTTIF) which is not shown here but is described in Section 5.4.5.

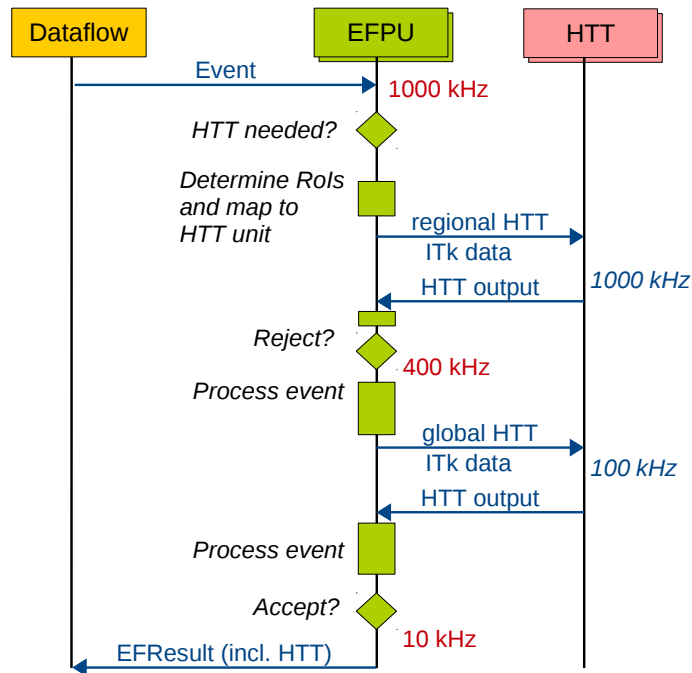


Figure 5.8: Diagram showing the interaction between the Event Filter processing unit (Event Filter Processing Unit (EFPU)), Dataflow system and the hardware tracking (HTT). After receiving the event data from the Dataflow system, the EFPU decides, based on the Level-0 decision, if regional hardware tracking is needed for this event. If so, requests are sent to the relevant HTT units along with the necessary ITk data. After receiving the HTT result the EFPU continues with the event processing including a possible second request for global hardware tracking based on the Level-0 decision or additional EF-based selections.

The EF overall receives events at 1 MHz, with each event assigned to a particular EF processing unit (EFPU). The full event is not read by the EFPU process at this rate. First information only from the L0 CTP and Global Trigger is used to determine the RoIs related to the L0 triggers which passed. The sequences of reconstruction algorithms to run and the choice of whether or not to use the HTT are driven by this information. If regional tracking is required, the EF reads the ITk module data corresponding to the relevant RoIs and sends this to the rHTT. The system is designed for up to 10% of the ITk data to be requested and used in this way, as explained in Section 6.12. If the rHTT was used, the EF then takes a trigger decision based on the regional tracks and Global Trigger objects such as topological calorimeter clusters, to reject some events and reduce the rate to 400 kHz. The dataflow sys-

tem can take advantage of the above partial event data access pattern and transfer the full event data only for accepted events. If global tracking is required, which will be determined by the trigger menu, around 10% of events, then the full **ITk** data is read and sent to the **gHTT**. During **HTT** requests, **EF** waits the short time (compared to **EF** event processing) for the tracks to be returned. Since the **EF** processing is multi-threaded, other events are processed during this time so there are no idle **CPU** cycles. Further software reconstruction and selection steps will take place until finally at a rate of 10 kHz, events are accepted.

The **EF** farm and software applications will be controlled by the Online Software system. The software will also provide physics and operational monitoring services. The **EF** processing is expected to occur promptly, i.e. with a minimum delay after the Level-0 accept decision. The large buffer in the Storage Handler will allow processing timeout settings to be less strict than are needed in the Run 2 and Run 3 system. It is in principle possible for the Storage Handler to buffer the data for longer, to allow processing to continue beyond the end of a data-taking run, thus utilising the inter-fill periods as well. This delayed processing introduces considerable operational complexity that outweighs the benefits unless there is a strong use case. The option of an **EF** calibration loop (similar to the prompt calibration loop deployed at Tier-0) that would derive certain calibrations before the **EF** processing starts has been studied. While there is no clear use-case for calorimeter or muons, a full **ITk** alignment could be derived within one hour of data-taking. However, assuming the new **ITk** detector is at least as stable as our current detector it is not clear that this would result in significant improvements for the **EF** selection.

### 5.4.3 Event Filter Farm Hardware

The most cost-effective computing platform for the **EF** software, the commodity **PC** server, shows a trend towards systems hosting multi-core **CPUs** with an increasing core-count and heterogeneous hardware architectures, incorporating General Purpose Graphical Processors (**GPGPUs**) or **FPGAs**. Therefore, the upgraded **EF** software should allow for both parallel algorithm execution as well as exploitation of internal parallelism for those which are the most costly in **CPU** terms. The optimal degree of parallelism (multiple events, intra-event i.e. concurrent Region of Interest (**RoI**) processing, or intra-algorithm) will be found by balancing the potential benefits in throughput and the effort needed to modify and maintain code.

Current estimates of the evolution of event processing times indicate the need for  $4.5^{+2.7}_{-0.7}$  million HEP-SPEC06<sup>5</sup> (MHS06) to handle a Level-0 rate of 1 MHz, of which an initial reduction to 400 kHz will be achieved entirely by using information from Level-0 and **rHTT** as shown in Table 6.4. The **CPU** estimate is based on an extrapolation of current (Run 2) **CPU** usage, taking into account scaling with pile-up, **CPU** time reduction due to the use of hardware tracking, and other software improvements. Figure 5.9 shows the expected **CPU**

---

<sup>5</sup> <http://w3.hepik.org/benchmarking.html>

Table 5.7: Summary of Event Filter farm size estimates, based on projections of compute capacity requirements and compute power of servers from current data, as described in the text.

Compute capacity required	4.5 MHS06
Equivalent dual-socket servers	3000
Racks	38

requirements versus pile-up assuming most of the **ITk** tracking is offloaded to **HTT** as described below in Section 5.4.4. More details on the model used for this estimation are given in Section 12.4. The largest uncertainties in this estimate are due to the expected number of RoIs, the possible improvements in the reconstruction software and the reduction in **CPU** requirements due to the use of hardware tracks where possible.

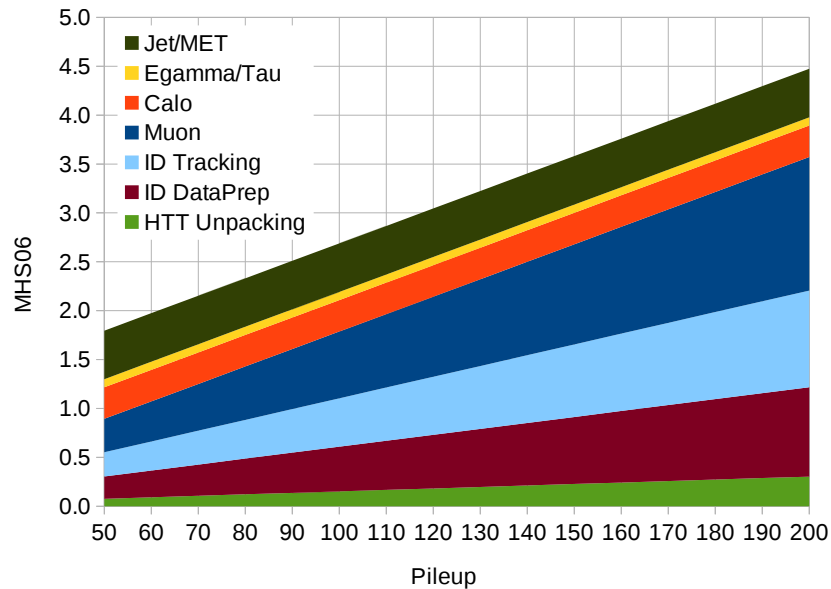


Figure 5.9: *EF CPU* extrapolation versus pile-up in million HEP-SPEC06 for the different components of the *EF* reconstruction software.

All the commodity compute power must be accommodated within a fixed rack-space located in the surface computing infrastructure. This rack-space is shared with other components of the data-acquisition system, such as storage and networking. An extrapolation based on the evolution of compute-power in the ATLAS **TDAQ** farm over the past ten years results in an estimated compute capacity of 1.5 kHS06 per dual-socket server or approximately 3000 motherboards on the time-scale of Phase-II. The scale of the system is summarised in Table 5.7.

**GPGPUs** are a potential commodity hardware accelerator to which suitable compute-intensive processing tasks can be offloaded from the main CPU. This has been studied for Phase-I: the findings are summarised in Section 12.3 where further references are provided. In that study it was estimated that using contemporary hardware it would cost approximately

the same, and have similar heating, cooling and space requirements, to increase the farm throughput by adding either [GPGPU](#) or CPU. The relative cost-effectiveness of these technologies for the Phase-II [EF](#) depends on the relative evolution of [CPU](#) and [GPGPU](#) in terms of price, performance and packaging.

The decision on the commodity compute hardware for the [EF](#) farm will be taken nearer the time of purchase following an evaluation of hardware accelerators. The above work on [GPGPUs](#) should be taken as an indication that the software can be successfully adapted to other architectures for a full cost/benefit evaluation.

It is envisaged that the Event Filter farm could be a rather heterogeneous system, possibly containing different classes of hardware such as [GPGPU](#) and commodity servers. Due to the rolling replacement strategy, it will host different hardware families within the same class. Since the technology decisions shaping the Event Filter infrastructure will be taken as late as possible, it is important to establish interfaces that will allow for operation in all of the above scenarios.

### 5.4.4 Event Filter Software

The selection software upgrades fall into two parts: further evolution of the framework and development of the selection algorithms to meet Phase-II requirements.

The present software framework is undergoing a major upgrade to provide the new functionality needed for the start of Run 3. The new framework, AthenaMT [5.6][5.7], is being implemented as a common trigger and offline computing framework. Key features are built-in multi-threading support, provision for seamless integration of offline algorithms and infrastructure to support external accelerators. The AthenaMT framework will be able to run existing [EF](#) and offline algorithms with limited changes, but additional changes will be needed to fully exploit the potential of the new framework. Most of the changes in the selection software are expected to be done in time for Run 3. Further changes for Run 4 will be needed to adapt the system to the evolved DAQ architecture, i.e. with regards to the Storage Handler and Hardware Tracking interfaces. In addition, new trigger hardware (e.g. [Global Trigger](#)) will have to be included into the trigger configuration database. In case [GPGPU](#) accelerators will be used the necessary framework services to offload the compute load to these devices will have to be developed.

Significant effort will be required to upgrade the selection software to provide the required rejection in the [EF](#) within [CPU](#) resource constraints. It is expected that an initial reduction from the Level-0 rate of 1 MHz to 400 kHz can be achieved by the use of hardware tracks and information from [Global Trigger](#). After this initial 'fast' rejection, the remaining required [EF](#) rejection power will be achieved by importing techniques currently used offline to provide selections that are robust against the effects of pile-up, and where necessary developing new techniques. Experience from Run 1 and Run 2 has shown that the rates of certain triggers, especially multi-jet triggers and triggers based on  $E_T^{\text{miss}}$ , rise rapidly with

increasing pile-up. Information on the average number of interactions per bunch-crossing is currently used to reduce the effect of pile-up. However, the most effective methods rely on event-by-event primary vertex reconstruction that provides the number and position of the primary vertices in the event. This allows, for example, the precise calculation of the contribution of pile-up vertices to the energy of a calorimeter cluster.

The biggest change is expected to the inner detector tracking software, as it will have to be adapted to the new **ITk** detector and make full use of the hardware tracks provided by **rHTT** and **gHTT**. Since the trigger tracking software is largely based on the offline tracking software, the additional trigger-specific changes are expected to be of limited scope. Significant experience in the inclusion of hardware tracks is expected to be gained once the **FTK** system is fully operational in Run 2. The current muon trigger reconstruction [5.8] is composed of a fast muon and inner detector track reconstruction, followed by a combination step and final precision reconstruction in both the inner detector and muon system. The reconstruction software will have to be adapted to include and leverage the information provided by the new **MDT** Level-0 trigger as well as hardware tracks, both of which should result in a significant reduction in processing time. It is assumed that at least a factor two improvement in the muon reconstruction time can be achieved for Run 4 and this improvement is included in the estimated compute requirements.

The calorimeter software will develop with the aim to further harmonise the trigger and offline reconstruction. With the fast rejection techniques used in the first steps of the **HLT** moving to the Level-0 trigger, there is both an opportunity to have these without any **EF** processing, and a consequence that more precision reconstruction will be needed to achieve the necessary rejection. Offline algorithms are the natural option but significant work will be required to speed them up for use at high rates early in the **EF**. It may be that new algorithms and techniques have to be developed, both for **EF** and offline reconstruction. The use in the **EF** of topoclusters produced by the **Global Trigger** will be studied. In principle they could be used with better object-finding algorithms to do a fast initial rejection in the **EF**.

### 5.4.5 Hardware-based Tracking for the Trigger Subsystem

The Hardware Track Trigger subsystem (**HTT**) aims to provide tracks for the **EF** quickly, hence reducing substantially the processing requirements at the **EF**. The **HTT** receives requests to find tracks, along with the related **ITk** data from the **EF**. It combines hits from the eight outermost **ITk** layers (including at least one Pixel layer) to quickly find track candidates. A second stage of processing uses information from all the remaining **ITk** layers to fit and refine the tracks can be performed when greater precision is required.

Two complementary **HTT** functionalities are envisaged in the baseline architecture, largely operating on different signatures and with different performance targets. The first is a regional **HTT** (**rHTT**) that finds tracks with  $p_T > 2$  GeV in limited regions around Level-0

trigger objects that can profit from track information (single high- $p_T$  leptons and multi-object triggers). The **rHTT** will operate at the full **L0A** rate of 1 MHz and process on average 10% of the **ITk** data in these events. The second function, called the global **HTT** (**gHTT**), performs full-scan event-level tracking at a nominal rate of 100 kHz. The **gHTT** finds tracks with  $p_T > 1$  GeV and, in contrast to **rHTT**, it includes the second stage processing mentioned above, hence providing high-quality tracks with better purity, for **EF** algorithms that rely more critically on this information (primary/secondary vertices,  $b$ -tagging, missing energy).

The input data to the **HTT** pattern recognition for the full event are **ITk** module data fragments from a subset of layers, about 2 MB per event. The second stage uses an additional 1.5 MB. For each track candidate, **HTT** provides the hits associated to it, together with five track parameters ( $p_T$ ,  $\eta$ ,  $\phi$ ,  $d_0$ , and  $z_0$ ) and the corresponding  $\chi^2$  value to assess the quality of the track fit. More sophisticated software algorithms to refit tracks can be used later in the **EF**, but the **CPU**-intensive pattern recognition does not have to be repeated.

The first-stage processing of the **HTT** finds track candidates in two steps: first it finds groups of **ITk** hits that match precomputed patterns stored in the Associative Memory (**AM**) **ASIC**, then processes the matched hit combinations with linearised track-fitting algorithms in **FPGAs**, to extract the tracking parameters. The capacity of the future **Associative Memory 09 ASIC (AM09)** chip is expected to be three times larger than the one used in the current **FTK** project ( $3 \times 128k$  patterns), with a processing speed of 250 MHz words per input bus. The event processing time is mainly driven by the **ITk** data transmission through the **AM** chips and the number of fits to be executed.

The second-stage processing takes as inputs the tracks found in the first stage and the clusters from **ITk** layers not used in the first stage. It then extrapolates each track to these additional **ITk** layers, finds clusters that may belong to the track and performs a full track fit that improves the track parameter resolutions significantly. The second-stage processing is implemented in **FPGAs**.

Being an **EF** co-processor, the **HTT** communicates with the **EF** through a dedicated interface, **HTTIF**. The **HTTIF** receives the tracking requests (regional in the case of **rHTT**) from the **EF**, together with the **ITk** data, propagates them to the appropriate **Tracking Processor main board in the HTT (TP)** boards, receives back the tracking results and returns them to the **EF** processor that formed the original request. The **HTTIF** will be implemented with **PC** servers holding **PCIe** cards, which communicate with the **EF** farm via commodity network, and with **HTT** units via optical links with custom protocol to be defined.

The **HTT** is organised as an array of independent tracking units called **HTT** units, as shown in Fig. 5.10. The main hardware building block of the **HTT** are **ATCA** boards called Tracking Processors (**TPs**). There are two types of **TP** boards: Associative Memory Tracking Processor, **AM Trigger Processors (AMTPs)**, and Second Stage Tracking Processor, **Second-Stage Tracking Processors (SSTPs)**, that will be implemented on identical **ATCA** boards, with different firmware. Each tracking unit comprises a set of six **AMTPs** and one **SSTP**

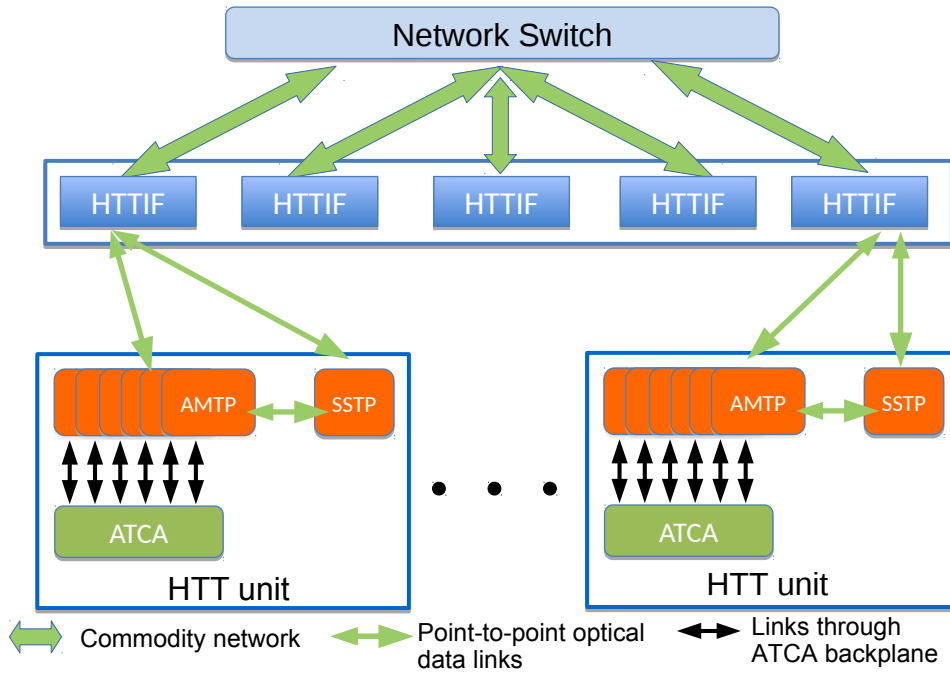


Figure 5.10: Overview diagram of the *HTT* system showing interconnections within *HTT* units and with the *HTTIF*.

Table 5.8: Summary of characteristics and size of the *HTT* system.

	<b>HTT</b>
rHTT minimum track $p_T$	2 GeV
rHTT Input rate	1 MHz @ 10%
gHTT minimum track $p_T$	1 GeV
gHTT Input rate	100 kHz
Number of HTTIF	48 (to be revisited)
Number of ATCA shelves for AMTPs	48
Total number of AMTPs	576
Total number of AM chips	18432
Number of ATCA shelves for SSTPs	8
Total number of SSTPs	96
Power estimate per TP	300 W

that will perform tracking in a specific  $\eta - \phi$  region of the track parameter phase space. Table 5.8 gives an overview of the main characteristics and size of the *HTT* hardware.

R&D of the Phase-II *AM ASIC* based on 28 nm technology is already underway. A first prototype called *Associative Memory 07 prototype ASIC (AM07)* has been recently re-



ceived and is under test. The next prototype called [Associative Memory 08 prototype ASIC \(AM08\)](#) will be a small area device including all functions of the final [AM09](#) with a smaller number of patterns (16k patterns). The logic and full-custom designed elements of [AM08](#) will be used as the base for the design of [AM09](#) that will be a large area [ASIC](#) to be produced with full-mask submission.

There is confidence that the [AM](#)-based baseline design will work. While this is design developed towards the Preliminary Design Review, other technologies such as commodity hardware accelerators are being kept under evaluation in case a major breakthrough occurs which would make it worth changing the design.

## References

- [5.1] ATLAS Collaboration, *Technical Design Report for the Phase-I Upgrade of the ATLAS TDAQ System*, Tech. Rep. CERN-LHCC-2013-018. ATLAS-TDR-023, CERN, Geneva, Sep, 2013. <https://cds.cern.ch/record/1602235>.
- [5.2] F. Lanni et al., *Interfaces with Detector Front-End Systems. Requirement Document for HL-LHC*, Tech. Rep. ATL-DA-ES-0051 v.1, EDMS Id 1563801 v.1, CERN, Geneva, March, 2016. <https://edms.cern.ch/document/1563801>.
- [5.3] P. Farthouat et al., *Report of the preliminary design review of the FELIX for Phase 1*, Tech. Rep. ATU-TC-ER-0048 v.1, EDMS Id 1518157 v.1, CERN, Geneva, Aug, 2015. <https://edms.cern.ch/document/1518157>.
- [5.4] ATLAS Collaboration, *Technical Design Report for the ATLAS Inner Tracker Pixel Detector*, Tech. Rep. CERN-LHCC-2017-021, CERN, Geneva, (*in preparation*).
- [5.5] ATLAS Collaboration, *Technical Design Report for the ATLAS Inner Tracker Strip Detector*, Tech. Rep. CERN-LHCC-2017-005, CERN, Geneva, Apr, 2017. <https://cds.cern.ch/record/2257755>.
- [5.6] J. Baines, T. Bold, P. Calafiura, S. Kama, C. Leggett, D. Malon, G. Stewart, and B. Wynne, *ATLAS Future Framework Requirements Group Report*, Tech. Rep. ATL-COM-SOFT-2014-048, CERN, Geneva, Dec, 2014. <https://cds.cern.ch/record/1974156>.
- [5.7] J. Baines, T. Bold, D. Oliveira Damazio, S. Martin-Haugh, J. Masik, T. Martin, F. Pastore, H. J. Stelzer, W. Panduro Vazquez, R. White, W. Wiedenmann, F. Winklmeier, B. Wynne, C. Bernius, M. Hodgkinson, C. Leggett, E. Moyse, I. Shapoval, S. Snyder, and G. Stewart, *Design of the High Level Trigger in AthenaMT*, Tech. Rep. ATL-COM-DAQ-2017-047, CERN, Geneva, Jun, 2017. <https://cds.cern.ch/record/2268736>.



- [5.8] ATLAS Collaboration, *Performance of the ATLAS Trigger System in 2015*, [Eur. Phys. J. C 77 \(2017\) 317](#), [arXiv:1611.09661 \[hep-ex\]](#).



## 6 Expected Performance

The identification of trigger-level objects (such as electrons, muons, and jets) is defined relative to each object's offline definition. The Phase-II [TDAQ](#) upgrade will enable the implementation of object identification algorithms that are much closer to the offline definitions than the Run 3 system. The [Global Trigger](#) processors deliver trigger objects based on full-granularity calorimeter information at Level-0. The [HTT](#) provides tracks to the Event Filter at the Level-0 rate (1 MHz) in regions of interest for tracks above 2 GeV and tracks above 1 GeV at a reduced rate (100 kHz) for the full-detector. In both cases, track reconstruction extends to the full [ITk](#) coverage of  $|\eta| < 4.0$ . This section describes the expected Phase-II trigger object performance relative to the current (Run 2) understanding of the offline definitions; it is expected that these definitions will continue to evolve beyond this document. In all cases, the performance is evaluated for a luminosity of  $\mathcal{L} = 7.5 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ .

### 6.1 Performance Estimation Procedures

The Level-0 and early rejection regional tracking in the Event Filter ([rHTT](#)) results presented in this section are based on full-simulation studies of the ATLAS detector [6.1] using Geant4 [6.2]. The planned algorithms to be implemented in hardware or software are then applied to the simulation to determine the performance. The trigger efficiencies for the various physics objects are evaluated using samples of signal events (e.g.,  $Z \rightarrow ee$  or  $Z \rightarrow \tau\tau$ ) overlaid with an average of  $\langle\mu\rangle \simeq 200$  minimum-bias events generated with PYTHIA 8 [6.3]. To reproduce the pile-up conditions expected in Phase-II, background rejections and trigger rates are obtained from a sample of overlapping minimum-bias events with  $\langle\mu\rangle \simeq 200$ . The metric by which the performance of a given trigger object is evaluated consists of the expected rate on minimum-bias events versus the corresponding effective offline thresholds in signal events. The effective offline threshold is the value of the relevant quantity (such as  $p_T$ ), above which the trigger reaches  $\sim 90 - 95\%$  of its plateau efficiency; the exact efficiency goal depends on the relevant object.

The Event Filter algorithms that are used in data-taking will follow an extensive programme of development of offline reconstruction from now until data-taking and beyond. These developments will increase the sophistication of the algorithms designed to manage the higher level of pile-up and exploit features of the upgraded detector. It is therefore not possible to accurately estimate Event Filter rates with the actual algorithms that will be used. Instead, the Event Filter rates are evaluated using selections based on the current

HLT algorithms applied to Run 2 data linearly extrapolated in luminosity to  $\mathcal{L} = 7.5 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ . For the majority of triggers, the Event Filter rates are approximately linear with luminosity. The exception is the missing transverse energy trigger, for which a dedicated analysis based on the Geant simulation described above is used. The ongoing development of offline reconstruction means the Event Filter must be able to follow those developments to avoid broad turn-on curves and object selection inefficiencies.

## 6.2 Topological Clusters

The proposed [Global Trigger](#) will allow the reconstruction of offline-inspired noise-suppressed “topological clusters”, or topoclusters [6.4]. Topoclusters are based on the full-granularity Liquid Argon and Tile calorimeter information for cells with  $|E_T| > 2\sigma$ , where  $E_T$  is the reconstructed cell transverse energy and  $\sigma$  is the sum in quadrature of the calorimeter electronics noise and the pile-up noise. In the “422” scheme, each topocluster is seeded from a cell with  $|E_T| > 4\sigma$ , and neighbouring cells (in three dimensions) with  $|E_T| > 2\sigma$  are added to complete the cluster. The full offline topological clustering algorithm [6.4], the “420” scheme, has an additional boundary layer of cells with  $E_T > 0$  adjacent to the cluster. In particular, the transverse energy of the reconstructed clusters is not significantly affected by the omission of the  $E_T > 0$  boundary layer, although the number of cells in each cluster is substantially smaller. The nominal plan for the upgrade is the implementation of the “422” scheme in the [Global Trigger](#) with minor modifications including the use of  $E_T$  provided by the calorimeters. The details of the implementation are described in Section 9.3.3. Comparisons of the number of cells per cluster, the number of clusters, and the transverse energy ( $E_T$ ) of the clusters are shown in Fig. 6.1 for  $ZH \rightarrow \nu b \bar{b}$  events. There are on average fewer cells per cluster in the “422” scheme (without the boundary layer). Furthermore, the splitting and merging algorithms are implemented independently, which results in the number of reconstructed clusters being slightly different. The effect of these differences on jets is described in Section 6.6. In addition, calorimeter-based topocluster moments may be calculated and calorimeter-based calibrations may be applied to improve jet and tau performance.

Once formed, the topoclusters will be used as the basis for jet and tau reconstruction. The trigger topocluster objects can also be used to calculate isolation-type variables for electrons, photons, muons, and tau leptons as well as to aid tau identification.

## 6.3 Electrons and Photons

**Level-0** Electron and photon object identification in Phase-II begins in the legacy Phase-I system (using the [eFEX](#)). The [Global Trigger](#) then refines the identification algorithm by taking advantage of the transmission of fine-granularity cells with  $|E_T| > 2\sigma$  from the [LAr](#)

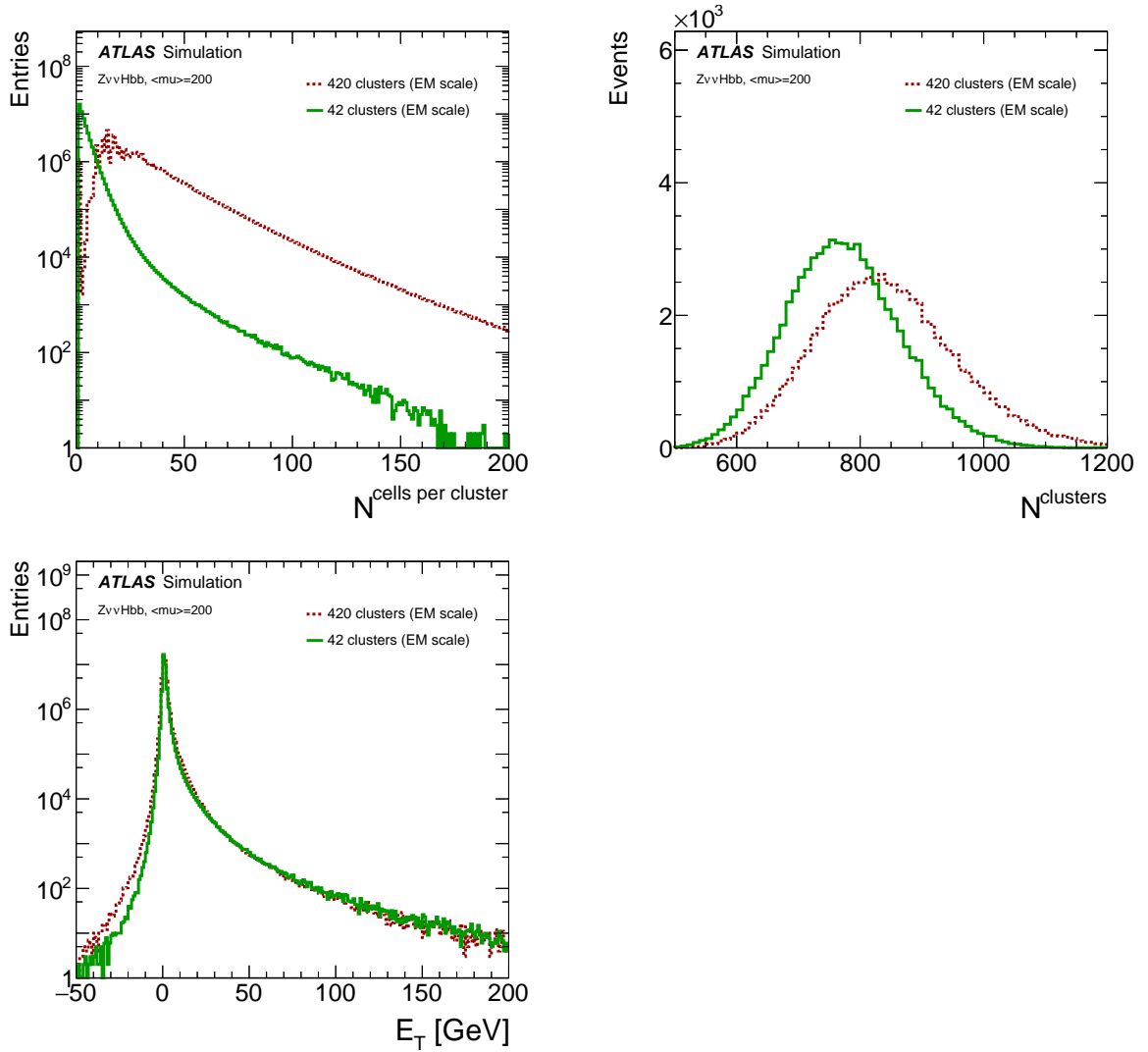


Figure 6.1: Performance of the proposed “422” (green, labeled “42” in the legend) topocluster reconstruction scheme compared to the offline “420” scheme (red) in  $ZH \rightarrow vbb$  events. The top plots are the number of cells in the clusters and the number of clusters found, and the bottom is the transverse energy of the clusters. The topocluster  $E_T$  is calculated with the individual cells calibrated to the electromagnetic scale and no further corrections (“EM Scale”).

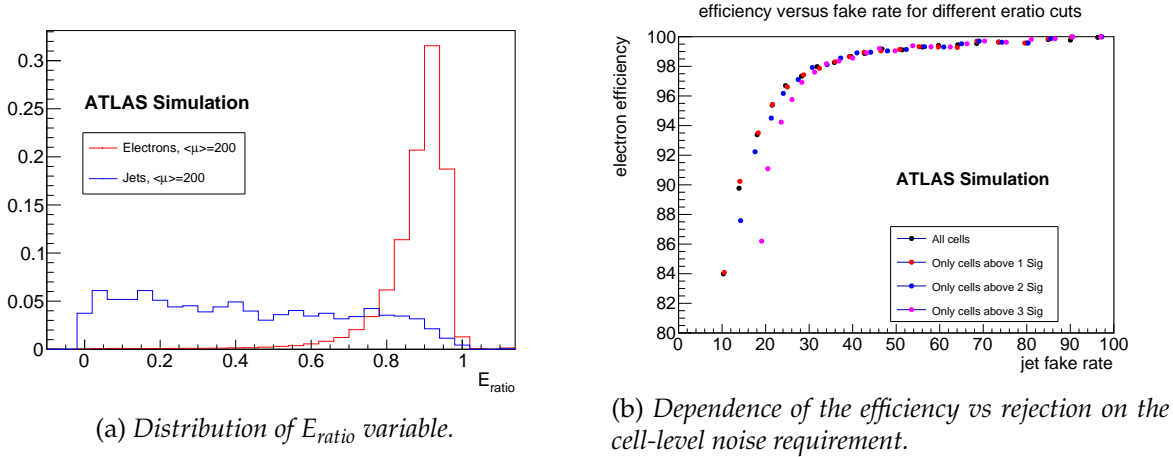


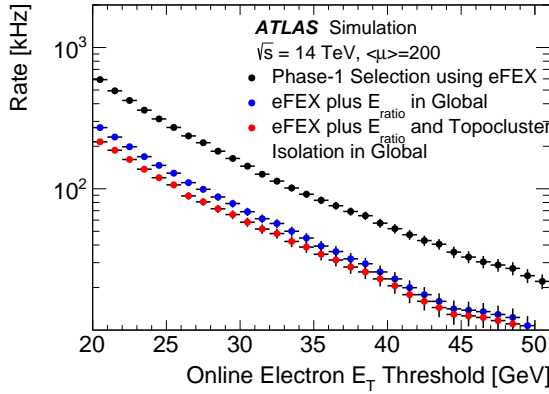
Figure 6.2: Performance of the  $E_{\text{ratio}}$  variable for electron identification assessed on  $Z \rightarrow ee$  signal and minbias samples. The electron candidate has a selection applied that is similar to the one that will be applied by the [eFEX](#) before the candidates reach the [Global Trigger](#) with a  $p_{\text{T}}$  threshold of 20 GeV.

calorimeter. One promising shower shape variable referred to as  $E_{\text{ratio}}$  (currently employed in the Run 2 [HLT](#) and offline electron and photon identification algorithms [6.5][6.6]) makes use of the first layer of the [LAr](#) calorimeter which is not available in the Phase-I [eFEX](#).  $E_{\text{ratio}}$  is defined as

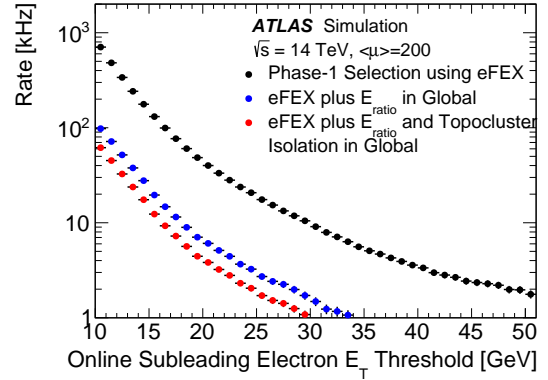
$$E_{\text{ratio}} = \frac{E_{\text{highest energy cell}} - E_{2\text{nd local maximum energy cell}}}{E_{\text{highest energy cell}} + E_{2\text{nd local maximum energy cell}}} \quad (6.1)$$

where  $E_{\text{highest energy cell}}$  is the energy in the cell with the largest energy deposit associated with the electron and  $E_{2\text{nd local maximum energy cell}}$  is the energy in the next largest local maximum (in a window of typically 0.0625 in  $\eta$ ). This discriminates between dominant background from  $\pi^0 \rightarrow \gamma\gamma$  showers (whose maximum and second-maximum energy depositions would be balanced) and isolated electrons or photons (whose  $E_{\text{ratio}}$  is close to 1). Figure 6.2 shows the distribution of  $E_{\text{ratio}}$  with a  $2\sigma$  energy threshold and the efficiency versus rejection when  $E_{\text{ratio}}$  is only computed for cells whose  $|E_{\text{T}}|$  is above the given noise threshold. The latter plot demonstrates that the  $|E_{\text{T}}| > 2\sigma$  threshold does not have a significant detrimental impact on the performance of this variable; a factor of five additional rejection can be achieved for  $\sim 99\%$  efficiency.

In addition, the availability of topoclusters in the [Global Trigger](#) allows for an implementation of the topocluster-based isolation currently used in the offline and [HLT](#) reconstruction. The resulting electron trigger rates at  $\mathcal{L} = 7.5 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$  are shown in Figure 6.3a for the [eFEX](#) selection only, then plus an  $E_{\text{ratio}}$  requirement, and finally with an additional topocluster-based isolation requirement. The rate for a 20 GeV threshold with the [eFEX](#) selection alone is 600 kHz. The  $E_{\text{ratio}}$  cut reduces the rate by a factor of  $\approx 2\times$  and then isolation requirement reduces the rate an additional 20 – 25%, bringing the total rate low



(a) Single-electron trigger rates as a function of leading lepton  $p_T$



(b) Di-electron trigger rates as a function of sub-leading lepton  $p_T$

Figure 6.3: The Level-0 rate as a function of leading  $p_T$  for the single-electron/photon triggers and sub-leading  $p_T$  for dielectron/diphoton triggers. Three levels of selection are shown solely from the *eFEX*, *eFEX* plus a requirement on  $E_{ratio}$ , and after a topocluster-based isolation requirement.

enough such that a 20 GeV threshold takes an acceptable fraction of the total readout bandwidth. The diphoton/dielectron rates are shown in Figure 6.3b.

### Regional Tracking Rejection

In order to support the high Level-0 rate, early rejection in the Event Filter using regional tracking (*rHTT*) is required. The rejection as a function of efficiency when requiring a track with a minimum  $p_T$  is shown in Fig. 6.4. The rejection is approximately a factor of 5 for 95-97% efficiency for an online electron threshold of 18 GeV and a factor of 3 for 95-97% efficiency for an online threshold of 10 GeV.

For dielectron triggers it is also possible to require the candidates be consistent with originating from the same vertex in the  $z$ . Study of the rejection expected from that requirement is underway.

### Event Filter Output

Finally in the Event Filter, selections similar to electrons and photons in the offline system are implemented including Gaussian sum filter for the electron tracks to account for radiation along their paths. The rates are summarised in Table 6.1. The single-electron trigger presented in the table has a track-based isolation requirement using tracks above 1 GeV  $p_T$  as is used in the current trigger selection.

## 6.4 Muons

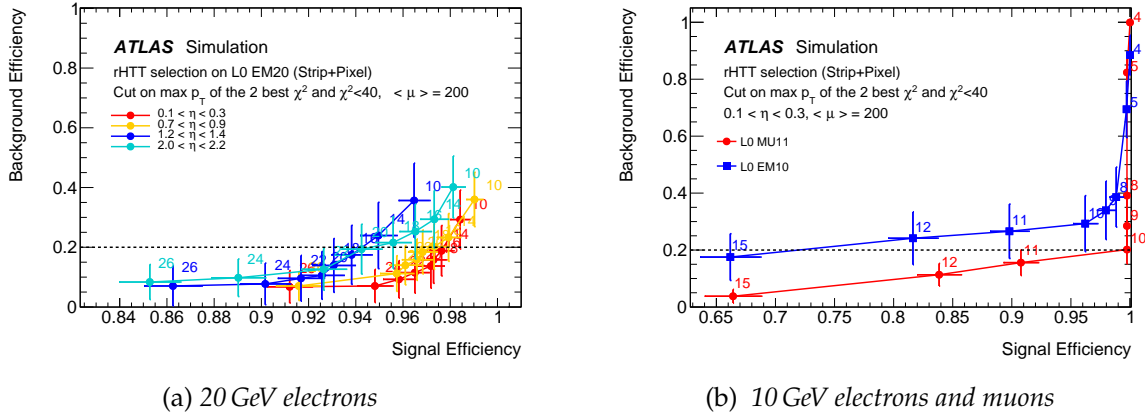


Figure 6.4: Rejection as a function of efficiency when requiring a track for (a) 20 GeV electrons and (b) 10 GeV electrons and muons. The track  $p_T$  requirement is indicated as numbers next to the lines. The electron and muon samples are single particle samples with  $p_T$  distributions that are flat in  $1/p_T$  from 4-400 GeV. The backgrounds are based on minimum bias simulation.

Table 6.1: Projected Event Filter rates for representative electron and photon triggers.

Trigger	Offline Threshold [GeV]	Rate
Level-0 Single Electron	22	200 kHz
After regional tracking cuts	22	40 kHz
Event Filter Isolated Single Electron	22	1.5 kHz
Level-0 Single Photon	120	5 kHz
No regional tracking cuts	-	-
Event Filter Single Photon	120	0.3 kHz
Level-0 Dielectron	10,10	60 kHz
After regional tracking cuts	10,10	10 kHz
Event Filter Dielectron	10,10	0.2 kHz
Level-0 Diphoton	25,25	20 kHz
No regional tracking cuts	-	-
Event Filter Diphoton	25,25	0.2 kHz

## 6.4 Muons

**Level-0** The Level-0 trigger for muons, described in Section 8, uses the Phase-I muon trigger [RPC](#), [TGC](#), and [NSW](#) detectors with new on-detector and trigger electronics to seed tracking in the high-precision [MDT](#) chambers. In addition, new [RPC](#) chambers will be added in the barrel region to increase the efficiency which would otherwise be significantly degraded from the Phase-I performance due to a lowering of voltage on the [RPC](#) chambers



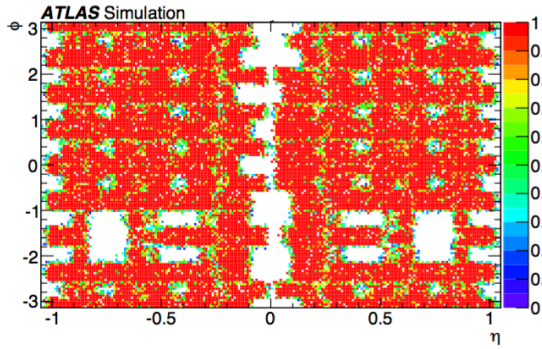
---

to avoid aging. Figure 6.5a shows the  $\eta - \phi$  coverage for the Phase-I system, which requires hits in all three RPC layers in order to manage the muon trigger rate. With the new RPCs this can be relaxed to requiring a coincidence in 3 out of 4 RPC layers. This significantly increases  $\eta - \phi$  coverage as seen in Fig. 6.5b. Finally, with the additional rate rejection from the MDT system, the RPC requirement can be further relaxed to a two-layer coincidence. The resulting improvement in  $\eta - \phi$  coverage is illustrated in Fig. 6.5c.

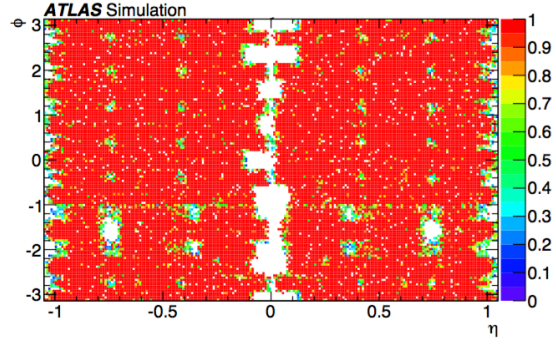
The corresponding impact on the maximum efficiency of the 20 GeV muon trigger is shown in Fig. 6.6a. The line for “3/3 chambers” corresponds to the Phase-I system. The line “3/4 chambers” corresponds to the new RPC configuration (the additional “3/4 chambers with BO hits” adds a requirement that the outer barrel layer be hit to remove the acceptance below 10 GeV). The new RPCs chambers increase the efficiency from  $\sim 80\%$  to  $\sim 90\%$  on the plateau. The 2 out of 4 requirement on the RPC layers is restricted to the barrel inner and barrel outer layers. This is shown in Fig. 6.6a as the “3/4 chambers + BIBO” line, which achieves an efficiency of  $\sim 96\%$ .

Also mentioned in Fig. 6.6a is that the RPC trigger rates increase from  $\sim 20$  kHz for the low-efficiency Phase-I system to  $\sim 85$  kHz for the high-efficiency “3/4 chambers + BIBO” scheme. The rate is dominated by the trigger acceptance below threshold, i.e. the trigger accepts a substantial number of muons below the nominal 20 GeV threshold. To reduce the rate due to those muons, the high-precision MDT chambers are used in the muon trigger to significantly sharpen the turn-on curve. The corresponding trigger efficiency turn-on curves for a 20 GeV threshold in the region  $|\eta| < 2.4$  before and after the inclusion of the MDT chamber information are shown in Fig. 6.6b. The MDT selection reduces the “2 out of 4” RPC rate from the  $\sim 85$  kHz to  $\sim 30$  kHz. It also reduces the end-cap TGC+NSW trigger rate for 20 GeV muons from  $\sim 30$  kHz to  $\sim 15$  kHz.

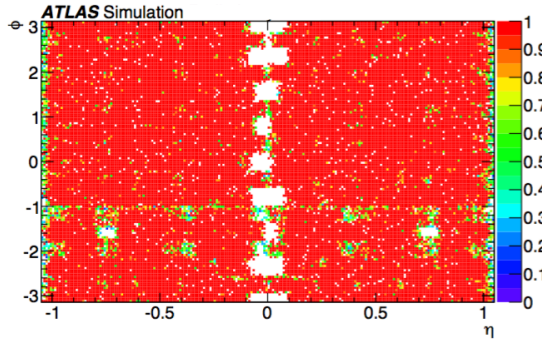
The dimuons rate estimates are still at an early stage. Each rate has three contributions: two muon candidates from the same collision, two muon candidates from two different collisions, and two muon candidates due to double counting the same muon. We expect to be able to remove the vast majority of the double counting with the muon trigger processor systems. For 10 GeV muons, the different collision contribution is estimated to be a small effect ( $\approx 1$  kHz based on binomial statistics for two independent collisions), and the total rate is estimated based on an overlay procedure to be less than 10 kHz. For 5 GeV muons, the different collision contribution is substantial (approximately 34 kHz) and the total rate will be in the range of 50-150 kHz depending on the details of the topological selection and overlap removal, which are still being optimised. For the 5 GeV dimuon rate and for even lower thresholds topological selections can be used to further reduce the rate for triggers targeting  $b$ -physics. In Run 2, such topological triggers give rates a factor of  $\sim 10$  lower. To support  $b$ -physics and exotic scenarios within the scope of the proposed hardware, algorithms will be optimised to increase nearby muon acceptance and keep the dimuon thresholds low.



(a)  $\eta - \phi$  acceptance for the Phase-I system in Phase-II conditions



(b)  $\eta - \phi$  acceptance for new RPCs operated with a three-station coincidence

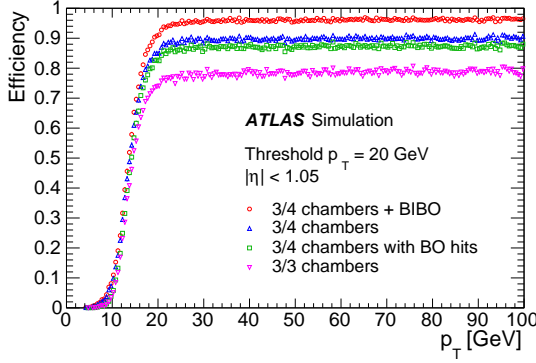


(c)  $\eta - \phi$  acceptance for new RPCs operated with a two-station coincidence (any 3 stations or "BIBO")

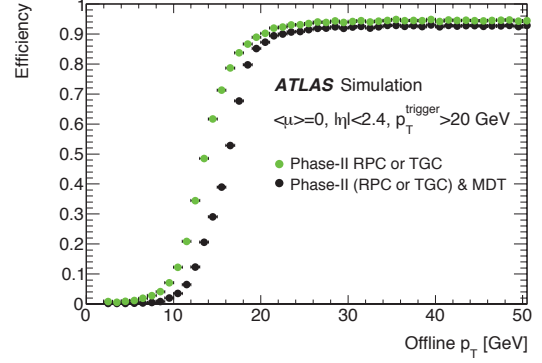
Figure 6.5:  $\eta - \phi$  acceptance for 20 GeV  $p_T$  trigger for (a) the Phase-I system in Phase-II conditions, (b) for new RPCs operated with a three-station coincidence, (c) for new RPCs operated with a two or three-station coincidence.

**Regional Tracking Rejection** Using `rHTT`, the muon trigger can add tracking requirements. For a 20 GeV trigger threshold, this is not necessary as the rate from the Level-0 is low enough to not require addition selections before the full Event Filter selection. For low- $p_T$  muons used in particular in the dimuon topology, a track requirement can reduce the rate with little loss of efficiency. Figure 6.4b shows that a factor of two rejection can be achieved with 99% efficiency. For a dimuon trigger, the rejection would be squared while the efficiency would drop to 98%. In addition, dimuons will be required to be consistent with a common vertex to reduce the impact of pile-up where the two muons are from different collisions.

**Event Filter Output** At the Event Filter, the muons are reconstructed with a method similar to the offline algorithm. Table 6.2 shows a representative set of rates. The single-muon trigger presented has a track-based isolation using a 1 GeV track  $p_T$  threshold.



(a) *RPC triggers for the Barrel Region ( $|\eta| < 1.05$ )*



(b) *Muon trigger ( $|\eta| < 2.4$ ) with or without *MDT*.*

Figure 6.6: Efficiency for a 20 GeV  $p_T$  muon trigger versus offline muon  $p_T$  for (a) the Phase-I system in Phase-II conditions, for new *RPCs* operated with a three-station coincidence, for new *RPCs* operated with a two or three-station coincidence (see text for details of the configurations), and (b) for *RPCs* operated with a two or three-station coincidence and a *MDT*-based requirement.

Table 6.2: Event Filter rates for representative triggers

Trigger	Offline Threshold [GeV]	Rate
Level-0 Single Muon	20	45 kHz
No regional tracking cuts	-	-
Event Filter Isolated Single Muon	20	1.5 kHz
Level-0 Dimuon	10,10	<10 kHz
After regional tracking cuts	10,10	<5 kHz
Event Filter Dimuon	10,10	0.2 kHz
Level-0 Low- $p_T$ Dimuon	5,5	50-150 kHz
no topological cuts		
Level-0 Low- $p_T$ Dimuon	5,5	< 15 kHz
with topological cuts		
Low- $p_T$ after regional tracking cuts	5,5	up to 10 kHz
Event Filter Low- $p_T$ Dimuon	5,5	up to 0.5 kHz

## 6.5 Tau Leptons

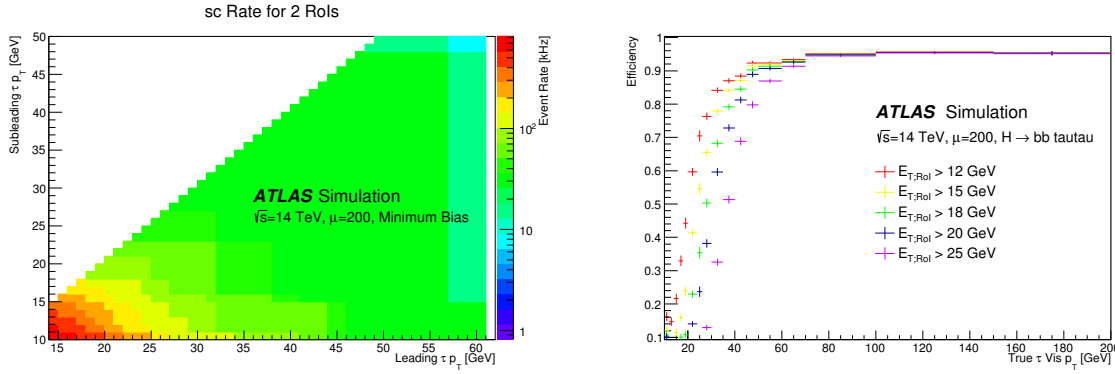


Figure 6.7: Rate as a function of leading and sub-leading online (uncalibrated)  $\tau p_T$  for a  $di\text{-}\tau$  trigger implemented in the *eFEX* and the corresponding efficiency per tau candidate (referred to as an “RoI” in the plot) with respect to offline for a set of online  $\tau p_T$  thresholds as a function of the offline  $\tau p_T$ .

## 6.5 Tau Leptons

**Level-0** The Level-0 tau lepton trigger can use the *eFEX* to define the initial candidate region and then makes use of the fine granularity calorimeter data reconstructed with the topocluster capability of the *Global Trigger*. The offline tau selection uses tracking and calorimeter information combined using a *Boosted Decision Tree (BDT)*. Two preliminary strategies are under development for tau identification. The cut-based trigger uses the medium granularity super-cell variables that are available in the *eFEX*. The second uses a recurrent neural network to select tau candidates in the *Global Trigger*. At the current early level of optimisation, both give comparable results, but it is expected that with further development a combined system can give better performance with moderate resource usage in the *Global Trigger*. Both algorithms approximately support a physics goal of reaching  $\approx 80 - 90\%$  of the efficiency plateau for offline  $p_T$  for leading  $\tau p_T > 40$  GeV and a sub-leading  $\tau p_T > 30$  GeV, which is motivated by the key  $H \rightarrow \tau\tau$  and  $HH \rightarrow bb\tau\tau$  analyses. This goal is not yet met by a small factor by either algorithm, but they are expected with further development.

The cut-based results use the energy in a narrow core, information on the depth profile of the energy deposition, and isolation requirements. The resulting rates and efficiencies are shown in Fig. 6.7. A 200 kHz rate corresponds to online  $\tau$  thresholds of  $p_T > 20$  GeV and  $p_T > 15$  GeV for the leading and sub-leading  $\tau$  candidates, respectively. These thresholds correspond to  $\approx 85\%$  efficiency for offline  $\tau$  candidates with  $p_T$  thresholds of 45 GeV and 35 GeV for the leading and sub-leading  $\tau$  candidates, respectively.

The recurrent neural network combines the information from up to six topoclusters. For each cluster, the input variables are cluster energy, the energy in the four layers of the electromagnetic calorimeter, the  $\Delta R$  distance from the seed jet axis, three moments of the

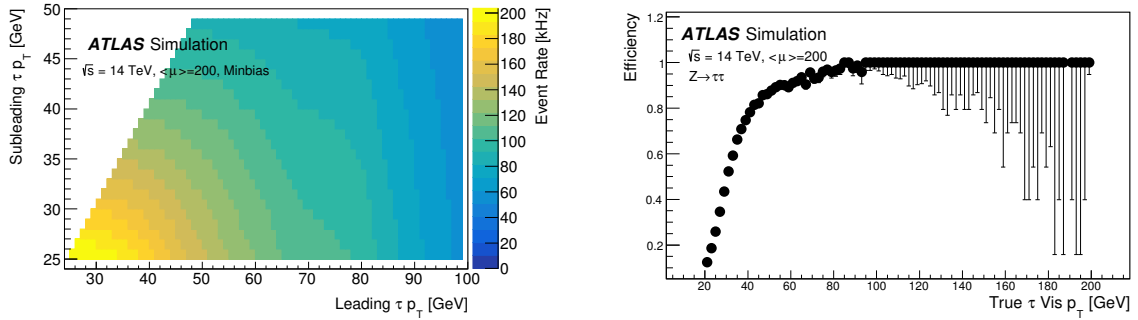


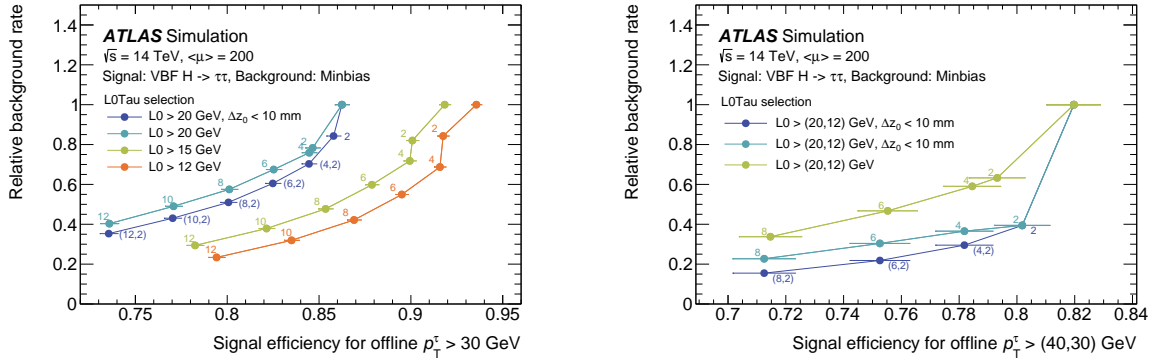
Figure 6.8: Rate as a function of leading and sub-leading online (uncalibrated)  $\tau p_T$  for a di- $\tau$  trigger implemented in the [Global Trigger](#) using an recursive neural-network (left) and the corresponding efficiency per tau candidate with respect to offline for a set of online  $\tau p_T$  thresholds as a function of the offline  $\tau p_T$  (right).

cluster shape, the depth of the cluster in the calorimeter, the fraction of energy contained in the highest energy cell, and the probability that the cluster is an electromagnetic shower based on energy distribution in the cluster. Figure 6.8 shows preliminary results for the di- $\tau$  rates as a function of the leading and sub-leading  $\tau E_T$  and the corresponding efficiencies for a range of online  $\tau E_T$  thresholds that correspond to a 200 kHz trigger rate. Such a selection reaches  $\approx 85\%$  efficiency for offline  $\tau$  candidate thresholds of  $p_T > 45$  GeV and  $p_T > 45$  GeV for leading and subleading  $\tau$  candidates, respectively.

**Regional Tracking Rejection** Regional tracking can be used to further reduce the background rate early in the Event Filter processing. Figure 6.9a shows the rejection versus efficiency for requiring one to five tracks in a cone of  $\Delta R < 0.2$  around the candidate tau as function of the track  $p_T$  requirement. A modified version of this requirement, also shown, only raises the  $p_T$  requirement of the leading track while keeping requirement on additional tracks at  $p_T > 2$  GeV and requiring that those tracks are consistent with coming from the same vertex in  $z$ ; this algorithm gives better performance. For di- $\tau$  triggers, two  $\tau$  candidates can additionally be required to be consistent with coming from the same vertex using a  $\Delta z_0$  cut. This is shown in Fig. 6.9b. The rate can be reduced by a factor of approximately 3.5 with a efficiency loss of  $\approx 4\%$ .

**Event Filter Output** After the regional tracking rejection, the Event Filter adds tracking down to 1 GeV for  $\tau$  candidates using [gHTT](#). The tracks are used in the [BDT](#) and for track-based isolation. Table 6.3 shows rates after the Event Filter selection. The rate of an inclusive isolated di- $\tau$  trigger is too large to record from the Event Filter. Additional requirements such as a  $\Delta R < 3$  selection on the  $\tau$  candidates to veto back-to-back dijets, additional jet and  $b$ -jet requirements, and  $E_T^{\text{miss}}$  requirements will likely be part of the menu. These additional

## 6.5 Tau Leptons



(a) The rejection versus efficiency for tau candidates as a function of the track  $p_T$  (small text next to the points) requirement for various Level-0 selections. Candidates are required to have one to five tracks. The dark blue curve maintains a 2 GeV requirement on the subleading tracks and requires them to be consistent with the  $z_0$  of the leading track.

(b) The rejection versus efficiency for di-tau candidates as a function of the tracking requirement. All three lines have the same Level-0 requirement of  $p_T > 20 \text{ GeV}$  on the leading  $\tau$  candidate and  $p_T > 12 \text{ GeV}$  on the sub-leading  $\tau$  candidate. The green line requires just one to five tracks above the track  $p_T$  shown in the text next to the points. Light blue line then requires the leading  $p_T$  track in each candidates to be consistent with the same vertex (suppressing pile-up). Finally the dark blue line allows the loosens the  $p_T$  requirement on the subleading tracks to 2 GeV and requires them to be consistent with the leading track  $z_0$ .

Figure 6.9: Tau performance using regional tracking (*rHTT*) in the EventFilter. The efficiencies indicated for the cut-based *eFEX* Level-0 algorithm with additional tracking requirements. The right most point on each curve is the Level-0 efficiency with no tracking requirements.

requirements will place significant demands on the Event Filter tracking after the regional-tracking-based cuts.

Table 6.3: Event Filter rates for representative  $\tau$  triggers

Trigger	Offline Threshold [GeV]	Rate
Level-0 Single Tau	150	3 kHz
No regional tracking cuts	-	-
Event Filter Isolated Single Tau	150	0.35 kHz
Level-0 Di- $\tau$	40,30	200 kHz
After regional tracking cuts	40,30	40 kHz
Event Filter Isolated Di- $\tau$	40,30	1.6 kHz
Event Filter Isolated Di- $\tau$ with $0.3 < \Delta R < 3$	40,30	0.5 kHz



---

## 6.6 Jets

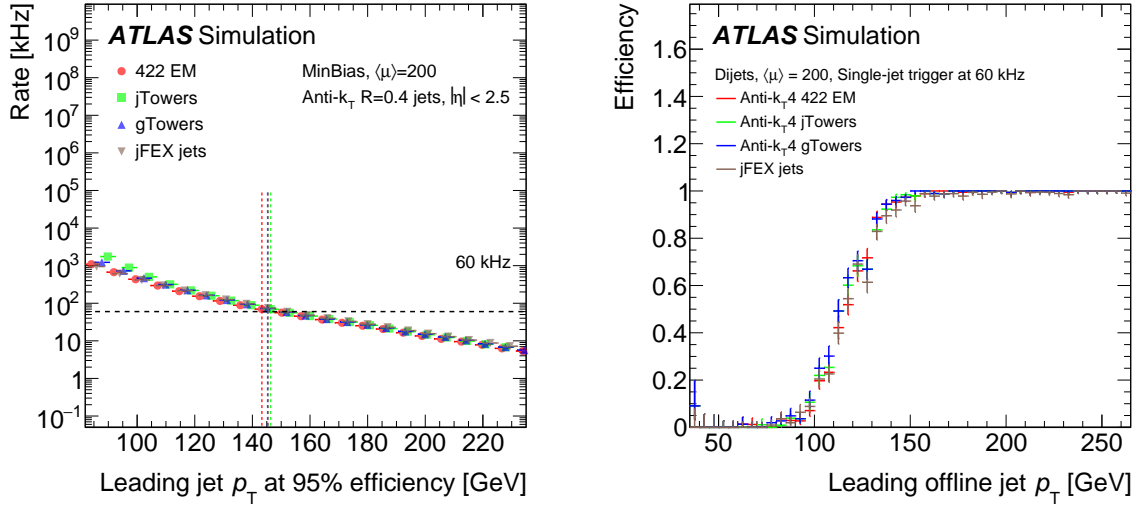
**Level-0** In addition to using offline-inspired topoclusters as inputs, a second important goal of the [Global Trigger](#) is the implementation of an offline-like jet-finding algorithm similar to the anti- $k_t$  algorithm. The anti- $k_t$  algorithm [6.7] used offline iteratively builds jets by merging nearby clusters starting with the highest- $p_T$  clusters until a maximum separation set by the parameter  $R$  is reached. The standard jets used offline in ATLAS are  $R = 0.4$  jets, though  $R = 1.0$  jets, referred to as large- $R$  jets, are used in special cases as discussed in [Section 6.8](#).

Figures [6.11b](#) and [6.11d](#) show that in multi-jet events the [jFEX](#) sliding window algorithm has a significant inefficiency for near-by offline jets, which leads the trigger turn-on to not reach full efficiency even at high offline jet  $p_T$ .

Figures [6.10](#) and [6.11](#) show the rate versus effective offline threshold and trigger efficiencies for single, three, and four jet triggers for the [jFEX](#) algorithm and anti- $k_t$  with [jFEX](#) towers, [gFEX](#) towers, and 422 topoclusters (which are described in [Section 6.2](#)). The effective offline threshold is defined to be the point where the trigger is at 95% of its plateau efficiency (maximum efficiency at high offline jet  $p_T$ ). For each trigger, the efficiencies are shown for an online threshold corresponding to that same rate, so that the 95% efficiency points can be compared. Also shown is the trigger efficiency for anti- $k_t$  clustering for either [jFEX](#) or [gFEX](#) towers. The anti- $k_t$ -based algorithms reach full efficiency at high offline jet  $p_T$ . In addition to improving the performance of close-by jets, the anti- $k_t$  422 topocluster algorithm reaches 95% efficiency about 10 GeV earlier for three and four-jet triggers than the tower-based algorithms. The results shown may be moderately improved through more detailed calibration of the 422 jets.

In order to make the anti- $k_t$  processing feasible in the [Global Trigger](#), it may be necessary to reduce the number of inputs to the iterative anti- $k_t$  algorithm. An example possibility is to bin the topoclusters in  $\eta$  and  $\phi$  to a coarser geometry. This reduces the number of constituents that the iterative algorithm needs to handle without significant impact on the jet performance. [Figure 6.12](#), shows the difference between the resulting jets due to the different thresholds and split/merge algorithms used for reconstructing the topoclusters, and due to binning the clusters before the anti- $k_t$  jet building. For both minbias and a  $Z' \rightarrow t\bar{t}$  signal, the binning versus not-binning makes a small difference in the number of reconstructed jets, while the change in the threshold (422 vs 420) makes a moderate difference. Similarly the variations make only a small difference in the minbias jet  $p_T$  distribution. Finally, the mass of the signal jets is unaffected by the binning and only slightly affected by the threshold difference.

**Regional Tracking Rejection** For multi-jet triggers, the regional tracking can be used to suppress the rate of events where the jets come from different interactions. Jets are required to be consistent with coming from a common vertex using tracks above 2 GeV associated to



(a) Single-jet trigger rates vs effective offline leading jet threshold.

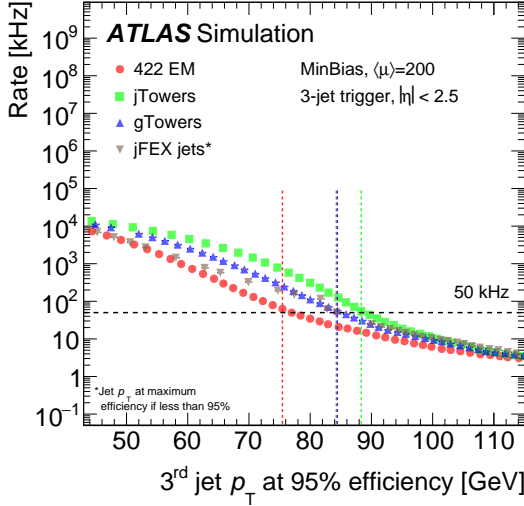
(b) Single-jet trigger efficiency versus leading offline jet  $p_T$  for triggers with a rate of 60 kHz.

Figure 6.10: Level-0 trigger rates versus threshold and event-level trigger efficiencies for single-jet triggers. The *jFEX* algorithm (grey) and anti- $k_t$  with *jFEX* towers (green), *gFEX* towers (blue), and 422 topoclusters (red) are compared.

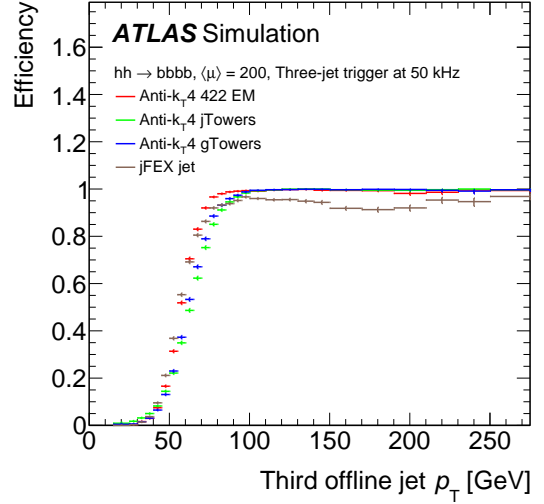
the jets within a cone of  $\Delta R < 0.4$ . The vertex is constructed from all jets above the jet  $p_T$  threshold. A ratio  $R_{p_T}$  is then constructed as the fraction of the sum of the track  $p_T$  associated to the jet and consistent with the primary vertex with respect to the total calorimeter-based measurement of the jet  $p_T$ . Jets associated with the primary vertex will have larger  $R_{p_T}$  than jets from other vertices. Scanning over  $R_{p_T}$  requirements, Fig. 6.13 shows the rejection of jets from other vertices (“pile-up jets”) as a function of the efficiency for jets from the primary vertex (“hard-scatter jets”). The efficiency for 20-40 GeV jets (Fig. 6.13a) reaches a maximum of  $\sim 90\%$  using 2 GeV tracks and 97% using 1 GeV tracks, because some 20-40 GeV jets have no tracks above 2 GeV. Similarly, the efficiency for the 40 GeV jets (Fig. 6.13b) reaches a maximum of  $\sim 90\%$  using 4 GeV tracks and  $\sim 95\%$  using 2 GeV tracks. The resulting rate versus threshold for a four-jet trigger after requiring a common vertex using  $R_{p_T}$  is shown in Fig. 6.14. The efficiency loss from the requirement is 1-2% for the  $HH \rightarrow 4b$  events passing a Level-0 with the same jet thresholds.

To estimate the needed hardware tracking for multi-jet triggers, Fig. 6.15 shows the average number of jets in an event as function of the minimum jet  $p_T$  for events that are selected at Level-0 to have at least four jets with an effective offline  $p_T$  threshold of 55 GeV or 65 GeV. For jets with offline  $p_T > 50$  GeV approximately 5 jets are present per event. This is relevant for the regional tracking needed to confirm a common  $z$  vertex.

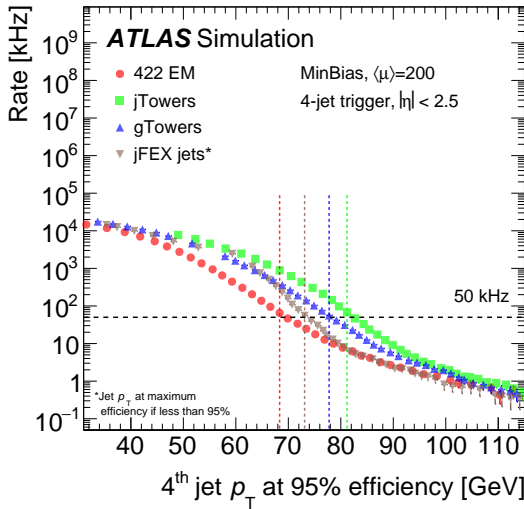




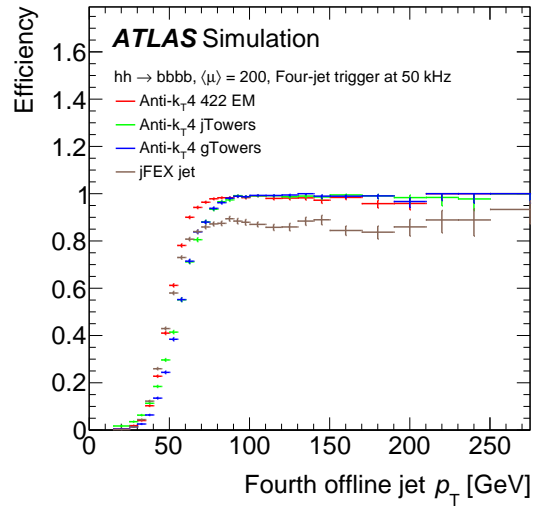
(a) Three-jet trigger rates vs effective offline threshold for the 3<sup>rd</sup> highest  $p_T$  jet.



(b) Three-jet trigger efficiency versus the 3<sup>rd</sup> highest offline jet  $p_T$  for triggers with a rate of 50 kHz.

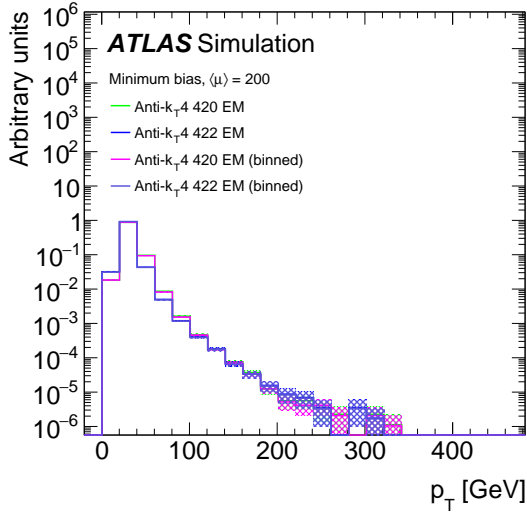


(c) Four-jet trigger rates vs effective offline threshold for the 4<sup>th</sup> highest  $p_T$  jet.

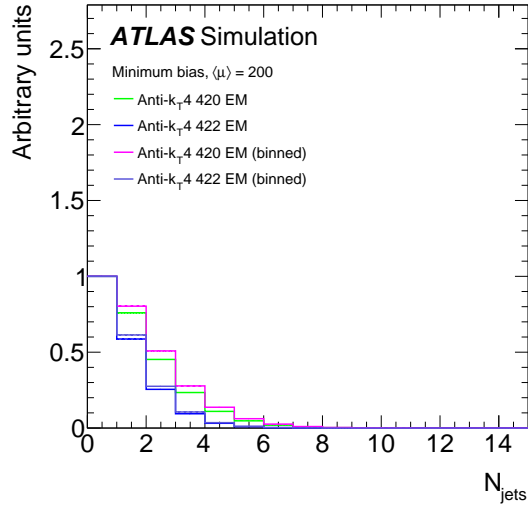


(d) Four-jet efficiency versus the 4<sup>th</sup> highest offline jet  $p_T$  for triggers with a rate of 50 kHz.

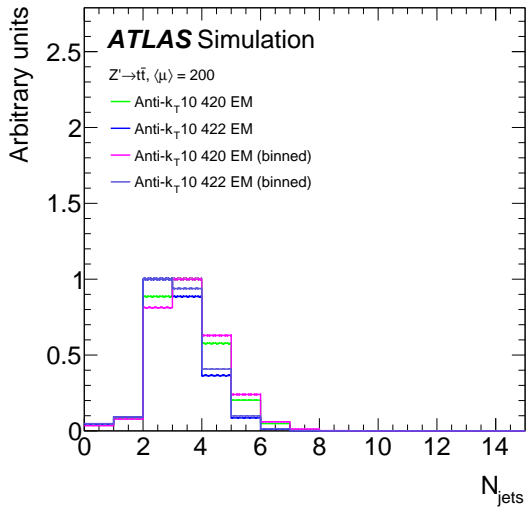
Figure 6.11: Level-0 trigger rates versus threshold and event-level trigger efficiencies for three-jet and four-jet triggers. The *jFEX* algorithm (grey) and anti- $k_t$  with *jFEX* towers (green), *gFEX* towers (blue), and 422 topoclusters (red) are compared.



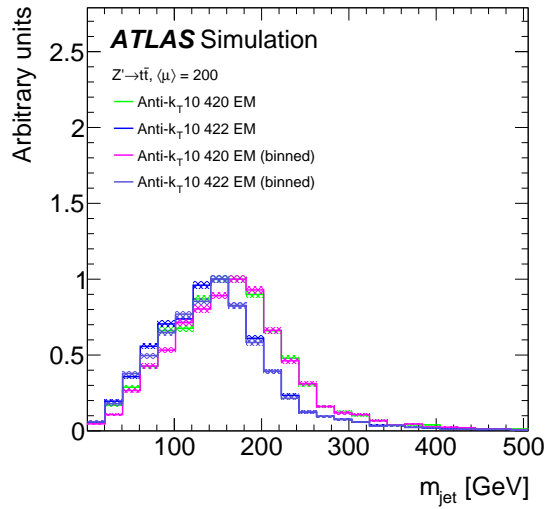
(a)  $p_T$  distribution of jets reconstructed in minimum bias simulation



(b) Number of  $R = 0.4$  jets reconstructed in minimum bias simulation

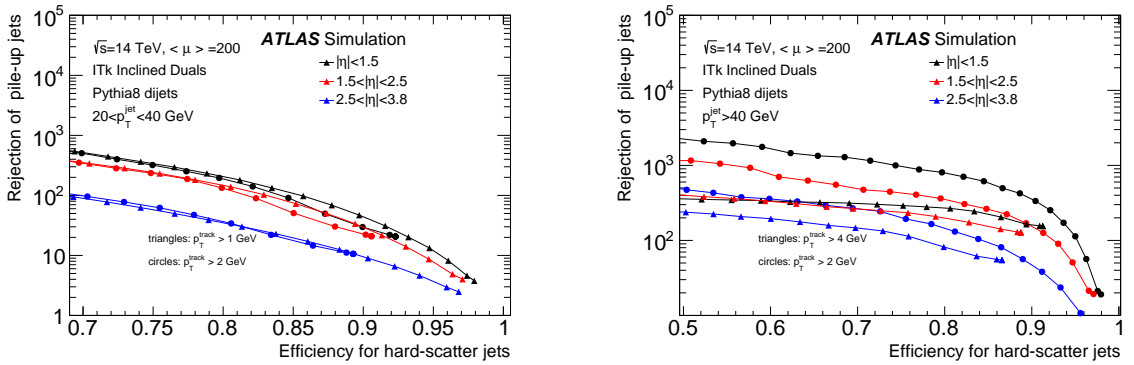


(c) Number of  $R = 1.0$  jets reconstructed in  $Z' \rightarrow t\bar{t}$  simulation



(d) Jet mass reconstructed in  $Z' \rightarrow t\bar{t}$  simulation

Figure 6.12: Comparisons of topocluster-building thresholds and binning on distributions related to jet reconstruction at Level-0.



(a) 20-40 GeV  $p_T$  jets comparing 1 GeV and 2 GeV track  $p_T$  requirements

(b)  $p_T > 40$  GeV jets comparing 2 GeV and 4 GeV track  $p_T$  requirements

Figure 6.13: The rejection of jets from other vertices (“pile-up jets”) as a function of the efficiency for jets from the primary vertex (“hard-scatter jets”) as a scan over the  $R_{p_T}$  requirement.

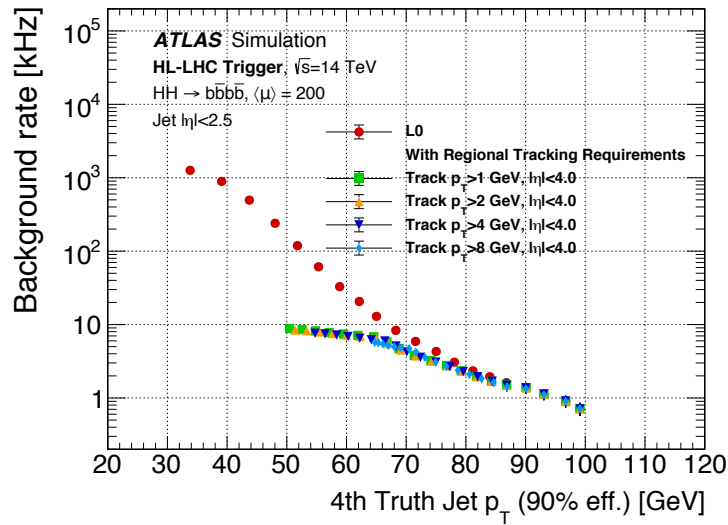


Figure 6.14: Rate for a four-jet trigger requiring the jets to be consistent with a common vertex using  $r_{HTT}$ . The rate is shown as a function of the effective offline threshold for the lowest (“4th”)  $p_T$  jet. The red points show the rate before the tracking requirements. The rate after tracking is shown using 1 GeV (green squares), 2 GeV (yellow upward pointing triangles), 4 GeV (dark blue downward pointing triangles), and 8 GeV (light blue diamonds). For this study, jets using the offline calorimeter only jets using “420” are used, but the tracking rejection is not expected to be correlated with this difference.

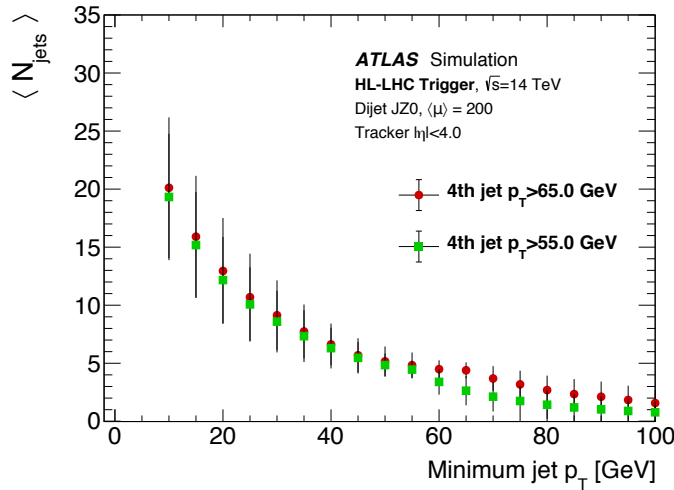


Figure 6.15: The mean number of jets in events as a function of the minimum jet  $p_T$  for events that are selected at Level-0 to have at least 4 jets with  $p_T$  above 50 GeV for two threshold scenarios. The error bars on the plot indicate the RMS of the jet distributions. The x-axis jet energy is the calibrated offline energy.

**Event Filter Output** The Event Filter then uses the full calorimeter data including cells below the  $2\sigma$  noise threshold to reconstruct jets with the same configuration as offline. Even with full offline jet reconstruction, the rate of four-jet events will be too high to reach the threshold needed for key physics goals such as  $HH \rightarrow 4b$ . An acceptable Event Filter output rate would require a  $\sim 100$  GeV threshold. To reduce this rate and preserve the main physics goals and opportunities, a set of additional selections will be applied in the Event Filter. These include  $b$ -tagging and additional jets and  $E_T^{\text{miss}}$  requirements. To allow the Event Filter requirements to be comparable to what is possible offline the target jet  $p_T$  threshold is 20-25 GeV. Figure 6.15 shows that for jets with  $p_T > 20$  GeV, an average of 13 jets per event passes the Level-0 trigger selection. These would cover a substantial fraction of the detector leading to the need for the full-detector tracking provided by the gHTT system. Furthermore, Fig. 6.13 shows that for this range of jet  $p_T$ , a 1 GeV track  $p_T$  threshold is needed to apply vertex constraints with high efficiency. These tracks are also key to enabling  $b$ -tagging and the offline track-based calibration to be used.

Offline jet reconstruction will continue to evolve prior to data-taking. It will likely use an increasing amount of tracking information as in the particle flow algorithm [6.8][6.9]. A key feature of the hardware-based track reconstruction is that it will allow the trigger to follow that evolution to maintain good turn-on curves. Figure 6.16 shows how the current implementation of particle flow on ATLAS depends on the  $p_T$  above which tracks are available. There is only a slight degradation between the current offline minimum  $p_T$  (0.5 GeV) and 1 GeV. By 2 GeV, the performance is significantly degraded.

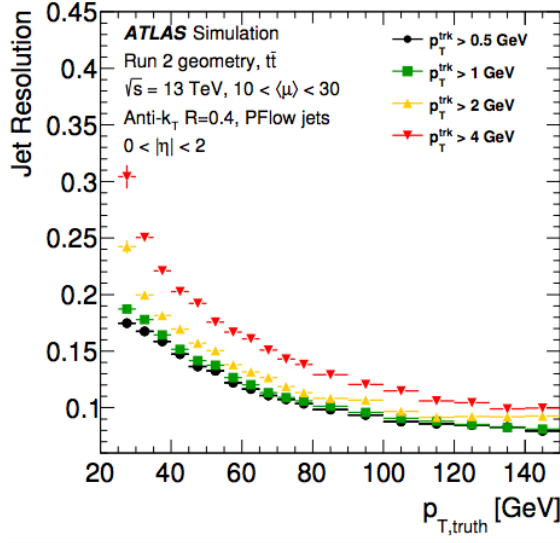


Figure 6.16: The dependence of the jet  $E_T$  resolution for particle flow jets in Run 2 on the input tracking  $p_T$  threshold.

## 6.7 Missing Transverse Energy

Missing transverse energy ( $E_T^{\text{miss}}$ ) reconstruction is a significant challenge at high pile-up for both the trigger and the offline reconstruction. The current offline makes a sum of all reconstructed objects using careful overlap removal to avoid double counting of energy. The  $E_T^{\text{miss}}$  is calculated as the magnitude of  $\mathbf{p}_T^{\text{miss}}$ , which is defined according to the 2-D vector sum

$$\mathbf{p}_T^{\text{miss}} = - \sum_{\text{electrons}} \mathbf{p}_T^e - \sum_{\text{photons}} \mathbf{p}_T^\gamma - \sum_{\text{muons}} \mathbf{p}_T^\mu - \sum_{\text{taus}} \mathbf{p}_T^\tau - \underbrace{\sum_{\text{jets}} \mathbf{p}_T^{\text{jet}}}_{H_T^{\text{miss}}} - \underbrace{\sum_{\text{tracks}} \mathbf{p}_T^{\text{track}}}_{\text{soft term}}. \quad (6.2)$$

The final term, referred to as the “soft” term, involves a sum over all tracks in the full detector coverage. All objects, including the tracks in the soft term, are constrained to be consistent with originating from the primary vertex. The trigger performance described in this section is measured relative to a representative offline  $E_T^{\text{miss}}$  defined as the sum in Equation 6.2, including only jets with  $p_T > 20$  GeV and a soft term, both of which use 1 GeV  $p_T$  tracks for vertex constraints. In general, the  $E_T^{\text{miss}}$  trigger is only used for signatures that are primarily hadronic (jets and soft term) and where the contributions from the other terms are small.

**Level-0** At Level-0, there are several calorimeter-only  $E_T^{\text{miss}}$  options under development. This section will focus on a trigger based on the calorimeter-only jets described in the previous section. This variable is referred to as  $H_T^{\text{miss}}$  in analogy to the standard  $H_T$  variable which is just the scalar sum of the jet  $p_T$ . Here the  $H_T^{\text{miss}}$  is the magnitude of the vector sum of the transverse components of the jets. Also under study is a similar variable which also includes a calorimeter-based soft-term using topoclusters. In addition, because the [Global Trigger](#) has the full event available to an [FPGA](#) algorithm, a wide variety of pile-up mitigation strategies are possible and are under investigation.

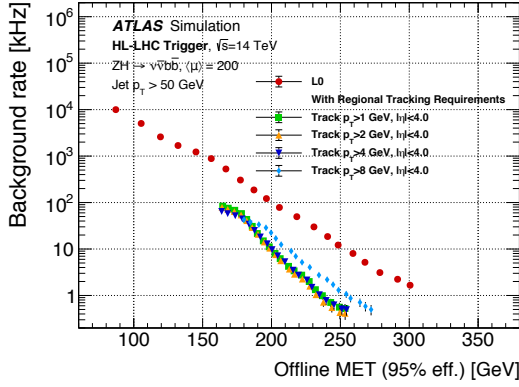
**Regional Tracking Rejection** Regional tracking is used to require that the jets in the  $E_T^{\text{miss}}$  calculation are consistent with a common primary vertex. Figures [6.17a](#) and [6.17b](#) show the Level-0 rate and the rate after regional tracking as a function of the offline  $E_T^{\text{miss}}$  at which the trigger is 95% efficient, for a range of minimum  $p_T$  and maximum  $|\eta|$  of the tracks used in the vertex constraint, respectively. The  $H_T^{\text{miss}}$  here is reconstructed with a 50 GeV threshold on the constituent jets, which is expected to be sufficiently high to suppress contributions from pile-up.

The Level-0 rate shows that the Phase-II upgrade supports a  $E_T^{\text{miss}}$  trigger that is efficient for offline missing energy at  $\sim 210$  GeV with a rate of 80 kHz. For the baseline configuration (tracks with  $p_T$  above 2 GeV and  $|\eta| < 4.0$ ) regional tracking reduces this rate to 4 kHz for the same threshold. While the 8 GeV minimum track  $p_T$  significantly degrades the results, the performance is not strongly sensitive to the  $p_T$  cut below 4 GeV. Reducing the  $\eta$  coverage from  $|\eta| < 4.0$  to  $|\eta| < 3.2$  does not have a strong impact, while reducing to  $|\eta| < 2.5$  raises the rate at a 210 GeV threshold from 4 kHz to 6 kHz.

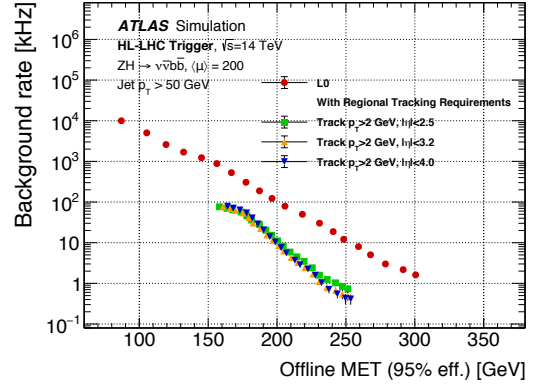
Figures [6.17c](#) and [6.17d](#) show the efficiency versus the true  $E_T^{\text{miss}}$  from invisible particles and the representative offline  $E_T^{\text{miss}}$ . The performance versus the representative offline  $E_T^{\text{miss}}$  is most relevant for analysis, while the performance versus the true  $E_T^{\text{miss}}$  demonstrates that this performance is due to a correlation similar correlation with the true  $E_T^{\text{miss}}$  and not specific to the representative offline  $E_T^{\text{miss}}$  chosen. These figures also show that the effect of the tracking requirement is mainly to remove events with low true  $E_T^{\text{miss}}$  and offline  $E_T^{\text{miss}}$ .

The [rHTT](#) will provide tracking on up to 10% of the [ITk](#) data averaging over all events. To estimate the tracking required to implement the  $E_T^{\text{miss}}$  trigger, Figure [6.18](#) shows the mean number of jets per event for which tracking would be needed as a function of the jet  $p_T$ . For a 50 GeV jet, this is 2.8 jets on average corresponding to 13-14% of the [ITk](#) data. When this is combined with the full set of Phase-II triggers, this is within the capability of the system.

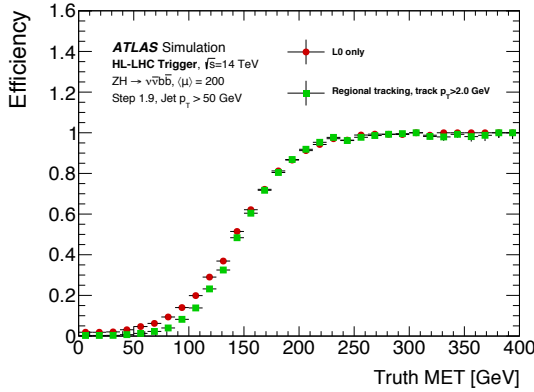
**Event Filter Output** Like the jet reconstruction,  $E_T^{\text{miss}}$  in the Event Filter makes use of the full calorimeter data including cells below the  $2\sigma$  noise threshold and tracking above 1 GeV. For the Event Filter, the  $E_T^{\text{miss}}$  reconstruction will follow as closely as possible the offline development which is expected to evolve in sophistication before the turn-on of the



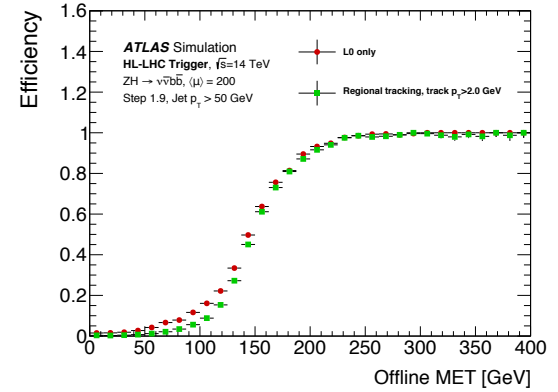
(a) Rate vs offline  $E_T^{\text{miss}}$  at which the trigger is 95% efficient for various minimum track  $p_T$  requirements used in the vertex matching. The full  $|\eta| < 4.0$  tracker coverage is used.



(b) Rate vs offline  $E_T^{\text{miss}}$  at which the trigger is 95% efficient for various track  $\eta$  coverage assumption used in the vertex matching. A minimum track  $p_T$  of 2 GeV is assumed.



(c) Efficiency as a function of true  $E_T^{\text{miss}}$  for the Level-0 and after a regional tracking requirement with  $|\eta| < 4.0$  and a minimum track  $p_T$  of 2 GeV.



(d) Efficiency as a function of offline  $E_T^{\text{miss}}$  for the Level-0 and after a regional tracking requirement with  $|\eta| < 4.0$  and a minimum track  $p_T$  of 2 GeV.

Figure 6.17: Performance of the calorimeter-based  $E_T^{\text{miss}}$  trigger at Level-0 based on the  $H_T^{\text{miss}}$  variable (50 GeV jets) and the same variable after regional tracking requirements that the jets used are consistent with a common vertex.

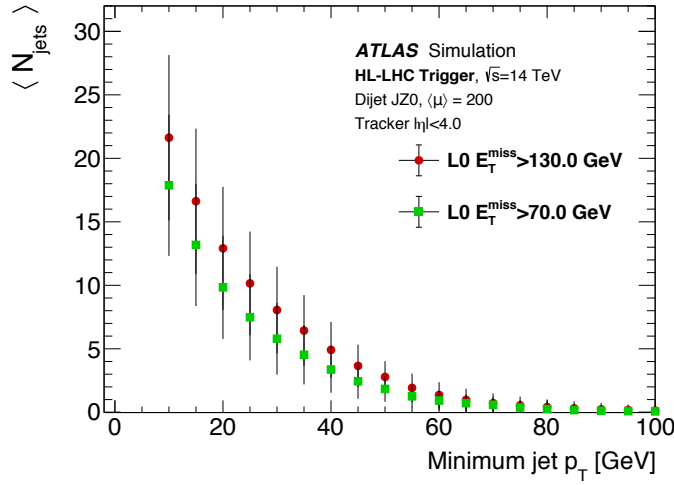


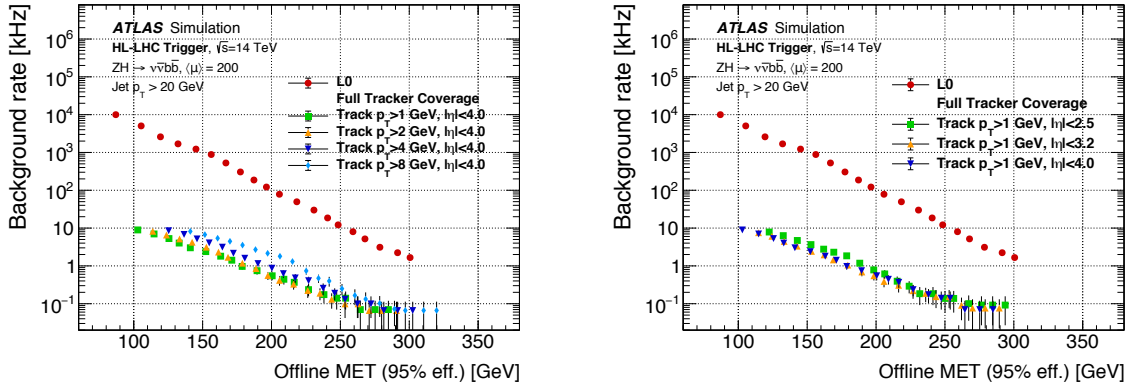
Figure 6.18: The mean number of jets for events selected by a Level-0  $E_T^{\text{miss}}$  trigger for two threshold scenarios (L0 MHT in legend). The errors bars on the plot indicate the RMS of the jet distributions. The x-axis jet energy is the calibrated offline energy.

**HL-LHC.** It is therefore important that the system be able to support a range of algorithms including the use of tracking.

In order to quantify the impact of the trigger system capabilities on the ability to reproduce the offline  $E_T^{\text{miss}}$ , a representative  $E_T^{\text{miss}}$  calculation similar to the current offline, previously described, is explored for its dependence of the minimum tracking  $p_T$  and  $\eta$  coverage. Figures 6.19a and 6.19b show the dependence of performance of this definition of  $E_T^{\text{miss}}$  on the minimum  $p_T$  and the  $\eta$  coverage of the tracking, respectively. It is assumed that 20 GeV jets will be used for the jet term. The 1 GeV minimum  $p_T$  scenario marginally outperforms the 2 GeV scenario and the 4 GeV has approximately 30% more rate for the same threshold. Similarly, reducing the  $\eta$  coverage from  $|\eta| < 4.0$  to  $|\eta| < 3.2$  has a small impact while reducing to  $|\eta| < 2.5$  raises the rate by  $\approx 40\%$ . It should also be noted that the impact of the  $\eta$  coverage is model dependent. In Run 2, it was seen that the optimisation that improved the performance from  $ZH \rightarrow \nu\nu$  degraded the performance for  $\text{VBF } H \rightarrow \text{invisible}$ , which is likely to be the case with reduced  $\eta$  coverage.

In order to perform the full  $E_T^{\text{miss}}$  calculation in Eqn. 6.2, full-detector tracking is needed to calculate the soft term. Figures 6.20a and 6.20b show the rates for the  $H_T^{\text{miss}}$  variable which does not include the soft-term. The rates are  $\approx 20\%$  larger and the degradation for reductions in  $p_T$  and  $\eta$  coverage are similar to the  $E_T^{\text{miss}}$  variable at a 210 GeV threshold. For lower thresholds, the degradation from removing the soft-term is larger ( $\approx 50\%$ ). The low  $E_T^{\text{miss}}$  region will be important when a  $E_T^{\text{miss}}$  requirement is used in combination with other requirements, e.g. in di- $\tau$  or multi-jet Level-0 triggers.





(a) Rate vs offline  $E_T^{\text{miss}}$  at which the trigger is 95% efficient for various minimum track  $p_T$  requirements used in the vertex matching. The full  $|\eta| < 4.0$  tracker coverage is used.

(b) Rate vs offline  $E_T^{\text{miss}}$  at which the trigger is 95% efficient for various track  $\eta$  coverage assumptions used in the vertex matching. A minimum track  $p_T$  of 1 GeV is assumed.

Figure 6.19: Rate vs offline  $E_T^{\text{miss}}$  at which the trigger is 95% efficient for a  $E_T^{\text{miss}}$  trigger based on the  $E_T^{\text{miss}}$  variable (20 GeV jets) which includes a track-based soft-term in the Event Filter using *gHTT*.

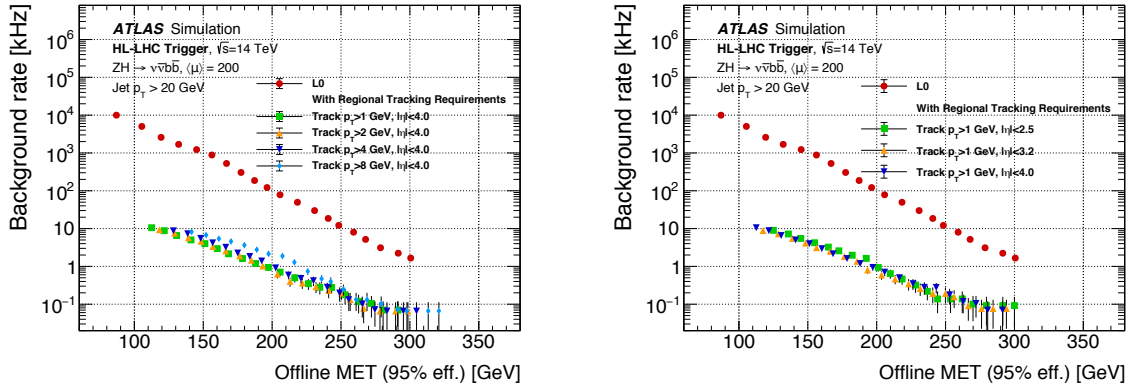
Finally, Fig. 6.21a illustrates the impact of soft-term on the efficiency as a function of true  $E_T^{\text{miss}}$ . However since analyses use the offline  $E_T^{\text{miss}}$ , Figure 6.21b shows the efficiency as a function of offline  $E_T^{\text{miss}}$  for  $H_T^{\text{miss}}$  and  $E_T^{\text{miss}}$  with a minimum track  $p_T$  to 2 GeV. By definition, the efficiency as a function of offline  $E_T^{\text{miss}}$  for the  $E_T^{\text{miss}}$  variable with a 1 GeV minimum track  $p_T$  is a step function at 200 GeV, so the shape of these curves shows the effect of raising the minimum track  $p_T$  to 2 GeV.

If the soft-term is excluded, a large area of the detector is still needed for tracking in order to support the 20 GeV jet threshold. Figure 6.18 shows that for a Level-0 threshold efficient at 200 GeV (130 GeV online), events have on average 13 jets with  $p_T > 20$  GeV.

## 6.8 Boosted-object Triggers

At Level-0, the *Global Trigger* will reconstruct  $R = 1.0$  jets from topoclusters using the anti- $k_t$  algorithm. The resulting performance for this trigger is shown in Fig. 6.22. The 422 topocluster-based jets outperform the *gFEX* jets give a threshold for 35 kHz of 365 GeV instead of 380 GeV. In order to further reduce the threshold for the same rate, two selections have been found to be very effective in Run 2 and can be implemented in the *Global Trigger*: jet trimming [6.10] requires that sub-jets within the large- $R$  jet have a minimum fraction of the jet energy to be counted as part of the jet, and jet mass requires that the remaining constituents have a minimum mass. Similar selections are used in offline analysis. Figure 6.23 shows the impact on the efficiency of these two selections on the Run 2 large- $R$  jet trigger.

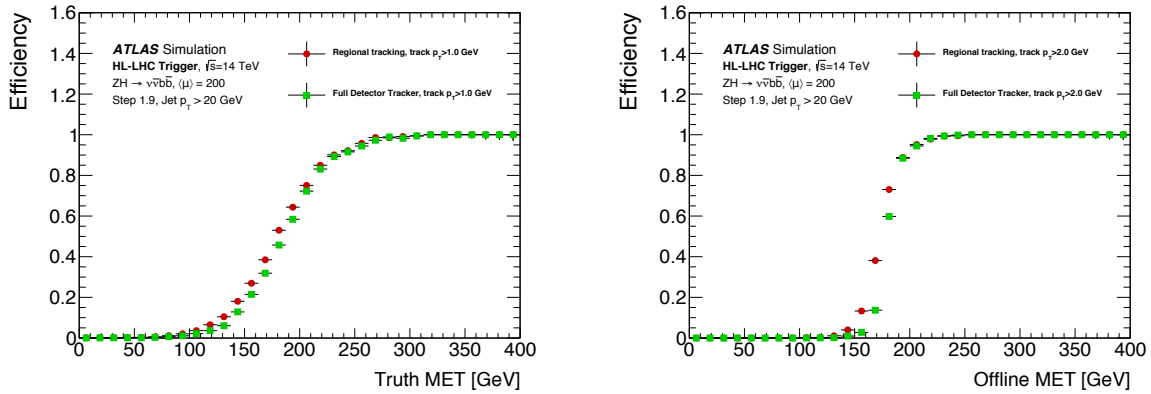
## 6.8 Boosted-object Triggers



(a) Rate vs offline  $E_T^{\text{miss}}$  at which the trigger is 95% efficient for various minimum track  $p_T$  requirements used in the vertex matching. The full  $|\eta| < 4.0$  tracker coverage is used.

(b) Rate vs offline  $E_T^{\text{miss}}$  at which the trigger is 95% efficient for various track  $\eta$  coverage assumptions used in the vertex matching. A minimum track  $p_T$  of 1 GeV is assumed.

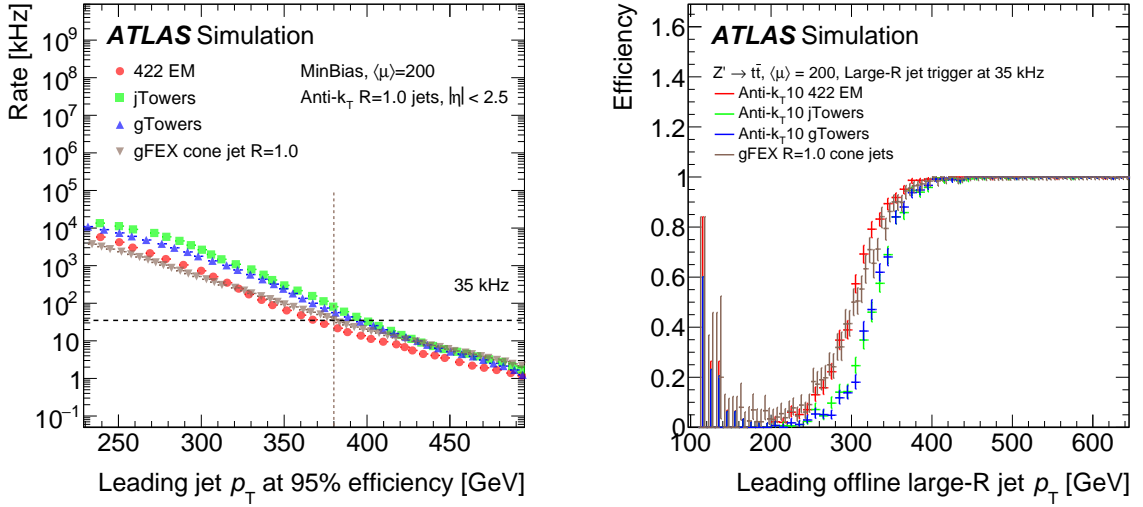
Figure 6.20: Rate vs offline  $E_T^{\text{miss}}$  at which the trigger is 95% efficient for a  $E_T^{\text{miss}}$  trigger based on the  $H_T^{\text{miss}}$  variable (20 GeV jets) in the Event Filter using *gHTT*.



(a) Efficiency as a function of true  $E_T^{\text{miss}}$  for an  $H_T^{\text{miss}}$  variable and an  $E_T^{\text{miss}}$  variable which includes the track soft term with tracking coverage over  $|\eta| < 4.0$  and a minimum track  $p_T$  of 1 GeV.

(b) Efficiency as a function of offline  $E_T^{\text{miss}}$  for an  $H_T^{\text{miss}}$  variable and an  $E_T^{\text{miss}}$  variable which includes the track soft term with tracking coverage over  $|\eta| < 4.0$  and a minimum track  $p_T$  of 2 GeV. The  $E_T^{\text{miss}}$  variable with a minimum track  $p_T$  of 1 GeV would be by definition a step function at 200 GeV.

Figure 6.21: Efficiency as a function of true  $E_T^{\text{miss}}$  and offline  $E_T^{\text{miss}}$  in the Event Filter using global tracking.



(a) Large- $R$  ( $R = 1$ ) jet trigger rates vs effective off-line threshold.

(b) Large- $R$  ( $R = 1$ ) jet trigger efficiency versus off-line jet  $p_T$  for triggers with a rate of 35 kHz.

Figure 6.22: Level-0 rate and efficiency for a large- $R$  jet trigger. The *gFEX* algorithm (grey) corresponds to the performance of the Phase-I system. The 422 EM (red) shows the performance of the proposed system which has a lower threshold for the same rate. The blue and green curves show performance of the anti- $k_T$  algorithm applied to the Phase-I granularity (a feature which is only possible in the Phase-II system) which is less performant than the 422 EM based result.

The red line shows the efficiency for large- $R$  jets without additional requirements. The blue line shows the efficiency after adding a jet trimming requirement that sub-jets reconstructed with  $R=0.2$  have at least  $f = 0.04$  of the total jet energy. The green line shows the efficiency after adding a further jet mass requirement of  $m_{\text{jet}} > 30$  GeV. The online thresholds for each curve are adjusted to give the same rate, so by adding the two additional requirements, the threshold can be reduced by  $\approx 50$  GeV. A similar improvement is expected from an implementation in the [Global Trigger](#).

## 6.9 Exotic-object Triggers

As described in Section 2.8, some new physics scenarios produce signatures which are not identified by the standard set of reconstructed objects. There are several on-going activities to address these cases. In particular, the loss of efficiency for near-by muons in the Run 2 and Phase-I system will be addressed with dedicated algorithms to search for second muon candidates near other muon candidates. Other related topics include investigations of delayed muon triggers where an event is triggered by a muon in subsequent event (1-2 bunch crossings later) and high-impact parameter muons, which are due to particles traveling a significant distance from the beam-spot before decaying.

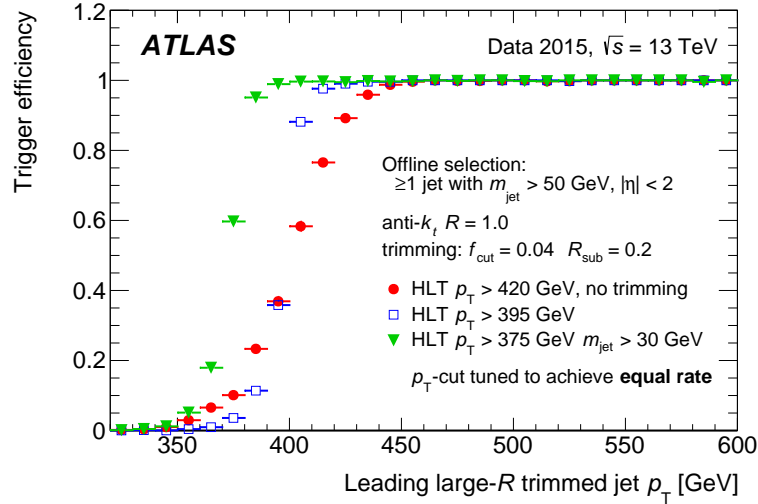


Figure 6.23: Trigger efficiency versus offline jet  $p_T$  for three large- $R$  triggers in Run 2. The threshold for each selection has been adjusted to give equal trigger rate, so better performance corresponds to having efficiency for lower jet  $p_T$  for the same rate. The red line shows large- $R$  jets without additional requirements. The blue line adds a jet trimming requirement that sub-jets reconstructed with  $R = 0.2$  have at least  $f = 0.04$  of the total jet energy. The green line adds a further jet mass requirement of  $m_{jet} > 30$  GeV.

## 6.10 Inclusive Vector Boson Fusion Trigger

A study has been conducted to implement an inclusive **VBF** Higgs trigger assuming that the **ffEX** provides calorimeter-based jets similar to the offline reconstruction. A candidate selection requires two trigger jets separated in pseudo-rapidity by  $\Delta\eta > 2.5$ . They are further required not to be back-to-back as in dijet production ( $\Delta\phi < 2.5$ ). The jet  $p_T$  threshold was then scanned to explore possible working points. The results show that an acceptance of 6.6% for all Higgs bosons produced by the **VBF** mechanism can be achieved with a jet  $p_T$  requirement of 75 GeV corresponding to a Level-0 rate of 33 kHz, which can be reduced to  $\sim 5$  kHz by requiring the jets to be consistent with a common  $z$  vertex using regional tracking. The cross-section for the **VBF** process at  $\sqrt{s} = 14$  TeV is 4.28 pb [6.11], compared to the  $WH$ ,  $W \rightarrow \ell\nu$  cross-section times branching fraction of 0.34 pb and 0.067 pb for  $ZH$ ,  $Z \rightarrow \ell\ell$ . The **VBF** inclusive trigger has comparable sensitivity to  $WH$  and far exceeding  $ZH$ . Furthermore, because of the neutrino in the leptonic  $W$  decay, the  $WH$  has a limitation with respect to channels involving  $E_T^{\text{miss}}$ .

This trigger allows 6.6% of the **VBF** Higgs processes to reach the Event Filter, but in order to read those events out to offline storage, additional analysis-specific requirements will be needed. These additional selections can then take advantage of the Event Filter capabilities, including full-detector tracking at 1 GeV as needed. In the menu presented in Section 6.11,

---

an Event Filter rate for the resulting trigger is allocated, which would be composed of a large set of analysis-specific selections.

## 6.11 Example Run 4 Trigger Menu

The full set of triggers used in the system is referred to as the menu. The menu, shown in Table 6.4, is constructed to meet the physics requirements given in Chapter 2 and is representative of the primary triggers that would be used. The full trigger menu for Run 2 consists of more than 2000 selections, so Table 6.4 should only be considered as representative.

The procedures used to estimate the rates are described in Section 6.1. The totals shown are the sum of the rates for the individual channels. Some reduction in the total is expected and is due to overlaps between the channels ( $\sim 10 - 20\%$ ). In addition, analysis-specific, monitoring, and support triggers would be included in a full menu. The Event Filter rate is near the limit, but the Event Filter system provides the largest flexibility for including additional selections to reduce the rates. If needed, this would include analysis-specific cuts which reduce rates at the expense of the generality of the resulting dataset.

The menu also includes resources reserved for supporting triggers, which enable trigger efficiency measurements and a wide variety of performance measurements. The allocations are based on the fraction of the current Run 2 menu used for supporting triggers at each level (10% at Level-0 and after regional tracking, and 20% in the Event Filter) and are indicated with the "Supporting Trigs" line.

The Phase-II Event Filter menu can be compared to the current Run 2 menu scaled by luminosity. The current Run 2 menu operates at 1500 Hz for a luminosity of  $\mathcal{L} = 2.0 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ . Scaled to  $\mathcal{L} = 7.5 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ , this gives 6.6 kHz. The majority of the additional rate in the Phase-II menu comes from these sources:

- Level-0 seeded large- $R$  jets, 0.5 kHz additional rate
- Forward electrons, 0.2 kHz additional rate
- Inclusive VBF triggers, 0.5 kHz additional rate
- Lower single-lepton  $p_T$  threshold, 0.9 kHz additional rate
  - For comparison, this 30% higher rate gives acceptance gains of 16% for inclusive  $W$  and  $t\bar{t}$  production, 28% acceptance gain for  $HH \rightarrow bb\tau\tau$  with one  $\tau$  decaying to leptons, and 47% for the "Well-tempered neutralino" compressed SUSY model introduced in Section 2.1.
- Lower dilepton  $p_T$  thresholds, 0.25 kHz additional rate
  - This leads to 70% more VBF  $H \rightarrow \tau\tau$  acceptance and  $\approx 3$  times more acceptance for the SUSY model shown in Fig. 2.4
- More inclusive di- $\tau$  trigger, 0.15 kHz additional rate

## 6.11 Example Run 4 Trigger Menu

Table 6.4: Representative trigger menu for 1 MHz Level-0 rate. The offline  $p_T$  thresholds indicate the momentum above which a typical analysis would use the data.

Trigger Selection	Run 1 Offline $p_T$ Threshold [GeV]	Run 2 (2017) Offline $p_T$ Threshold [GeV]	Planned HL-LHC Offline $p_T$ Threshold [GeV]	L0 Rate [kHz]	After regional tracking cuts [kHz]	Event Filter Rate [kHz]
isolated single $e$	25	27	22	200	40	1.5
isolated single $\mu$	25	27	20	45	45	1.5
single $\gamma$	120	145	120	5	5	0.3
forward $e$			35	40	8	0.2
di- $\gamma$	25	25	25,25		20	0.2
di- $e$	15	18	10,10	60	10	0.2
di- $\mu$	15	15	10,10	10	2	0.2
$e - \mu$	17,6	8,25 / 18,15	10,10	45	10	0.2
single $\tau$	100	170	150	3	3	0.35
di- $\tau$	40,30	40,30	40,30	200	40	0.5 <sup>+++</sup>
single $b$ -jet	200	235	180	25	25	0.35 <sup>+++</sup>
single jet	370	460	400			0.25
large- $R$ jet	470	500	300	40	40	0.5
four-jet (w/ $b$ -tags)		45 <sup>†</sup> (1-tag)	65(2-tags)			
four-jet	85	125	100	100	20	0.1
$H_T$	700	700	375	50	10	0.2 <sup>+++</sup>
$E_T^{\text{miss}}$	150	200	210	60	5	0.4
VBF inclusive			2x75 w/ ( $\Delta\eta > 2.5$ & $\Delta\phi < 2.5$ )	33	5	0.5 <sup>+++</sup>
$B$ -physics <sup>++</sup>				50	10	0.5
Supporting Trigs				100	40	2
Total				1066	338	10.4

<sup>†</sup> In Run 2, the 4-jet  $b$ -tag trigger operates below the efficiency plateau of the Level-1 trigger.

<sup>++</sup> This is a place-holder for selections to be defined.

<sup>+++</sup> Assumes additional analysis specific requires at the Event Filter level

The [Global Trigger](#) and [HTT](#) elements give the proposed system significant flexibility to construct such specialised triggers either to address science goals that are not met by the inclusive object-based selections shown in the example menu, such as exotic signatures, or where inclusive selections have a rate that is too high. This is particularly important for the Event Filter where additional algorithms are not difficult to include provided a large computing cost is not implied. In many cases, the Run 2 performance is limited by the computing cost of tracking which in the upgraded system can be addressed with the [HTT](#).

Table 6.5: Estimation of regional tracking needs. The fourth column indicates the  $\eta - \phi$  fraction of the detector for which tracking is needed, but as described in the text, a significantly larger fraction of the data is needed to find the corresponding tracks because of the length of the beam spot in  $z$ , the spread of the tracks in  $\phi$  due to the track curvature, and the effects of module boundaries. The impact of those effects on three example layers are shown in columns 5-7. The small differences between bottom row numbers and the layer by layer maxima in Figure 6.26 are due to a more detailed accounting of the  $\eta$  dependence of the objects for the figure.

Trigger Selection	Object Multiplicity	Region Size (% detector)	Regional Tracking (rHTT) Need Per Event			
			% $\eta - \phi$ coverage	Detector Data Fraction		
				Pixel Layer 5	Strip Layer 1	Strip Layer 4
isolated single $e$	1	$0.2 \times 0.2$	0.13%	1.0%	1.3%	0.4%
isolated single $\mu$	Not Used					
single $\gamma$	Not Used					
forward $e$	1	$0.2 \times 0.2$	0.13%	1.0%	1.3%	0.4%
di- $\gamma$	Not Used					
di- $e$	2	$0.2 \times 0.2$	0.25%	2.0%	2.6%	0.8%
di- $\mu$	2	$0.2 \times 0.2$	0.25%	2.0%	2.6%	0.8%
$e - \mu$	2	$0.2 \times 0.2$	0.25%	2.0%	2.6%	0.8%
single $\tau$	Not Used					
di- $\tau$	2	$0.2 \times 0.2$	0.25%	2.0%	2.6%	0.8%
single jet	Not Used					
large- $R$ jet	Not Used					
four-jet	5	$0.8 \times 0.8$	10.2%	23.0%	25.5%	14.5%
$H_T$	5	$0.8 \times 0.8$	10.2%	23.0%	25.5%	14.5%
$E_T^{\text{miss}}$	3	$0.8 \times 0.8$	5.7%	12.9%	14.3%	8.1%
VBF inclusive	2	$0.8 \times 0.8$	4.1%	9.2%	10.2%	5.8%
Supporting Trigs	10% of total rate		0.2%	0.5%	0.6%	0.3%
Averages per event weighted by rates			2.3%	6.0%	6.8%	3.5%

## 6.12 Tracking Needs

The rHTT and gHTT are critical to achieving the menu presented in the previous section. An estimate of the tracking rates required to achieve this menu is presented in Tables 6.5 and 6.6 respectively.

The rHTT system provides regional tracking at the full Level-0 accept rate (1 MHz). It uses a reduced set of ITk layers which provides sufficient resolution for verifying the presence of tracks down to 2 GeV and significantly reducing backgrounds which have objects from multiples vertices. Because of the reduced number of layers used by the rHTT system, the resulting tracks are not appropriate for  $b$ -tagging. The regional tracking estimate is calculated assuming that  $e$ ,  $\mu$ , and  $\tau$  use  $\eta \times \phi$  regions that are  $0.2 \times 0.2$ , small- $R$  jets use tracking in regions that are  $0.8 \times 0.8$ , and large- $R$  jets use tracking in regions that are  $2.0 \times 2.0$ . Be-



Table 6.6: Estimation of global tracking needs.

Trigger Selection	Object Multiplicity	Region Size (% detector)	Global Tracking (gHTT) Need Per Event		Use Cases
			% $\eta - \phi$ coverage	Effective Rate (kHz)	
isolated single $e$	1	$0.4 \times 0.4$	0.51%	0.20	isolation
isolated single $\mu$	1	$0.4 \times 0.4$	0.51%	0.13	isolation
single $\gamma$	1	$0.4 \times 0.4$	0.51%	0.03	isolation
forward $e$	1	$0.4 \times 0.4$	0.51%	0.04	isolation
di- $\gamma$	2	$0.4 \times 0.4$	1.02%	0.20	isolation
di- $e$	2	$0.4 \times 0.4$	1.02%	0.10	isolation
di- $\mu$	2	$0.4 \times 0.4$	1.02%	0.05	isolation
$e - \mu$	2	$0.4 \times 0.4$	1.02%	0.10	isolation
single $\tau$	1	$0.4 \times 0.4$	0.51%	0.02	isolation
di- $\tau$		full event	100%	5	$b$ -tagging, $E_T^{\text{miss}}$ soft-term
single jet		full event	100%	25	$b$ -tagging, soft hadronic
large- $R$ jet	1	$2.0 \times 2.0$	13%	5.1	$b$ -tagging, pile-up
four-jet		full event	100%	20	$b$ -tagging, soft jets to seed high jet multiplicity
$H_T$		full event	100%	10	$b$ -tagging, soft jets to seed high jet multiplicity
$E_T^{\text{miss}}$		full event	100%	5	$b$ -tagging, $E_T^{\text{miss}}$ soft term
VBF inclusive		full event	100%	5	$E_T^{\text{miss}}$ soft-term, soft jets to seed high jet multiplicity
Supporting Trigs				15	20% of total
Total Rate				91	

cause of the beam spot spread in  $z$ , the bending of tracks in  $\phi$ , and the tracking detector module boundaries, significantly more data must be used than is represented by the geometrical  $\eta - \phi$  regions for which tracking is needed. To estimate this effect, the frequency that a detector module would be included in a tracking request is calculate for various size objects ( $0.2 \times 0.2$ ,  $0.8 \times 0.8$ , or  $2.0 \times 2.0$ ) distributed uniformly over either  $|\eta| < 2.5$  or  $|\eta| < 4.0$ . For example, Figures 6.24 shows  $0.2 \times 0.2$  requests distributed over  $|\eta| < 2.5$  which are relevant for  $e$ ,  $\mu$ , and  $\tau$  objects and Figure 6.25 shows  $0.8 \times 0.8$  request distributed over  $|\eta| < 4.0$  which are appropriate for jets in  $E_T^{\text{miss}}$  calculations. The multiplicities of  $e$ ,  $\mu$ , and  $\tau$  are assumed to be the number in the trigger (a small underestimate). For the multi-jet,  $E_T^{\text{miss}}$  triggers, and  $H_T$  triggers, the jet multiplicities are estimated from simulation as a function of the  $p_T$  relevant for the selection, see Figs. 6.15 and 6.18. For regional tracking, a 50 GeV jet threshold is expected. The result is that averaged over the representative menu,



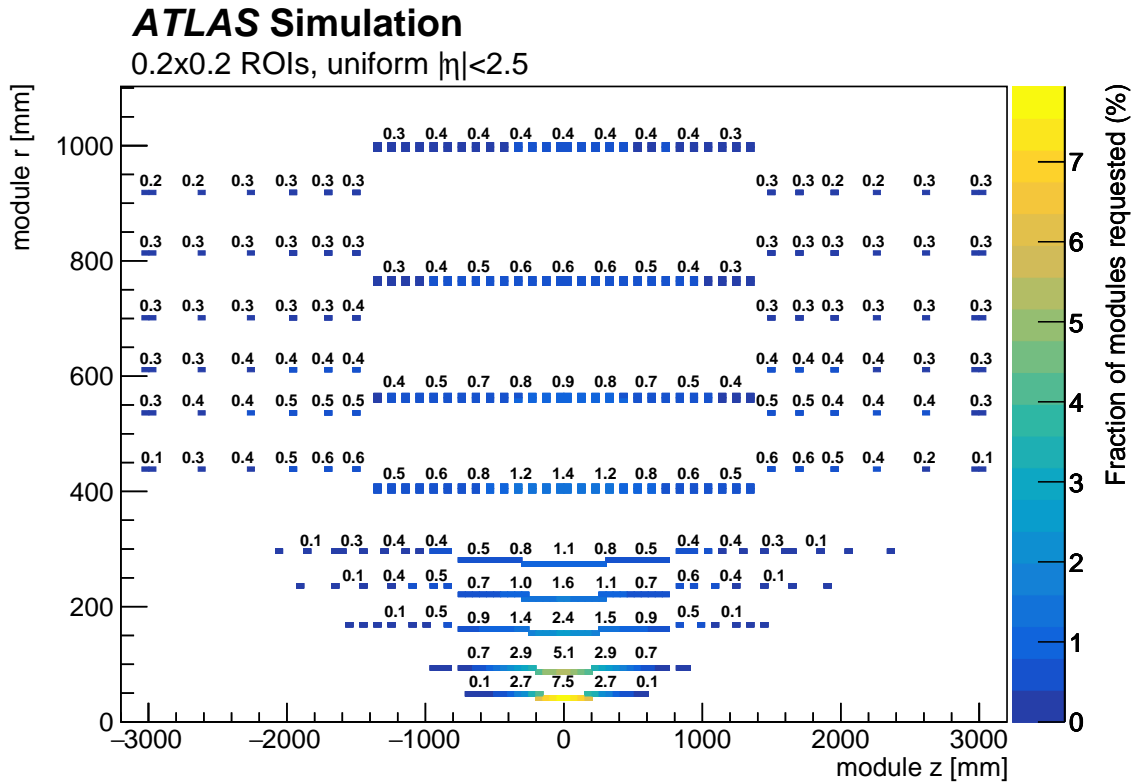


Figure 6.24: Fraction of tracking requests needing a module for  $0.2 \times 0.2$  objects uniformly distributed over  $|\eta| < 2.5$ .

approximately 2.3% of  $\eta - \phi$  coverage is need for regional tracking, but the data required for that tracking is as much as 3.5 to 6.8% for the more central modules in the layers used in the [rHTT](#) system. This is shown for all the modules positions in the ITK system in Figure 6.26. This figure is constructed by summing readout fraction distributions for individual objects, such as those shown in Figures 6.24 and 6.25, multiplied by the object multiplicity for the event type and the fraction of events for which that trigger is active (there small differences between bottom row numbers in Table 6.5 and the layer by layer maxima in Figure 6.26 due to a more detailed accounting of the  $\eta$  dependence of the objects used in the figure). To leave a margin, and because the menu is not complete, regional tracking is planned to be available for 10% of the detector data per event on average (individual events may have significantly more than the average).

The [gHTT](#) system provides high-resolution full-scan tracking down to 1 GeV appropriate for  $b$ -tagging either in regions or for the full-detector. The offline reconstruction is also expected to maintain a 1 GeV track threshold [6.12][6.13]. Use cases that are being considered for [gHTT](#) include  $b$ -tagging and track-based calibration on all jets with  $p_T > 20$  GeV, calculation of variables such as the  $E_T^{\text{miss}}$  soft-term, pile-up correction/mitigation, and searching

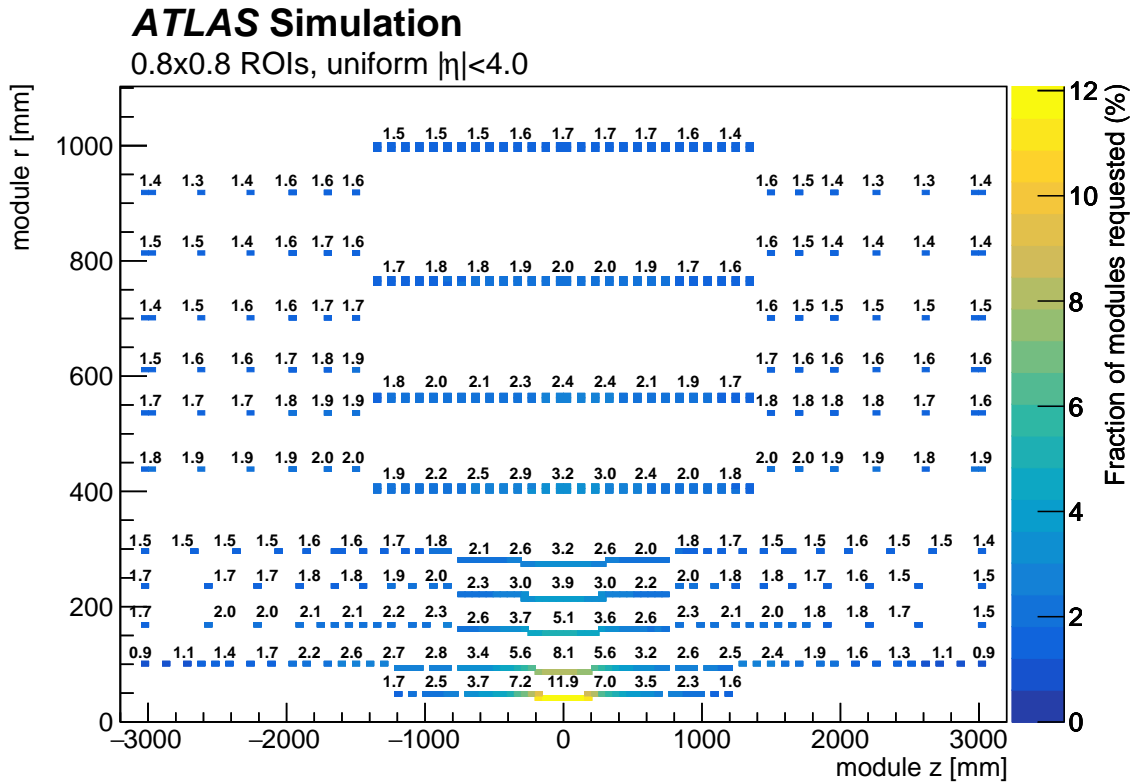


Figure 6.25: Fraction of tracking requests needing a module for  $0.8 \times 0.8$  objects uniformly distributed over  $|\eta| < 4.0$ .

for additional soft-jets using full-detector tracking. Supporting triggers for efficiency calculations and other performance studies are assumed to operate at 20% of the primary triggers. These use cases are based on examination of the current Run 2 menu which has been developed based on extensive collaboration between the trigger group and the data analysis teams. The  $b$ -tagging is used in single jet, multi-jet, large- $R$  jet,  $E_T^{\text{miss}}$ ,  $H_T$ , and di- $\tau$  triggers to reduce significantly reduce rates while preserving physics with flavour-specific final states, such as  $HH \rightarrow 4b$  or  $bb\tau\tau$ , electroweak **SUSY** to  $4b + E_T^{\text{miss}}$ , and **RPV SUSY**. Searching for additional soft-jets in the Run 2 menu is primarily used in the multi-jet triggers where the rate that can be brought to the Event Filter is substantially higher than what can be recorded; in particular, higher jet multiplicity triggers are build on four jet triggers. With the novel **VBF** inclusive trigger and the  $E_T^{\text{miss}}$  trigger, it is expected that soft hadronic final states will be important, for example for the search for the exotic Higgs boson decay  $H \rightarrow 4b$ .

Figs. 6.15 and 6.18 show that for both 4-jet and  $E_T^{\text{miss}}$  events, there are on average 13 jets with  $p_T > 20$  GeV. This large number combined with the effect of the beam spot spread in  $z$  and the bending of tracks in  $\phi$  means that essentially the entire detector is needed for  $b$ -tagging

## ATLAS Simulation Phase-II TDAQ Trigger Menu (TDR Table 6.4)

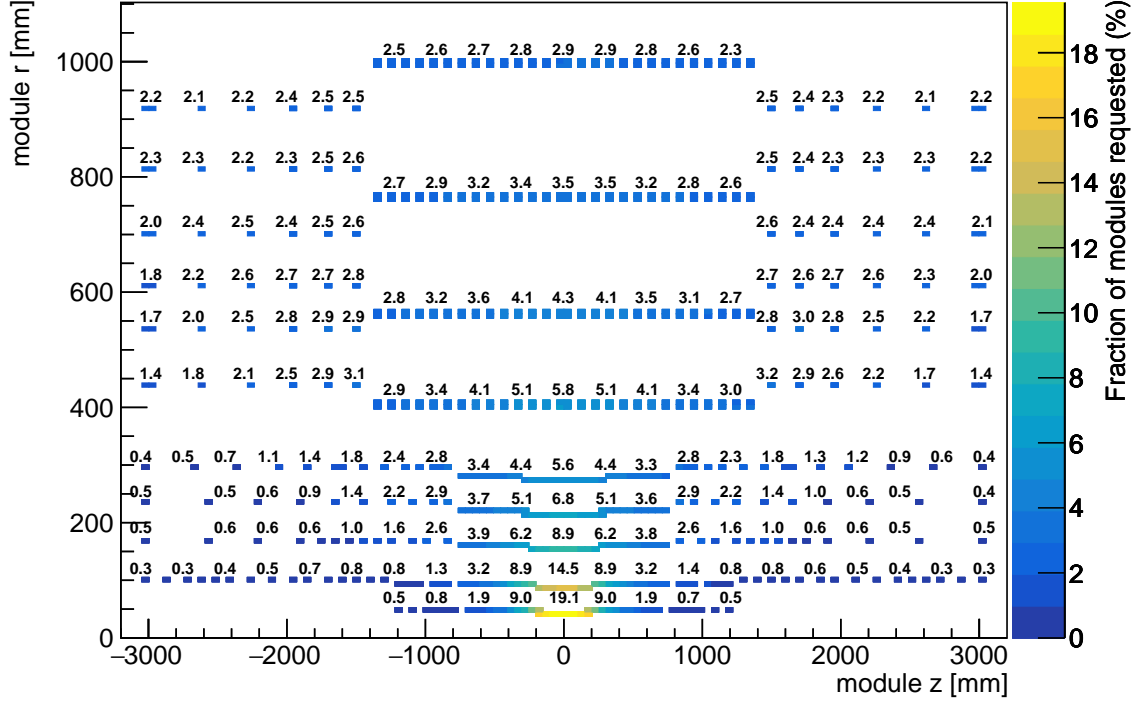


Figure 6.26: Fraction of events passing Level-0 in the representative menu for which a tracking request needs a given module. This figure is constructed by summing readout fraction distributions for individual objects, such as those shown in Figures 6.24 and 6.25, multiplied by the object multiplicity for the event type and the fraction of events for which that trigger is active.

and track-based calibration of these jets (an estimate not accounting for duplicate requests of the same region 98% of the outer **ITk** pixel layer). It is therefore assumed that single jet, multi-jet, di- $\tau$ ,  $E_T^{\text{miss}}$ , and  $H_T$  triggers will all use full-detector tracking. In addition,  $e$ ,  $\mu$ ,  $\tau$ ,  $\gamma$ , and large- $R$  jet triggers will use **gHTT** as a regional tracker with a 1 GeV minimum track  $p_T$  reconstructed.

Without the **HTT**, the nominal tracking requirements would require  $\approx 10$  times more CPUs than the baseline design with **HTT** (see Section 12.4). If less tracking is available, then the thresholds for the objects requiring tracking would need to be raised to reduce the rates. For example to reduce the tracking needs for the  $b$ -tagged four-jet the Level-0 threshold would have to be raised to an effective offline threshold of 85 GeV, which would raise the limit on the search for  $HH \rightarrow 4b$  by approximately 50%.

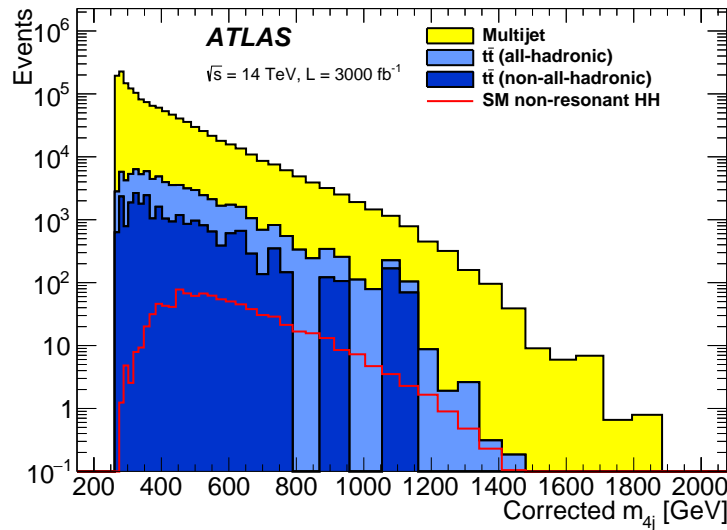


Figure 6.27: Four-jet invariant mass distribution for the  $HH \rightarrow 4b$  search.

### 6.13 Impact of the System on $HH \rightarrow 4b$

The upgrade system design has several key components which work together to provide excellent performance across a wide range of physics topics. In order to understand the interplay of the components, this section describes the  $HH \rightarrow 4b$  sensitivity predictions and how they are related to the individual sub-systems.

Measurement of the non-resonant  $HH \rightarrow 4b$  is a major goal of the HL-LHC because it places constraints on the Higgs self-coupling,  $\lambda_{HHH}$ , in the Standard Model. This parameter is related to the form of the Higgs potential, so its measurement is a direct test of the model of the spontaneous symmetry breaking mechanism in the SM. The Higgs self-coupling is also related to the order of the phase transition in the earlier universe [6.14], an important condition for the generation of the matter-antimatter asymmetry of the universe.

The projections are based on the Run 2 analysis [6.15] scaled to 14 TeV and an integrated luminosity of  $3000 \text{ fb}^{-1}$ . The analysis is based on a selection of four  $b$ -tagged jets with two pairs each having an invariant mass consistent with the Higgs boson mass. These events would be recorded by a four-jet trigger. Hadronic top quark decays are vetoed explicitly by constructing top quark candidates in the events. Additional mass-dependent requirements are made on the leading and sub-leading Higgs boson candidate  $p_T$ , and the  $\Delta\eta$  and  $\Delta R$  between the two Higgs boson candidates. Figure 6.27 shows the four-jet invariant mass,  $m_{4j}$ , of the resulting sample which is used as the final discriminant in a fit. Multi-jet QCD is the dominant background ( $\approx 90\%$ ) followed by  $t\bar{t}$  ( $\approx 10\%$ ).

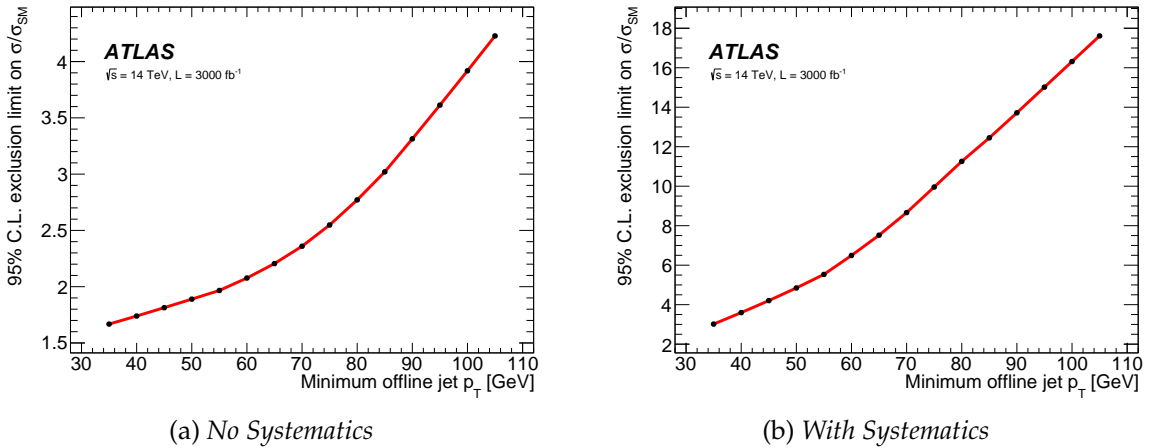


Figure 6.28: *Non-resonant  $HH \rightarrow 4b$   $\sigma/\sigma_{SM}$  95% exclusion sensitivity as a function of the minimum offline jet  $p_T$  used in the analysis.*

The backgrounds are modelled using the data-driven model from Run 2 analysis because this background is difficult to model with simulation. The impact of the improved **ITk**  $b$ -tagging efficiency is included, but the potential degradation of the performance due to pile-up and detector ageing is not. There are also many potential future improvements in analysis technique that are possible but not yet explored.

Systematics uncertainties play an important role in this analysis, but are difficult to project out to high luminosity and **HL-LHC** conditions. Two extrapolations are therefore considered: one without systematic uncertainties, and one with the Run 2 systematics uncertainties applied without any scaling. The resulting sensitivity for the non-resonant cross-section as a function of the minimum jet  $p_T$  requirement is shown in Fig. 6.28 for the two systematics scenarios. The sensitivity to  $\lambda_{HHH}$  is shown in Fig. 6.29. Since the minimum jet  $p_T$  used in the offline analysis is set by the four-jet trigger threshold, which in turn depends on the trigger architecture, these plots can be used to assess the impact on the analysis of various hardware scenarios.

All cases show substantial degradation with increased minimum jet  $p_T$  requirements. Table 6.7 shows the impact on this analysis for three scenarios with reduced upgrades: a) no **Global Trigger**, b) no **HTT**, and c) no-upgrade as previously defined to mean the Phase-I system with an output rate for 100 kHz. Without the **Global Trigger** system, the threshold for the lowest  $p_T$  jet in a four-jet trigger would have to be raised to approximately 75 GeV instead of 65 GeV to maintain the same Level-0 output rate. In addition, the loss of acceptance for near-by jets would cause an further efficiency loss of 10-15%. The impact of that reduction on the  $HH \rightarrow 4b$  analysis would be to reduce the cross-section sensitivity by  $\approx 25\%$ . Without the **HTT**, the tracking would be **CPU**-limited, and the Level-0 trigger rate would need to be reduced by a factor of  $\approx 10\times$  to allow **CPUs** in the Event Filter to do the required tracking. Such a rate reduction corresponds to a 85 GeV threshold, and an  $\approx 45\%$  loss of sensitivity. With no upgrade at all the loss is greater than  $\approx 65\%$ . A scenario with

## 6.14 System Flexibility and Dataset Quality

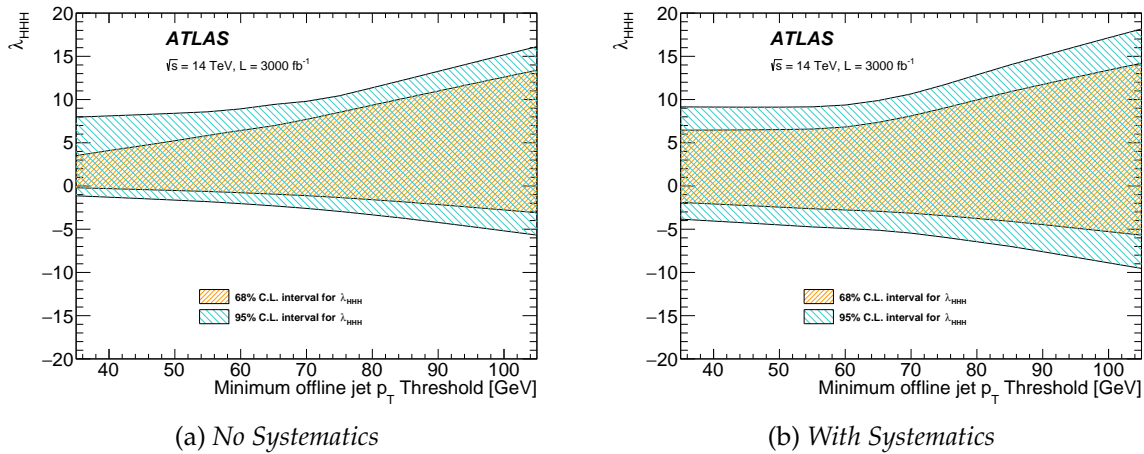


Figure 6.29: Allowed intervals for the  $\lambda_{HHH}$  parameter assuming the standard model as function of the minimum offline jet  $p_T$  used in the analysis.

Table 6.7: Estimated loss of sensitivity in the  $HH \rightarrow 4b$  analysis in various reduced upgrade scenarios

System Modification	Threshold raise from 65 GeV to ...	Cross-section Sensitivity no syst	Cross-section Sensitivity w/ syst	$\lambda_{HHH}$ Sensitivity no syst	$\lambda_{HHH}$ Sensitivity w/ syst
No <a href="#">Global Trigger</a>	75 GeV (and 10% loss of efficiency)	25%	62%	28%	26%
No <a href="#">HTT</a>	85 GeV (needed to reduce tracking CPU by $\approx 10\times$ )	47%	92%	43%	47%
No upgrade	100 GeV (Phase-I system with 100 kHz output rate)	97%	154%	79%	67%

a [DAQ](#) upgrade to 1 MHz, but no trigger upgrade would give similar performance to the no-[HTT](#) scenario, because the system would be limited by the Event Filter [CPU](#).

## 6.14 System Flexibility and Dataset Quality

The menu presented in Section 6.11 shows that the system is capable of achieving the goals set in Section 2. The [HL-LHC](#) will run for a period of approximately a decade, and these data may be the best/only sample collected for a significant period after the end of data-taking. It is therefore critical that the resulting sample be sufficiently inclusive that innovations either in experimental technique or theory can be addressed with the sample. In particular, new models often have new signatures, and precision measurements require

calibration and control datasets, the full scope of which is difficult to anticipate a decade before the final result is complete. The menu presented is therefore based on inclusive triggers that are not tuned to specific physics analysis targets, but designed to include high priority targets and provide a complete dataset for the future. The Level-0 is designed to maximise the physics entering the Event Filter, because the Event Filter being software-based can be modified near to or during data-taking and can support significantly more algorithm complexity. By providing tracking capabilities to the trigger similar to those of the offline reconstruction, the hardware tracking provides the computing resources to allow for future optimisations without being computing limited. This flexibility in the Event Filter also provides the capability to reduce the Event Filter output rate if needed by moving analysis-specific selections online for known analysis targets at the expense of reducing the richness and long-term value of the dataset.

## References

- [6.1] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, [JINST 3 \(2008\) S08003](#).
- [6.2] GEANT4 Collaboration, S. Agostinelli et al., *GEANT4: A simulation toolkit*, [Nucl. Instrum. Meth. A 506 \(2003\) 250](#).
- [6.3] T. Sjostrand, S. Mrenna, and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, [Comput. Phys. Commun. 178 \(2008\) 852](#), [arXiv:0710.3820 \[hep-ph\]](#).
- [6.4] ATLAS Collaboration, G. Aad et al., *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, [Eur. Phys. J. C77 \(2017\) 490](#), [arXiv:1603.02934 \[hep-ex\]](#).
- [6.5] ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton-proton collision data*, ATLAS-CONF-2014-032, 2014.
- [6.6] ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data*, ATLAS-CONF-2016-024, 2016.
- [6.7] M. Cacciari, G. P. Salam, and G. Soyez, *The Anti-k(t) jet clustering algorithm*, [JHEP 04 \(2008\) 063](#), [arXiv:0802.1189 \[hep-ph\]](#).
- [6.8] ALEPH Collaboration, D. Buskulic et al., *Performance of the ALEPH detector at LEP*, [Nucl. Instrum. Meth. A360 \(1995\) 481–506](#).
- [6.9] ATLAS Collaboration, *Jet reconstruction and performance using particle flow with the ATLAS Detector*, [Eur. Phys. J. C77 \(2017\) no. 7, 466](#), [arXiv:1703.10485 \[hep-ex\]](#).
- [6.10] D. Krohn, J. Thaler, and L.-T. Wang, *Jet Trimming*, [JHEP 02 \(2010\) 084](#), [arXiv:0912.1342 \[hep-ph\]](#).

## 6.14 System Flexibility and Dataset Quality

---

- [6.11] LHC Higgs Cross Section Working Group Collaboration, D. de Florian et al., *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, [arXiv:1610.07922](https://arxiv.org/abs/1610.07922) [hep-ph].
- [6.12] ATLAS Collaboration, *Technical Design Report for the ATLAS Inner Tracker Strip Detector*, Tech. Rep. CERN-LHCC-2017-005, CERN, Geneva, Apr, 2017. <https://cds.cern.ch/record/2257755>.
- [6.13] ATLAS Collaboration, *Technical Design Report for the ATLAS Inner Tracker Pixel Detector*, Tech. Rep. CERN-LHCC-2017-021, CERN, Geneva, (in preparation).
- [6.14] A. Noble and M. Perelstein, *Higgs self-coupling as a probe of electroweak phase transition*, *Phys. Rev.* **D78** (2008) 063518, [arXiv:0711.3018](https://arxiv.org/abs/0711.3018) [hep-ph].
- [6.15] ATLAS Collaboration Collaboration, *Search for pair production of Higgs bosons in the  $b\bar{b}b\bar{b}$  final state using proton–proton collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector*, Tech. Rep. ATLAS-CONF-2016-049, CERN, Geneva, Aug, 2016. <https://cds.cern.ch/record/2206131>.



## **Part II**

### **Detailed Description of System Components**



## 7 Level-0 Calorimeter Trigger

The Level-0 Trigger consists of the Level-0 Calorimeter Trigger, the Level-0 Muon Trigger, the [Global Trigger](#) and the [CTP](#). Whereas the Calorimeter- and Muon-Trigger identify and determine the properties of [TOBs](#), the [Global Trigger](#) refines the information using higher granularity calorimeter information and combines the information from the calorimeters and the muon system. It applies topological algorithms and sends the information to the [CTP](#) which issues the [LOA](#) signal. This chapter describes the Level-0 Calorimeter Trigger.

The basis of the Level-0 Calorimeter Trigger consists of the Level-1 Calorimeter Trigger which will be installed during the LHC Long Shutdown 2 and operated during Run 3 [7.1]. It is designed with forward compatibility to be operated during Run 4. Various hardware, firmware and software modifications are needed to accommodate the changes of the overall [TDAQ](#) system and the new calorimeter signal paths. We expect to implement improvements to the feature extraction algorithms based on the experience of operation during Run 3. The performance of the system will be enhanced by an additional processing system, the [fFEX](#), which targets at the precise reconstruction of forward electrons and jets. In this chapter the [LOCalo](#) system is described with special focus on the technical changes to the existing system and potential improvements.

### 7.1 Evolution of the Hardware Calorimeter Trigger

#### 7.1.1 Enhancements and Interface changes of the Phase-I Level-1 Calorimeter Trigger

The Level-1 Calorimeter Trigger that is currently being built within the Phase-I upgrade is constrained by various requirements. These constraints become less stringent at Phase-II, due to changes of the overall detector and readout-electronics.

- The latency budget will be significantly relaxed. This will allow the implementation of more complex algorithms with a potentially longer execution time.
- The signals from the Tile Calorimeter are currently provided as analogue sums, so-called trigger towers. The [L1Calo](#) PreProcessor system which will digitise the [Tile](#) analogue trigger signal as input to the [FEXs](#) will become obsolete after the Phase-II upgrade. The processing will be performed within the Tile backend electronics similar to the [LAr](#) calorimeter.

## 7.2 Performance of the Level-0 Calorimeter Trigger

---

- The [LAr](#) Calorimeter will provide signals, in particular cell information from the forward calorimeter [FCal](#), in addition to the already provided [super cell](#) and Trigger Tower information. This will enable the implementation of better performing algorithms.
- The output signals will be sent to the [Global Trigger](#) as well as to the legacy [L1Topo](#) System, in the last case for commissioning purposes.

### 7.1.2 Requirements for the Level-0 Calorimeter Trigger

The core of the Level-0 Calorimeter Trigger will be the Level-1 Calorimeter Trigger of the Phase-I upgrade. The proposed upgrade of the calorimeter electronics and the [TDAQ](#) system leads to a number of new requirements which the Level-0 Calorimeter Trigger needs to follow. These requirements concern in particular the interfaces to upstream and downstream systems as well as the implementation of the new [fFEX](#) processing system.

- The Level-0 Calorimeter Trigger needs to implement the remapping of the digital input signals from the Tile Calorimeter.
- Inputs from the full granularity cell signal electronics for  $|\eta| > 2.5$  need to be received.
- The processing of the fine granularity input signals in the forward region needs to be implemented.
- The acceptance to trigger on electrons and photons should be extended beyond the region which will not be covered by the Phase-I system ( $2.5 < |\eta| < 4.9$ ).
- The triggering on jets for  $|\eta| > 3.2$  should be improved using the high granularity inputs.
- The firmware that formats the output signals to the [Global Trigger](#) and to [L1Topo](#) needs to be revised.
- The arrival time differences between signals from different partitions need to be compensated.

## 7.2 Performance of the Level-0 Calorimeter Trigger

The Level-1 Calorimeter Trigger of the Phase-I system is designed such that it copes with the conditions of the LHC Run 3. The conditions during Run 4 will be harsher, e.g. the pile-up will increase to about  $\mu = 200$ . This will naturally increase the trigger rates and worsen the purity of the object identification by the trigger processors. In order to counteract this effect, the Level-0 system will be enhanced by the [Global Trigger](#). The performance of the combined system ([FEXs](#) and [Global Trigger](#)) will be such that the trigger thresholds can be chosen reasonably low to cover the whole physics programme. The performance of the combined system is discussed in Section 6. Here we discuss the performance of the Level-0 Calorimeter Trigger object finding.

---

As discussed in Section 6.3, the electron selection of the eFEX will be enhanced by additional cluster shape variables which are determined in the Global Trigger. The performance of the combined system (Level-0 Calo and Global Trigger) allows to set the single electron threshold at about 20 GeV. Figure 7.1 a) displays the rates for triggering on electrons with the Level-0 system. The different curves show the effect of applying veto condition successively. The variables  $R_\eta$  and  $R_{\text{HAD}}$  will be used in the eFEX algorithm. The variable  $E_{\text{ratio}}$  will be used in the Global Trigger algorithm. A substantial rate reduction from the initial selection which is based on finding energy maxima in the super cell grid, can be observed. Figure 7.2 shows the efficiencies for reconstructed taus as function of the offline transverse momentum of the tau candidates. Sharp turn-ons can be observed over a large range. The 90 % efficiency is reached between 40 GeV and 60 GeV depending on the exact threshold value. The performance is sufficient to support the physics programme as discussed in Section 6.5.

Single jets are well triggered by the jFEX system (see Figures 6.10 a) and 6.10 b)). Multijets are handled by the Global Trigger. Jets in the forward direction suffer from harsher pile-up conditions compared to central jets. The fFEX is designed to process the cell level calorimeter input for  $|\eta| > 2.5$ . It will improve the current jet reconstruction for high  $\eta$  which is based on trigger towers. Figure 7.1 b) shows the rate for forward jets using different reconstruction algorithms. The standard trigger algorithm leads to a rate which is about an order of magnitude higher compared to the offline algorithm. The usage of an anti- $k_t$  algorithm does not improve the situation. This leads to the conclusion that the amount of information is not sufficient to decrease the trigger rate. The cell information which is available for fFEX can be employed to reduce the rates below the values given by the Phase-I system.

The gFEX system will continue to trigger on boosted topologies by identifying large-radius jets that are characteristic of Lorentz-boosted objects. The jet-finding algorithm in the gFEX consists of building  $\eta \times \phi = 0.6 \times 0.6$  fully overlapping regions as described in Section 3.2.1. The performance of the gFEX at  $\langle mu \rangle = 200$  is shown in Fig. 7.3.

## 7.3 Architecture and Hardware Realisation

The core components of the Run 3 hardware described as the Level-1 Calorimeter Trigger in the Phase-I TDR [7.1] will stay untouched. Nevertheless various modifications, especially of the interfaces with upstream and downstream systems, are necessary.

### 7.3.1 Overview

Figure 7.4 shows a functional diagram of the Level-0 Calorimeter Trigger. The L0Calo system receives digital input signals from both calorimeter systems (LAr and Tile) through optical fibres. The signals are routed through the Fibre Optic eXchange (FOX) (described

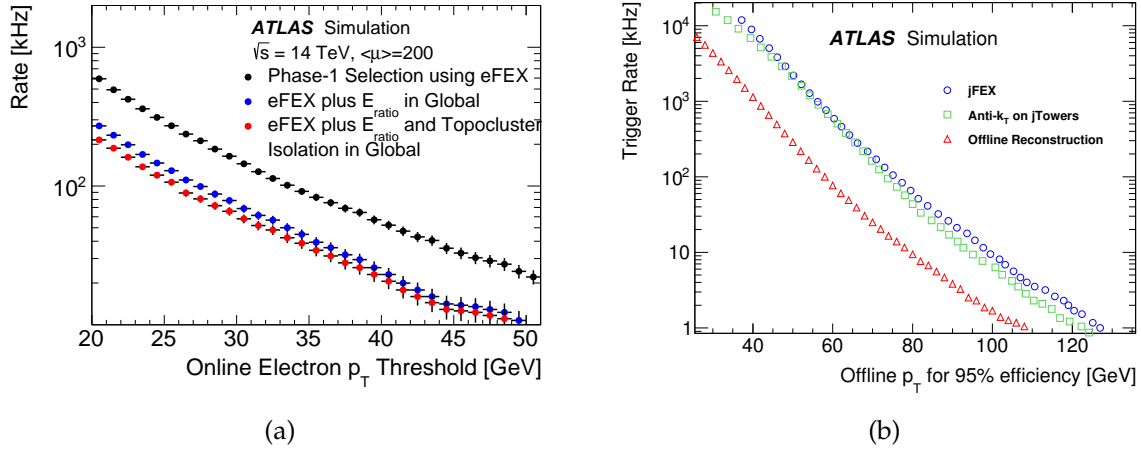


Figure 7.1: a) Level-0 trigger rates for electrons. The different curves are for the successive application of veto conditions. b) The forward ( $|\eta| > 3.2$ ) single-jet trigger rate vs. offline  $p_T$  thresholds for jets reconstructed in the *jFEX*. The efficiency is evaluated using  $HH \rightarrow bbbb$  signal events, and the trigger rate is evaluated based on minimum bias background events at  $\langle \mu \rangle \simeq 200$ . The *jFEX* algorithm, the offline anti- $k_t$  algorithm (run over  $\eta \times \phi = 0.1 \times 0.1$  towers), and the full offline reconstruction are compared.

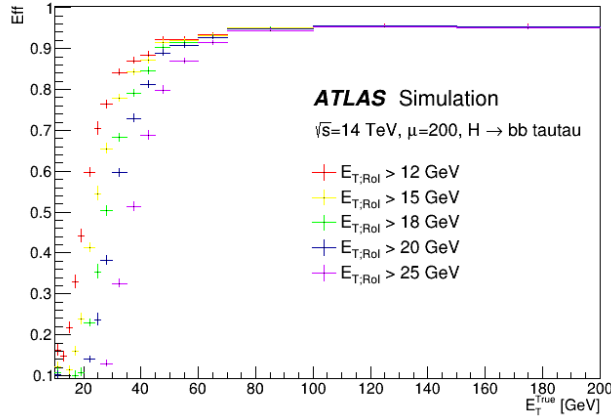


Figure 7.2: Efficiency to trigger on taus for different trigger thresholds as function of the true transverse momentum of the tau.

below) to the Feature Extractors (FEXs) which constitute the core of the Level-0 Calorimeter Trigger. Three different FEX systems, which are currently being built for the Phase-I upgrade, will operate complex algorithms to identify electrons, photons and taus (eFEX), jets, large area taus, missing and total energy (jFEX) as well as large area jets and alternat-

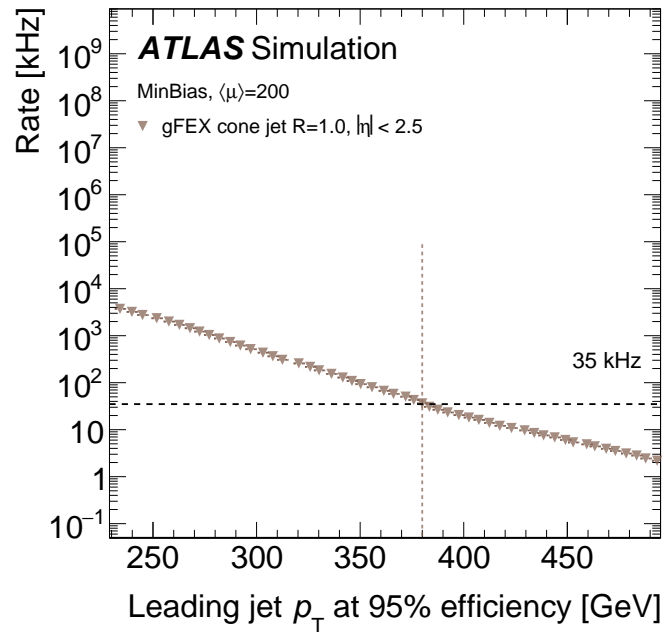


Figure 7.3: The expected performance of the [gFEX](#) at  $\langle \mu \rangle = 200$ . The trigger rate vs. the leading  $R = 1.0$  jet  $p_T$  at 95% efficiency is shown. The [gFEX](#) “cone jet” algorithm is used.

ive missing transverse energy algorithms ([gFEX](#)). During Run 3 the objects are sent to the [L1Topo](#), which counts the number of trigger objects above configurable thresholds and performs topological algorithms like e.g. the calculation of invariant masses. With the Phase-II upgrade, these tasks are taken over by the [Global Trigger](#) (described in Section 9). The trigger objects will be sent optically to the [Global Trigger](#) which can use the full calorimeter granularity to run more sophisticated algorithms. The information sent to the [CTP](#) consists of multiplicities of objects (e.g. electrons above a  $p_T$  threshold) and flags indicating if topological requirements, etc., have been satisfied.

In order to enhance the capabilities of the system, an additional [FEX](#) system will be built, the forward FEX ([fFEX](#)). It will operate on the full granularity [FCal](#) and [HEC](#) data as well as the cell information of the inner part of the [EMEC](#) ( $|\eta| > 2.5$ ), which enables the reconstruction of forward electrons and forward jets. The hardware implementation will be based on the [jFEX](#) modules, which are currently being built within the Phase-I upgrade. At the beginning of Run 4 the [L1Topo](#) will be operated in parallel to the [Global Trigger](#). Once the [Global Trigger](#) is fully commissioned, [L1Topo](#) will be decommissioned.

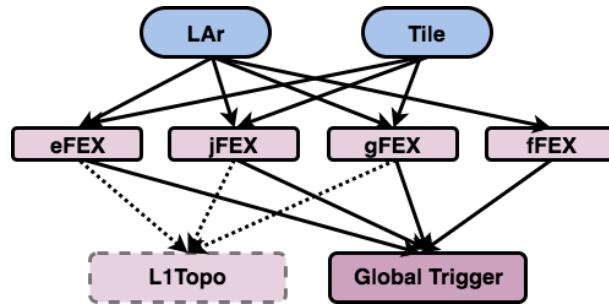


Figure 7.4: Functional diagram of the Level-0 Calorimeter Trigger: The input signals from calorimeters (LAr and Tile) are sent to the processing systems. Feature extractors eFEX, jFEX and gFEX will be built as part of the Phase-I upgrade programme. fFEX will be a new component. The identified trigger objects are sent to the Global Trigger as well as to L1Topo. The L1Topo will be operated during the commissioning phase. Once the Global Trigger is fully operational, L1Topo will be decommissioned. The different components of the detector backend electronic systems as well as the Fibre-Optic eXchange Plant (FOX) are not shown explicitly in this diagram.

### 7.3.2 Input Signals

**Requirements** Figure 7.5 shows the connectivity between the different partitions of the calorimeter and the FEX processing systems. The input signals are sent as transverse energies per bunch crossing partially corrected for pile-up backgrounds. The preprocessing of the digitised signals take place in the calorimeter backend systems. Special algorithms are employed to determine the bunch crossing where the interaction of an energy deposition took place. The signal to noise ratio is improved using digital filter algorithms which correct also for in- and out-of-time pile-up.

**LAr Calorimeter Signals** The Level-0 Calorimeter Trigger receives digital input data from the backend electronics of the calorimeter systems. The calorimeter cells are summed up to super cells in the FE electronics. They consist of up to 32 detector cells. The information is sent to the LDPS and from there to the FEXs. The shape and the size of the super cells depends on the LAr layer and they are optimised for algorithms which aim to identify electromagnetic energy deposits. These data are sent to the eFEX system. The LDPS also provides coarser granularity data for the jFEX and the gFEX systems. For the jFEX system energy sums are computed over the full longitudinal depths and areas which extend over  $\Delta\eta \times \Delta\phi$  of  $0.1 \times 0.1$  whereas these sums for the gFEX extend over  $0.2 \times 0.2$ . The LDPS is part of the LAr Phase-I upgrade [7.2].

The Phase-II upgrade of the LAr electronics consists of the complete replacement of the cell readout chain from the frontend to the backend (the Phase-I upgrade trigger path will be kept) [7.3]. The 180k LAr cells will be digitised at 40 MHz on detector and signals will be available on the backend system (LASP). It provides detailed strip layer information to



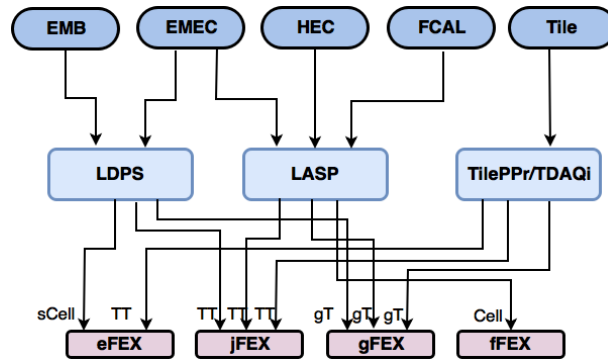


Figure 7.5: The connectivity of the calorimeter partitions with the *FEX* systems through the different backend systems is shown. *LDPS* is the LAr Digital Processing System. The labels at the input to the different *FEX* systems describe the granularity of the input data. The acronym *TT* refers to trigger towers which cover an area in  $\Delta\eta \times \Delta\phi$  of  $0.1 \times 0.1$  for  $|\eta| < 2.5$ . Beyond this value the granularity increases 5.1; *gT* refers to the inputs to the *gFEX* and cover an area of  $0.2 \times 0.2$ ; *sCell* are super cells, and *Cell* means cell-level information.

the [Global Trigger](#). The *LASP* system will also provide data for the Level-0 Calorimeter Trigger. It sends the signals from the hadronic endcap part (*HEC*), the inner part of the electromagnetic endcap (*EMEC*) and the forward calorimeter (*FCal*). The change of the data source for the legacy *FEX* systems (*eFEX*, *jFEX* and *gFEX*) will be transparent. The signals will be sent with the same granularity and data content as will be done for the Phase-I system, i.e. no hardware or firmware modification will be needed. The new feature extractor will receive cell level information for  $|\eta| > 2.5$  including data from *EMEC*, *HEC* and *FCal*. This system will enhance the capabilities to trigger on forward electrons and jets using the high granularity inputs. The new feature extractor makes modifications of the *FOX* system necessary.

**Tile Calorimeter Signals** The signals from the Tile Calorimeter during Run 3 will be transmitted electrically and digitised and pre-processed within the Level-1 Calorimeter Trigger. For the Phase-II upgrade a complete redesign of the signal chain will be implemented. The Tile signal digitisation takes place within the Tile frontend electronics followed by optical transmission to the (*TPPr*) [7.4] where the cell energies are determined. The signals for the Level-0 Calorimeter Trigger are prepared within the *TDAQ* interface (*TDAQi*) and transmitted to the *FEX* systems. The *TDAQi* system consists of electronic boards which are connected at the rear of the shelves filled with *TPPr* modules. There are in total 32 *TDAQi* boards being built. Since the individual modules of the Level-1 Calorimeter PreProcessor (Phase-I) and the Tile PreProcessor (Phase-II) cover different calorimeter areas a redesign of the Fibre Optic eXchange (*FOX*) responsible for the distribution of the Tile signals is necessary.

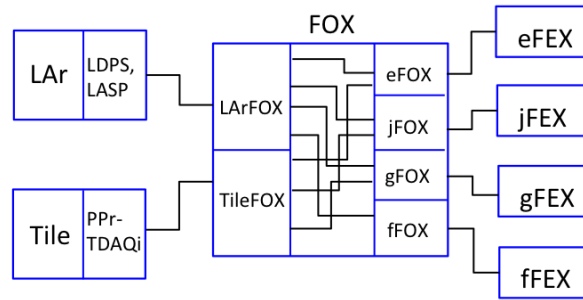


Figure 7.6: This diagram shows the connectivity between the calorimeter backend electronics and the processing system of the Level-0 Calorimeter Trigger consisting of four FEX systems. The signals are processed within the LDPS and LASP modules for the LAr Calorimeter and the PPr-TDAQi modules for the Tile Calorimeter. The signals for the trigger processors are sent digitally on optical fibres. A Fibre Optic eXchange system (FOX) redistributes the signals to the according FEX processing system. The FOX consists of an input stage a remapping stage and an output stage.

**The Fibre Optic eXchange (FOX)** The FOX (shown in Fig. 7.6) performs a remapping of the input signals between the calorimeter electronics and the FEX system. The signals are received in fibre bundles of 12 fibres and transmitted in bundles of 48 or 72 fibres. The receiving part of the FOX is partitioned in a LArFOX and a TileFOX. The output section is partitioned into 4 parts according to the four FEX systems. In between the signals are redistributed with custom made distribution ribbons. Figure 7.7 shows the detailed interconnectivity for one of the FOX boxes. The Phase-II upgrade necessitates a redesign of the TileFOX and the LArFOX as well as the connected distribution ribbons which connect the input to the corresponding output FOX units. A special fFOX needs to be built as well. The whole system consists of passive components. The requirements concerning the signal integrity do not change, since the data rates are the same as before. The new components, the TDAQi and the LASP transmit data with the same rate as the legacy systems.

Table 7.1 lists the number of fibres which connect the backend electronics of the various detector regions and the FEX systems. Each fibre carries the signals of a fixed number of channels. The signals have a resolution of 10 bits for the LAr signals and 10 bits for the Tile signals. The optical transmission operates at a data rate of 11.2 Gb/s. Since the data transfer is synchronous, no algorithms for data reduction are employed. The amount of transmitted data is always the same for each event.

### 7.3.3 Processing System

The processing system consists of four FEX systems three of which are being built within the Phase-I upgrade. The three systems consist of a different number of processing modules: eFEX (24), jFEX (6) and gFEX (1). The hardware will stay untouched. All the necessary

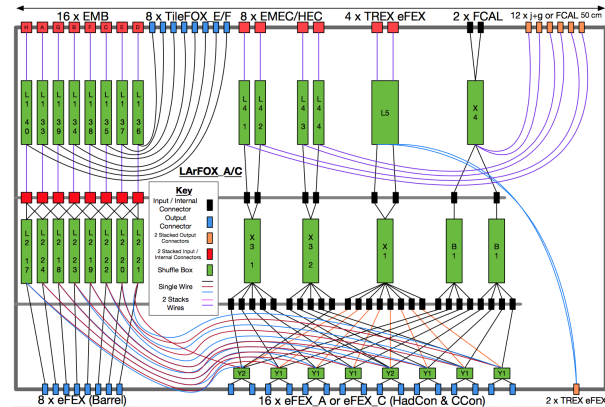


Figure 7.7: Diagram of the interconnectivity within an example FOX box.

Table 7.1: Number of fibres connecting the backend electronics of the respective detector regions and the FEX systems as well as the type of information which is transmitted. TT corresponds to the trigger towers which represent transverse energy sums over cells in areas of  $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$  and in full depths. gT refers to gTowers, which represent sums in  $\Delta\eta \times \Delta\phi = 0.2 \times 0.2$  and in full depths. sCell corresponds to super cells and Cell corresponds to the individual detector cells.

	EMB		EMB/EMEC		EMEC		EMEC/HEC		FCal		Tile	
	fibres	type	fibres	type	fibres	type	fibres	type	fibres	type	fibres	type
eFEX	800	sCell	800	sCell	640	sCell	240	sCell	64	sCell	384	TT
jFEX	192	TT	192	TT	160	TT	336	TT	64	TT	384	TT
gFEX	32	gT	32	gT	32	gT	64	gT	18	gT	32	TT
fFEX	-	-	-	-	-	-	320	Cell	162	Cell	-	-

requirements to operate the system after Phase-II is taken into account during their design at Phase-I. The additional FEX, the fFEX will consist of two modules. It is described below.

The modules are characterised by a large number of input links, several high performance FPGAs for processing and a smaller number of output links sending the processing results to L1Topo during Run 3 and to the Global Trigger after the Phase-II upgrade. The main characteristics of the modules are summarised in Table 7.2.

The infrastructure firmware which handles the input and output signals on the FEX processor boards needs to be changed to accommodate the different routing paths of the input signals and to change the output to the Global Trigger instead of L1Topo.

### 7.3.4 The forward FEX System (fFEX)

The forward FEX system aims at the reconstruction of electron/photon and jet candidates in the forward direction ( $|\eta| > 2.5$  for electrons and  $|\eta| > 3.2$  for jets). It profits from the

Table 7.2: Main characteristics of the *FEX* systems.

	modules	shelves	power per module	coverage	available input links per module	processing FPGAs per module	available output links per module
eFEX	24	2	400 W	$ \eta  < 2.5$	144	4	48
jFEX	6	1	400 W	$ \eta  < 4.9$	240	4	48
gFEX	1	1	300 W	$ \eta  < 4.9$	312	4	108
fFEX	2	1	400 W	$2.5 <  \eta  < 4.9$	252	4	48

changes in the *LAr* readout electronics, which allow for the transmission of the individual cell information rather than trigger towers or *super cells*. This allows the application of sophisticated algorithms and a better rate control in this region where the background from pile-up is particularly large.

In order to implement this functionality a processing system consisting of two modules is foreseen. Each board covers one detector side ( $|\eta| > 2.5$ ) and receives 241 fibres that carry the cell signals of the *FCal*, the inner part of the *EMEC* and the *HEC* part of the *LAr* calorimeter [7.3]. The implementation is based on the *jFEX* module, which is currently being designed in the context of the Phase-I upgrade.

### Requirements of the *fFEX*

- The modules need to connect to 241 input fibres.
- The input data need to be received and distributed to the processing FPGAs.
- Algorithms need to be implemented which identify electron/photon and jet candidates
- The processing output needs to be provided to the *Global Trigger*.

**Technical Implementation of the *fFEX*** The technical implementation of the *fFEX* is based on the *jFEX* modules which are currently being built in the framework of the Phase-I upgrade project. Figure 7.8 shows the final prototype of the *jFEX* module. The *fFEX* modules will be implemented using the ATCA standard. They are equipped with optical receivers and transmitters to connect to the upstream and downstream components (calorimeters, FELIX, TTC and DCS). The processing will be performed using four high speed grade FPGAs which provide the appropriate input resources to receive the calorimeter signals. In order to share information between FPGAs, inter FPGA communication will be implemented through high speed links. The output data will be formatted and transmitted through optical links to the *Global Trigger*. The configuration and control will be implemented and real time monitoring data will be sent to *DCS* using the *FELIX* interface.

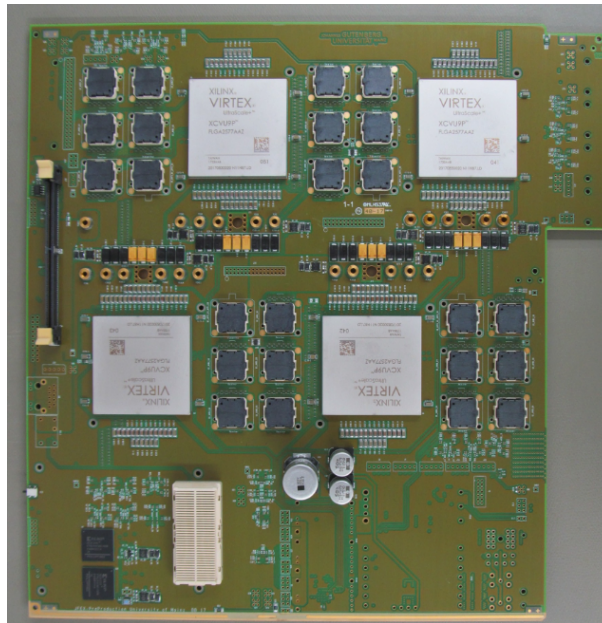


Figure 7.8: Photograph of the final prototype of a *jFEX* blade. The board is equipped with four high speed grade FPGAs and the minipod connectors for the optical input and output. Missing on this module are various infrastructure components for control and power management of the board.

### 7.3.5 Output Signals

During Run 3 the *FEX* systems send the information of the identified Trigger Objects (*TOBs*) to *L1Topo*, which counts the number of *TOBs* above various configurable thresholds and performs topological algorithms like the determination of the invariant mass of pairs of objects. The information is then sent to the central trigger processor (*CTP*), which determines the final trigger decision and issues the Level-1 accept signal leading to the readout of the event.

The task of *L1Topo* is taken over by the *Global Trigger* after the Phase-II upgrade. For this purpose the output fibres are connected to the *Global Trigger* hardware and the *TOB* information is sent to the *Global Trigger*. The *TOBs* will be sent directly from each of the *FPGAs* to the *Global Trigger*. The *TOBs* contain the information about the type of the objects, their transverse momentum and their position.

In order to manage the transition during the commissioning phase the signals are still sent to *L1Topo*, which remains in operation until the *Global Trigger* is fully commissioned. A duplication of the signal output is therefore needed. For the *gFEX* system this can be performed at module level. There exists enough spare output capacity. The number of output fibres from the *eFEX* and *jFEX* system are limited, making the splitting of the optical output signals necessary.

Table 7.3: *Number of output links and bits per TOB from each FEX system to the Global Trigger system.*

	links to Global per FEX module	bits per TOB (Global)
eFEX	48 (after splitting)	64
jFEX	48 (after splitting)	32
gFEX	32	64
fFEX	48	32

Apart from the changes of the routing of the optical fibres also firmware changes are foreseen to allow to send more data from the FEX systems to the Global Trigger. This allows sending more detailed information about the Trigger Objects. The amount of output for the Phase-I system is limited by the L1Topo design which allows a certain amount of input data only. The Global Trigger will be able to accept more information. The content of the TOBs will be enhanced by detailed information about the variables which are used for identification rather than whether a certain condition has been fulfilled. This additional information can then be used by the Global Trigger as input for e.g. a BDT. Table 7.3 summarises the number of output links to the Global Trigger along with the number of bits available for each TOB.

#### 7.3.6 Latency

The input signals from the calorimeters arrive  $1.1 \mu\text{s}$  (LAr) and  $1.45 \mu\text{s}$  (Tile) after the time of the collision at the input to the FEX processors. The available processing time is 500 ns. This can be compared with the Phase-I processing time of 125 ns. The quadrupled available processing time therefore allows for significantly more complex algorithms to be deployed. The time needed to submit signals from the output of the FEXs to the inputs of the Global Trigger is 175 ns. Further details can be found in Section 5.2.8.

#### 7.3.7 Readout and Monitoring

The main purpose of the readout of the Level-0 trigger data is the possibility to perform a detailed monitoring of the processing algorithms. This helps to identify possible hardware malfunctions, problems with the algorithms or potential problems with the input data (e.g. noise bursts). The readout data consist of the information of the TOBs. The calorimeter information will be read out by the calorimeter systems such that the algorithms can be emulated and the trigger decision simulated and compared to the actual hardware result.

The readout proceeds to the data acquisition system via FELIX after receiving a L0A signal. The system is equipped with several components that support the core part of the system. Each shelf, except the gFEX shelf, is equipped with two HUB modules that carry ROD modules to perform the event readout. The connectivity to the processing modules

---

is implemented through backplane connections. The [gFEX](#) module is directly read out via [FELIX](#). The hardware and firmware stays untouched and will be used as it is used during Phase-I.

## 7.4 Firmware

### 7.4.1 Algorithmic Firmware

The general goal is to keep the firmware stable over long periods and we do not expect to update the algorithms except for bug fixing or in the case of a very substantial performance improvement. The Phase-II upgrade opens the possibility to upgrade also the algorithmic firmware based on the experience from Run 3. The relaxed latency budget allows to increase the available processing time from 125 ns to 500 ns (see Section [5.2.8](#)).

### 7.4.2 Infrastructure Firmware

There is no major upgrade of the infrastructure firmware foreseen except for small adaptations due to the change of the interfaces to [FELIX](#) and the [DCS](#). Depending on the mapping details of the input signals, minor changes of the input firmware of the [FEX](#) systems might be necessary. After the Phase-II upgrade also the bandwidth limits to the subsequent system ([Global Trigger](#)) are not anymore as stringent as in Run 3. This opens the possibility to send more information about the [TOBs](#). This also necessitates a change of the output data formats.

## 7.5 R & D Programme

A large part of the Level-0 Calorimeter Trigger will consist of legacy components which are constructed in the framework of the Phase-I upgrade. The following components need to be modified or newly developed.

- Hardware
  - The [fFEX](#) module needs to be designed and constructed.
  - The [FOX](#) needs to be modified.
  - The [eFEX](#) and [jFEX](#) output signals to [L1Topo](#) need to be split.
- Firmware
  - The algorithms for the [fFEX](#) need to be developed.
  - Possible improvements to the algorithms of the legacy components need to be updated.



- Anticipated modifications to the readout FW need to be implemented.
  - Anticipated modification to the DCS FW need to be implemented.
- Software  
The configuration and control software which supports the different components needs to be updated to account for the hardware and firmware changes.

## 7.6 Commissioning

The Level-0 Calorimeter Trigger (except the **fFEX**) should be fully operational at the beginning of Phase-II since it consists of the Phase-I legacy system. The functional commissioning of the new interfaces to the calorimeters (**LASP** and **TDAQi**) as well as to **FELIX** and **DCS** can be finalised well in advance of data taking, since the signals are all digital. The calibration of the calorimeter inputs can be performed using well established procedures using test pulses. The **fFEX** needs time for commissioning although a large part of its functionality can be tested before first collisions since it is a fully digital system. In order to support a managed transition of the tasks from **L1Topo** to **Global Trigger** we plan to operate **L1Topo** during the initial phase of Run 4. Once the **Global Trigger** is fully functional and has taken over all the **L1Topo** tasks in a reliably manner we will decommission **L1Topo**.

## References

- [7.1] ATLAS Collaboration, *Technical Design Report for the Phase-I Upgrade of the ATLAS TDAQ System*, Tech. Rep. CERN-LHCC-2013-018. ATLAS-TDR-023, CERN, Geneva, Sep, 2013. <https://cds.cern.ch/record/1602235>.
- [7.2] ATLAS Liquid Argon Calorimeter Group, *ATLAS Liquid Argon Calorimeter Technical Design Report: Trigger Electronics for the LHC Phase-I Upgrade*, Tech. Rep. CERN-LHCC-2013-017, CERN, Geneva, (in preparation). <https://twiki.cern.ch/twiki/bin/view/LAr/LArPhaseITDR>.
- [7.3] ATLAS Collaboration, *Technical Design Report for the Phase-II Upgrade of the ATLAS LAr Calorimeter*, Tech. Rep. CERN-LHCC-2017-018. ATLAS-TDR-027, CERN, Geneva, Sep, 2017. <https://cds.cern.ch/record/2285582>.
- [7.4] ATLAS Collaboration, *Technical Design Report for the Phase-II Upgrade of the ATLAS Tile Calorimeter*, Tech. Rep. CERN-LHCC-2017-019; ATLAS-TDR-028, CERN, Geneva, Sep, 2017. <http://cds.cern.ch/record/2285583>.



# 8 Level-0 Muon Trigger

## 8.1 Introduction

The current Level-1 Muon Trigger System will be upgraded to form the Level-0 Muon Trigger System for HL-LHC. The Level-0 muon trigger will be based on the data of the upgraded muon spectrometer and the Tile calorimeter. The muon spectrometer upgrade is described in another TDR [8.1]. Figure 8.1 shows a cross-sectional view of the upgraded muon spectrometer. It consists of RPC ( $0 < |\eta| < 1.05$ ), TGC ( $1.05 < |\eta| < 2.4$ ), NSW ( $1.3 < |\eta| < 2.7$ ), and sMDT/MDT ( $|\eta| < 2.7$ ). The NSW consists of sTGC and MM [8.2]. The sMDT is implied by the word MDT throughout the chapter. The detectors constitute three groups ('stations') in both the barrel and the endcap: inner, middle, and outer stations in  $|\eta| < 1.05$  and  $1.3 < |\eta| < 2.7$  and inner, extra, and middle stations in  $1.05 < |\eta| < 1.3$ .

This chapter describes the design of the Level-0 Muon Trigger System, including the hardware and the trigger logic. Section 8.2 shows an overview of the current Level-1 Muon Trigger System and the Phase-I upgrade as well as the limitations. Section 8.3 gives an overview of the Phase-II upgrade. Sections 8.4, 8.5, and 8.6 describe the components of the Level-0 Muon Trigger System, the Sector Logic, the NSW Trigger Processor, and the MDT Trigger Processor, respectively. Section 8.7 summarises the latency estimates. Section 8.8 provides the R&D items.

## 8.2 Overview of the Current System and the Limitations

The original Level-1 muon trigger at the ATLAS experiment is based on the hit data from RPC in the barrel region ( $|\eta| < 1.05$ ) and TGC in the endcap region ( $1.05 < |\eta| < 2.4$ ) [8.3]. The muon track candidates are identified by simple coincidence logic in the on-detector boards. The transverse momentum is evaluated by look-up tables in the off-detector boards. At the beginning of Run 2, the signals from TGCs in the endcap inner station ( $1.05 < |\eta| < 1.9$ ) were introduced in the coincidence to reduce the rate of triggers caused by the particles not directly from the interaction point ('fake triggers'). The coverage of the TGCs in the endcap inner station is about half in  $1.05 < |\eta| < 1.3$ , and thus the energy deposits in the outermost cells of the Tile calorimeter are being incorporated during Run 2 to reduce the fake triggers [8.4]. The detectors in the endcap inner station in  $1.3 < |\eta| < 2.7$  will be

## 8.2 Overview of the Current System and the Limitations

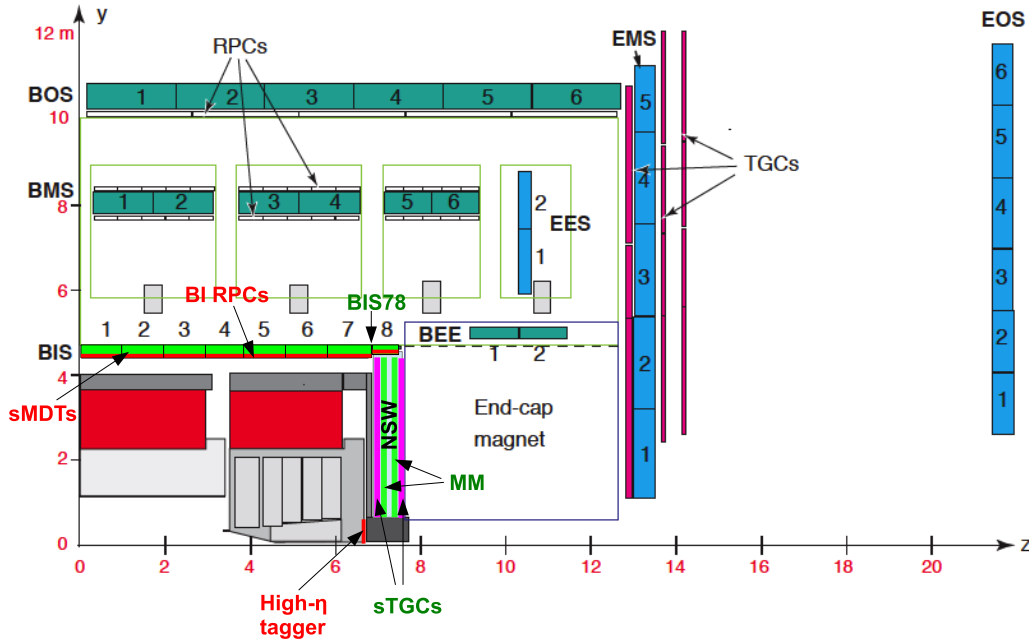


Figure 8.1: Cross-sectional view of the Phase-II ATLAS muon spectrometer layout, showing a so-called small sector, one of the azimuthal sectors that contain the barrel toroid coils. The drawing shows the new detectors to be added in the Phase-II upgrade (red text: BI RPC, sMDT, high- $\eta$  tagger), those to be installed during Long Shutdown 2 (green text: Micromegas and sTGC in the NSW and BIS78 RPC and sMDT), and those that will remain unchanged from the Run 1 layout (black text). In the so-called large sectors, which are the sectors in-between the barrel toroid coils, the TGCs and MDT chambers in the endcap inner station covers the  $\eta$ - $\phi$  range of the BIS78.

replaced by NSW during Long Shutdown 2, which consists of sTGC and MM, for further suppression of the fake triggers [8.2].

The main reasons to upgrade the current Level-1 Muon Trigger System are the limited available latency and rate and the relatively low efficiency and momentum resolution. It is impossible for the original RPC and TGC electronics to cope with the latencies and the rates of the trigger and readout scheme for HL-LHC. The original systems of RPC and TGC are designed for maximum latencies of  $6.4 \mu\text{s}$  and  $3.2 \mu\text{s}$ , respectively, and for a maximum rate of 100 kHz. The limitations arise from the depth of the readout buffer to store the hit data before the arrival of the trigger accept signals and also from the readout bandwidth. The NSW was designed for a  $60 \mu\text{s}$  latency and a 1 MHz rate to be compatible with the Phase-II system.

The current Level-1 muon trigger has limitations on the performance. The product of the acceptance and the efficiency for the barrel region is about 70% (Fig. 8.2). This is due to limited coverage of RPC in the barrel middle station, which is originating from the sup-

port structures of the toroid magnets [8.5]. The product for the endcap region is about 90%, which is due to the accumulation of minor effects including a relatively tight coincidence criterion. Additionally, there is a substantial contribution in the trigger rates from the muons with the transverse momentum  $p_T$  smaller than the threshold (Fig. 8.3). This is due to a rapid increase of the production cross section of muons for lower  $p_T$  and the limited  $p_T$  resolution of the Level-1 muon trigger.

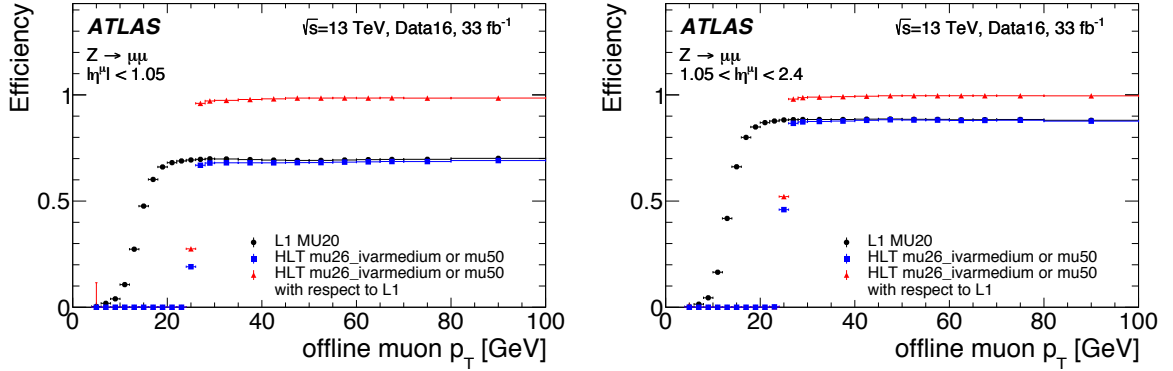


Figure 8.2: Product of the acceptance and the efficiency for single muon as a function of transverse momentum  $p_T$  evaluated by an offline analysis. The plots are obtained from Run 2 data. The left and right plots are for the barrel and endcap regions, respectively. The circles show the values for the Level-1 muon trigger with 20 GeV threshold. The triangles show the values for the high-level trigger targeting isolated muons with 26 GeV threshold and non-isolated muons with 50 GeV threshold. The squares show the product of them. The values of the Level-1 muon trigger for  $p_T$  higher than the threshold is about 70% in the barrel region.

### 8.3 Overview of the Upgrade

The upgrade of the Level-0 Muon Trigger System is designed to increase the latencies and the rates of the trigger and readout system and to improve the trigger efficiency. The detector and on-detector electronics upgrades are described in Ref. [8.1]. This TDR focuses on the off-detector electronics upgrades.

Figure 8.4 shows a simplified block diagram of the proposed system. It consists of the Barrel Sector Logic, the Endcap Sector Logic, the NSW Trigger Processor, and the MDT Trigger Processor. The Sector Logic and the NSW Trigger Processor installed by the end of Long Shutdown 2 will be replaced, while the MDT Trigger Processor will be newly installed. The Barrel Sector Logic receives the hits and the energy flags from RPC and Tile calorimeter, respectively. The Endcap Sector Logic receives the hits from TGC and RPC in  $1.05 < |\eta| < 1.3$ . It also receives the track segments and the energy flags from NSW and Tile calorimeter, respectively. Since the track-segment reconstruction with NSW hits

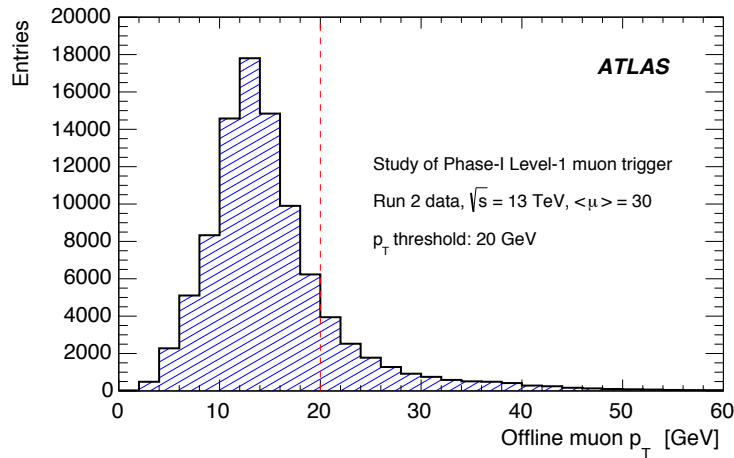


Figure 8.3: Distribution of the transverse momentum  $p_T$  evaluated by an offline analysis for the muon candidates selected by the current Level-1 muon trigger with 20 GeV threshold. The distribution is obtained from Run 2 data. The requirement based on [NSW](#), which will be integrated after Run 2, has been simulated and applied based on the signals from the current detectors in the endcap inner station. Both barrel and endcap regions are included. The entries are dominated ( $\sim 85\%$ ) by the muon candidates with  $p_T$  smaller than the threshold.

requires relatively large amount of resources, the hardware will be separate from the Sector Logic. The Sector Logic provides track candidates, and sends the results to the [MDT](#) Trigger Processors. The [MDT](#) Trigger Processor filter the track candidates with a better  $p_T$  determination, and sends the results to the Sector Logic. The Sector Logic sends the track candidates to Level-0 [MUCTPI](#).

#### 8.3.1 Extension of the Latency and the Rate

To increase the latencies and the rates of the trigger and readout system, all the [RPC](#) and [TGC](#) electronics boards except for the amplifier-shaper-discriminator boards will require the replacements [8.1]. In the new design, all hit data will be transferred from the on-detector boards to the off-detector boards over high-speed optical links. The limitations by the bandwidth between the on-detector and off-detector boards will be eliminated. No trigger logic will be implemented in the on-detector boards. The trigger logic will be fully implemented in the Sector Logic, providing the flexibility of modifying and tuning the logic. The readout buffer will be implemented in the off-detector boards. It is designed to have a depth sufficient for the maximum latency of the Level-0 trigger for [HL-LHC](#).

The [NSW](#) on and off-detector electronics delivered for Phase-I will remain for Phase-II. The main update foreseen for the system is for additional [FELIX](#) boxes to manage the additional data from the 1 MHz trigger rate [8.2]. The upgrade of the [NSW](#) Trigger Processor

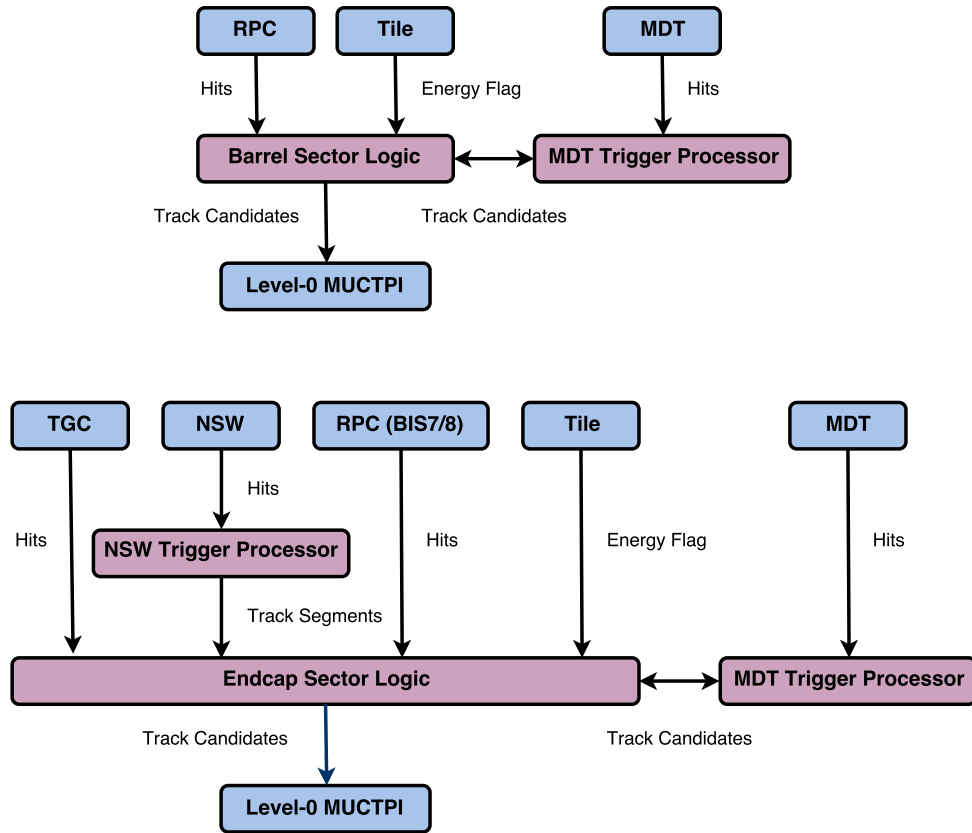


Figure 8.4: A simplified block diagram of the proposed Level-0 Muon Trigger System for *HL-LHC*. The top and bottom parts show the diagrams for the barrel and the endcap, respectively.

(see Section 8.5) is proposed to improve the robustness and redundancy of the trigger by performing a combined track segment reconstruction by *sTGC* and *MM*.

### 8.3.2 Improvement of the Trigger Performance

Improvements in trigger performance will be achieved by increasing the detector acceptance, the trigger logic, and the momentum resolution. The upgrade of the muon system towards Phase-II includes the addition of *RPCs* in the barrel inner station as described in Ref. [8.1]. By allowing for coincidences of hits in the *RPCs* in the inner and outer stations, the acceptance gaps in the barrel trigger system which are caused by the non-instrumented regions in the barrel middle station due to the presence of the structures of the barrel toroid coils will be closed. The expected value of the product of the acceptance and the efficiency of the muon trigger in the barrel region is more than 90% depending on the coincidence criteria. The Barrel Sector Logic in the new design includes the coincidence of the addi-

tional [RPCs](#). The results of the performance studies are described in the former parts of Section [8.4.1](#). The Barrel Sector Logic also includes the coincidence of the [Tile](#) calorimeter for a trigger rate suppression.

The Endcap Sector Logic in the new design is based on a loosened [TGC](#) coincidence for an improvement of the efficiency by a few per cent. The Endcap Sector Logic reconstructs the track segments with [TGC](#) hits in the endcap middle station with a 4 mrad resolution. The [NSW](#) Trigger Processors reconstruct the track segments in the endcap inner station with a 1 mrad resolution [[8.2](#)]. The deflection angle between the [TGC](#) and [NSW](#) segments makes it possible to reduce the trigger rate in  $1.3 < |\eta| < 2.4$ , while retaining the efficiency for the muons with the transverse momentum higher than the threshold. The fake triggers in  $1.05 < |\eta| < 1.3$  are suppressed by the coincidence between [TGCs](#), [RPCs](#) in the BIS78 region [[8.1](#)], and the [Tile](#) calorimeter. The results of the performance studies of the trigger efficiency and rate are described in the latter parts of Section [8.4.1](#). The performance of the segment reconstruction based on [NSW](#) hits is described separately in Section [8.5.2](#).

In order to improve the  $p_T$  resolution at the Level-0 muon trigger, it is proposed to include the [MDT](#) chambers, which are used for the high-level trigger and the precision tracking in the current system, in the Level-0 Muon Trigger System for [HL-LHC](#). In the new design, the data of all [MDT](#) hits are transferred to the off-detector boards over high-speed optical links [[8.1](#)]. The data include the hit time measurements. The [MDT](#) trigger processing is only applied when muon track candidates are provided by the Sector Logic. The [MDT](#) drift time is measured with the time origin determined from the bunch crossing identified by [RPCs](#) and [TGCs](#). The [MDT](#) Trigger Processor reconstructs the track segments and evaluates  $p_T$  from the relation of the angles and the positions of the segments in different stations. The  $p_T$  evaluated by [MDT](#) makes it possible to reduce the trigger rate. The results of the performance studies are described in Section [8.6.2](#).

## 8.4 Sector Logic

The trigger logic is provided for each sector defined by the boundaries of combinations of [RPC](#) and [TGC](#) chambers [[8.3](#)], and hence is called the Sector Logic. The Sector Logic is implemented in [FPGAs](#) on the off-detector boards. This section introduces the logic scheme, the expected performance, the hardware and firmware design, and the interfaces of the Sector Logic upgrade.

## 8.4.1 Trigger Scheme and Performance

### Barrel Trigger Scheme

The RPC trigger logic in Phase-II will use nine measurement planes, provided by four groups of RPC chambers: three planes (RPC0) on the barrel inner (BI) station, two planes (RPC1) in the inner part of the barrel middle (BM) station, two planes (RPC2) in the outer part of the BM station, and two planes (RPC3) on the barrel outer (BO) station. Figure 8.5 shows an example of a so-called “small” barrel sector, showing the positions of RPC0, RPC1, RPC2, and RPC3 chambers. The acceptance holes in the RPC1 and RPC2 chambers, caused by the magnet coils and their supports, are clearly visible.

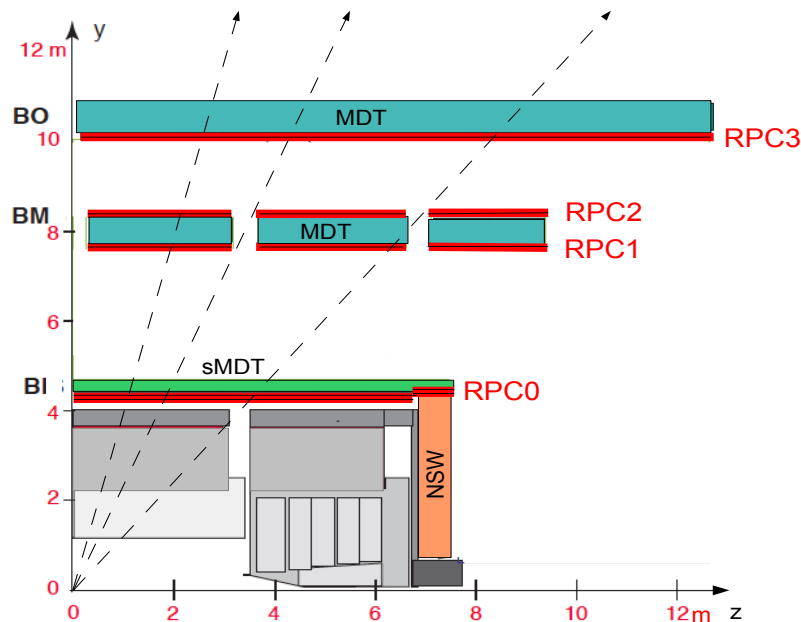


Figure 8.5: Sketch of a transverse section of the barrel region. The four groups of RPC chambers (red) are shown as well as the MDT chambers (green and cyan) on the BI, BM, and BO stations. The three dashed lines represent muon trajectories traversing four, two, and three RPC chambers. The drawing represents one of the sectors that contain a barrel toroid coil and its support structures which cause the holes in the chamber coverage of the BM station.

To take advantage of the redundancy of detector planes, a trigger algorithm that does not make use of a fixed pivot plane (as in present ATLAS scheme) has been developed, which selects patterns of hits with a minimal deviation from a straight line passing from the nominal interaction point, both in  $\eta$  and in  $\phi$ . This makes it possible to define different trigger coincidence logic schemes without the constraint of requiring at least one hit on a fixed

pivot plane. The following logic schemes have been considered, based on different requirements on the four groups of **RPC** chambers which are explained in the following and summarised symbolically in Table 8.1:

- **3/3 chambers.** Hits in at least three out of four planes of the RPC1+RPC2 chambers and in at least one out of two planes of RPC3. This is equivalent to the present high- $p_T$  trigger.
- **3/4 chambers.** The previous requirement in logical OR with the requirement of hits in at least two planes out of three in RPC0 and in at least three planes out of six in RPC1+RPC2+RPC3. In this way, all combinations of three-chamber coincidences (satisfying the above hit requirements) are accepted.
- **3/4 chambers + BI-BO.** The previous requirement in logical OR with the requirement of at least two hits in RPC0 and at least one hit in RPC3. This enhances the trigger coverage in the regions where no BM RPCs are installed due to the mechanical support structure of the toroid coils.

Table 8.1: Detail of the hit requirements used in different **RPC** triggers. The cell with “ $x$  out of  $y$ ” indicates the requirement of at least  $x$  planes out of  $y$  planes. The cell with “–” indicates no use of the layer. The requirements in a row is combined with logical AND. For each set of “3/4 chambers” and “3/4 chambers + BI-BO”, the requirements in different columns are combined with logical OR.

Requirement	RPC0	RPC1	RPC2	RPC3
3/3 chambers	–	3 out of 4		1 out of 2
3/4 chambers	–	3 out of 4		1 out of 2
	2 out of 3	3 out of 6		
3/4 chambers + BI-BO	–	3 out of 4		1 out of 2
	2 out of 3	3 out of 6		
	2 out of 3	–	–	1 out of 2

The baseline of the **RPC** trigger uses the “3/4 chambers + BI-BO” scheme for all  $p_T$  thresholds. The BI-BO coincidence is expected to be prone to accidental coincidences of uncorrelated background hits that are negligible in three-chamber coincidences. There is an option to apply the BI-BO coincidence only to the regions with acceptance holes in the “3/4 chambers” scheme (approximately 15% of the barrel) to control the trigger rate. The “3/4 chambers” and “3/3 chambers” schemes are other options to further control the trigger rate. For the “3/4 chambers” scheme, an additional requirement of at least one hit on the BO station can be included. The **Tile** calorimeter energy flags are used to suppress the accidental coincidences.



### Barrel Trigger Acceptance and Efficiency

The performance of the RPC-based trigger was studied using simulated data with a simplified description of the BI RPCs. In the simplified description, the geometrical coverage of the BI RPC chambers was included according to the preliminary layout presented in Ref. [8.1]. A dead region of 15 mm at the borders of each RPC chambers is included in the simulation. The presence of these cut-outs reduces the geometrical acceptance of the BI station, calculated with respect to reconstructed muons with  $|\eta| < 1.05$ , from 95% to 91%.

A Phase-II RPC-based trigger algorithm was implemented in the simulation. The algorithm searches for sequences of hits in different RPC planes based on pre-defined coincidence windows that define the maximum distance in  $\eta$  or  $\phi$  to associate hits belonging to the same muon candidate. Not only the hits of the bending plane ( $\eta$ ) but also the hits of the perpendicular plane ( $\phi$ ) are used to suppress the accidental coincidence. The requirement that two hits are inside a coincidence window is applied to each pair of hits in consecutive planes, as illustrated in Fig. 8.6. The centre of the coincidence window on the outer layer is obtained assuming a straight line passing from the nominal interaction point and the inner hit. This algorithm effectively measures the muon momentum from the deflection of the trajectory with respect to a straight line from the interaction point. Therefore the momentum resolution that can be obtained with this approach is limited by the spread of the interaction point along the beam line rather than the spatial hit resolution.

In the simulation a luminous region with a Gaussian distribution around the nominal interaction point with  $\sigma = 5$  cm was used. Only a simple digital readout was simulated, and neither strip clustering nor any use of charge information was implemented. In addition to the spatial coincidence, a timing requirement has been applied, by considering only hits with  $|t_{\text{hit}} - t_0| < 6.25$  ns. Here,  $t_{\text{hit}}$  is the time of hit detection and  $t_0$  is the time of an infinite momentum muon which hits the centre of the strip and was generated at the nominal interaction point at the nominal bunch crossing time.

The acceptance of the three coincidence logic schemes with respect to reconstructed muon tracks is shown in Fig. 8.7. The acceptance limitation of the “3/3 chambers” trigger, corresponding mostly to the ATLAS support structures at  $-2.2 < \phi < -1$  and to the supports of the toroid magnets (at  $|\eta| \simeq 0.4, 0.75, 1$ ) are largely recovered by the “3/4 chambers” trigger which includes the new BI RPC chambers. The residual acceptance holes are recovered with the “3/4 chambers + BI-BO” trigger, with the exception of the region at  $|\eta| \simeq 0$  where the acceptance is limited by the calorimeter services.

To study the robustness of the trigger against possible efficiency reductions due to aging of the old RPCs in the BM and BO stations, the performance of different assumptions on the old RPC detector efficiency has been studied with a simulation of single muons generated with a fixed  $p_T$  of 25 GeV and uniform distributions in the  $\eta$  and  $\phi$  plane. Simulations were performed with 100%, 90%, and 80% hit efficiency for the BM and BO RPCs and with the “worst case scenario”. The worst case scenario corresponds to a reduction of high voltage

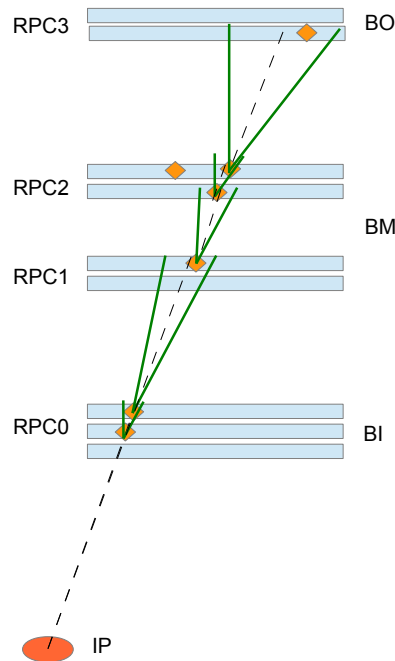


Figure 8.6: Illustration of the implemented *RPC*-based trigger algorithm. Hits are linked between consecutive planes based on pre-defined coincidence windows centred on a straight-line extrapolation from the nominal interaction point. Patterns of linked hits that satisfy quality requirements are selected as muon track candidates. The width of the coincidence windows defines the  $p_T$  threshold.

of old *RPCs* such that the expected current on the *RPC* chambers is always below the safe operation limit, estimated including a safety factor of two on the currents. In addition to the inefficiency from reduced HV, a 95% hit efficiency as measured in Run 2 was applied. In this scenario the *RPC* hit efficiency decreases with  $|\eta|$  from approximately 90% at  $\eta \simeq 0$  to 56%–75% at  $\eta \simeq 1$ , depending on the chamber type. Table 8.2 shows the product of the acceptance and the efficiency for the difference assumptions. Figure 8.8 shows the product of the acceptance and the efficiency for the worst case scenario. Adding the new BI *RPC* station, the dependency of the trigger efficiency on the hit efficiency of the current *RPCs* is strongly reduced.

### Barrel Momentum Selection and Trigger Rate

The performance of the proposed *RPC* trigger was studied for transverse momentum thresholds of  $p_T > 20$  GeV, 15 GeV, 10 GeV, and 5 GeV. The dependencies of trigger efficiency on the muon  $p_T$  are shown in Fig. 8.9. For the 20 GeV threshold, the plots for different coincidence schemes are shown. Some residual efficiency below 10 GeV is related to the class of

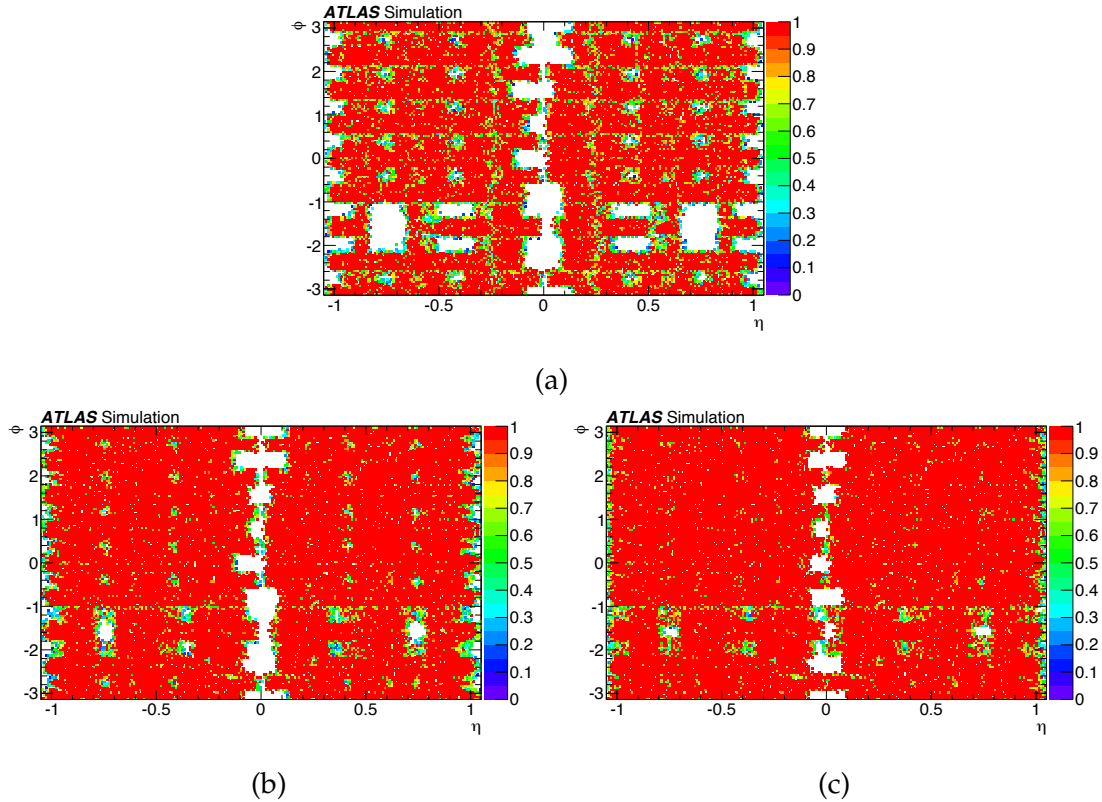


Figure 8.7: Geometrical acceptance of the Level-0 barrel muon trigger with respect to reconstructed muons with  $p_T = 25$  GeV in the  $\eta$ - $\phi$  plane. The plots are obtained by assuming 100% hit efficiency and no pile-up. Figures (a), (b), and (c) show the acceptance for the different trigger coincidence logic “3/3 chambers”, “3/4 chambers”, and “3/4 chambers + BI-BO”, respectively. The white areas correspond to zero acceptance.

Table 8.2: Acceptance  $\times$  efficiency in per cent for the *RPC* trigger under different hypotheses on the hit efficiency of the present *RPC* detectors. The row with 100% “old” *RPC* efficiency gives the geometrical acceptance of the trigger requirements. The “worst case” corresponds to a scenario in which the *RPC* HV is reduced depending on the expected chamber rate to maintain the currents within a safe limit.

“Old” <i>RPC</i> efficiency (%)	Trigger acceptance $\times$ efficiency (%)		
	3/3 chambers	3/4 chambers	3/4 chambers + BI-BO
100	78	91	96
90	73	90	95
80	62	87	93
Worst case	63	85	92

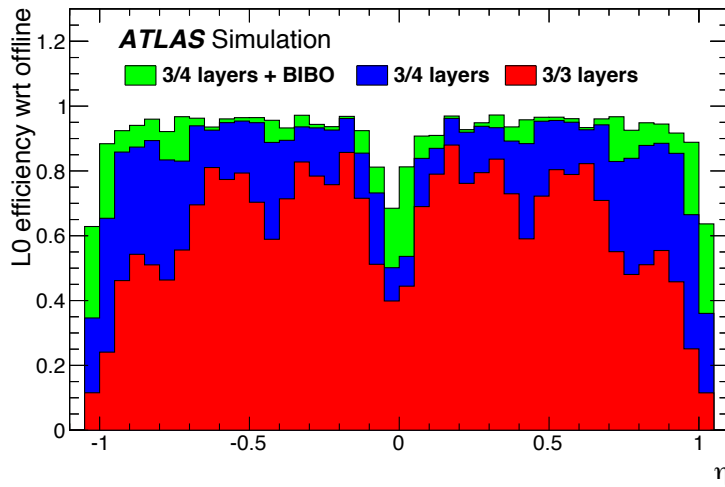


Figure 8.8: *Acceptance  $\times$  efficiency of the RPC trigger with respect to reconstructed muons as a function of  $\eta$  for the “worst case” hit efficiency scenario in which the RPC HV is reduced to maintain the chamber currents within a safe limit with a safety factor of two. The red histograms show the efficiency of the existing “3/3 chambers” trigger, the blue histograms the “3/4 chambers” trigger, and the green histograms the “3/4 chambers + BI-BO” trigger. The values are evaluated by the Monte-Carlo (MC) samples of muons with a fixed  $p_T$  of 25 GeV and no pile-up involved.*

muons that leave hits only in BI and BM chambers, for which the resolution is worse than for muons with hits in the BO chambers.

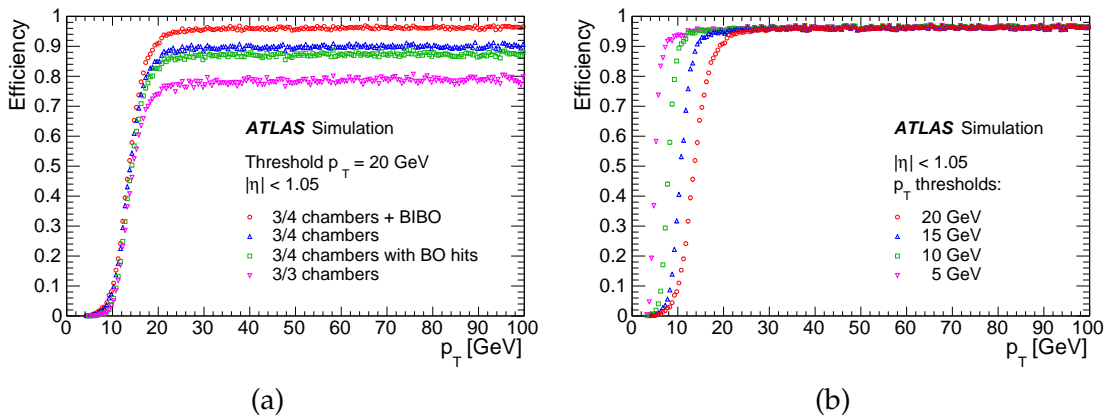


Figure 8.9: *Efficiency curves as a function of  $p_T$  for RPC Level-0 trigger. Figure (a) shows the result for different coincidence schemes for a  $p_T$  threshold of 20 GeV. Figure (b) shows the result for different  $p_T$  thresholds for the “3/4 chambers + BI-BO” scheme. The figures assume 100% RPC hit efficiency and no pile-up.*

The rates of single muon selection based on [RPC](#) have been estimated using a simulation of minimum bias events generated with PYTHIA8 and passed through the standard GEANT4 ATLAS simulation, including the long-lifetime background component originating from slow neutrons (cavern background). Additional corrections have been applied by scaling the rate of muons from heavy-flavour decays in the simulation to agree with ATLAS measurements [8.6] and FONLL calculations [8.7] and by scaling the [RPC](#) hit rate obtained from the simulation in order to reproduce rate measurements. Figure 8.10 shows the estimated rate depending on the luminosity for a  $p_T$  threshold of 20 GeV. The rate of “3/3 chambers” and “3/4 chambers” triggers increase linearly with luminosity as they are mainly generated by real muons from the decay of pions, kaons, and heavy flavours that scale linearly with the number of  $pp$  interactions. The BI-BO trigger instead has an additional quadratic dependence on pile-up due to accidental coincidences of background hits. Figure 8.11 shows the estimated rate depending on the  $p_T$  threshold for the “3/4 chambers + BI-BO” trigger. The estimated trigger rates for different  $p_T$  thresholds and different schemes are summarised in Table 8.3.

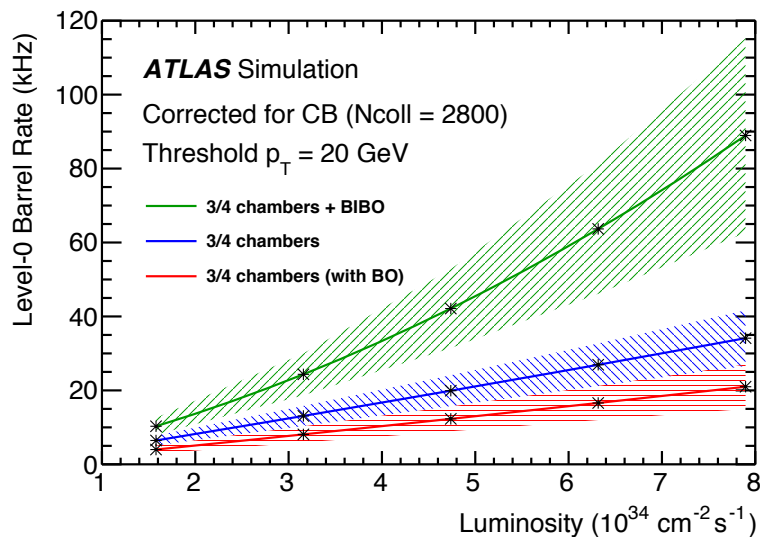


Figure 8.10: Estimated rate of Level-0 single muon selection based on [RPC](#) depending on the luminosity. The result is shown for a  $p_T$  threshold of 20 GeV and for different coincidence schemes.

For the subset of muons with hits in the three stations of chambers (BI, BM, and BO), which correspond to an acceptance of approximately 80%, it is possible to obtain a  $p_T$  measurement based on three points. This measurement is independent of the spread of the interaction point and therefore provides an improved  $p_T$  resolution.

The effect of random coincidences on BI-BO triggers can be reduced by applying tighter timing cuts. The intrinsic time resolution of the [RPC](#) detectors is 1 ns for BM and BO and

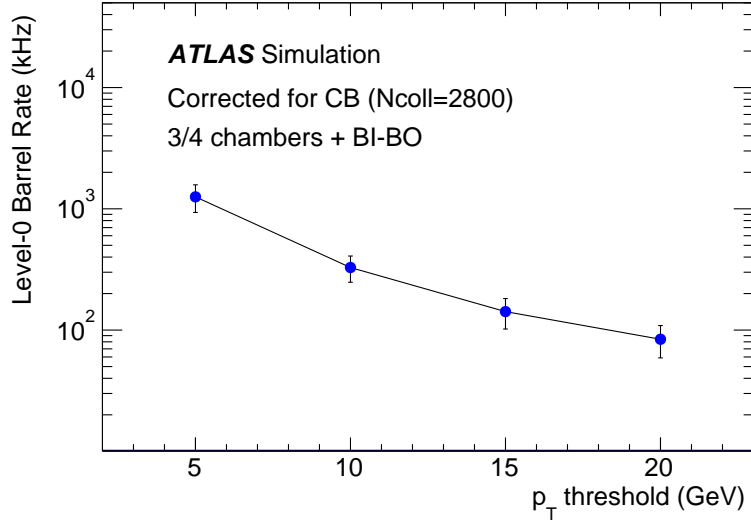


Figure 8.11: Estimated rate of Level-0 single muon selection based on *RPC* depending on the  $p_T$  threshold. The result is shown for a coincidence scheme of “3/4 chambers + BI-BO”.

Table 8.3: Estimated rate of Level-0 single muon selection based on *RPC* for different  $p_T$  thresholds and different schemes at a luminosity of  $7.5 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ . The values are shown in units of kHz. These rates are reduced by the *MDT* trigger before the muon track candidates are sent to the Level-0 *MUCTPI*.

	$p_T > 20 \text{ GeV}$	$p_T > 15 \text{ GeV}$	$p_T > 10 \text{ GeV}$	$p_T > 5 \text{ GeV}$
3/4 chambers + BI-BO	$84 \pm 25$	$142 \pm 40$	$328 \pm 80$	$1255 \pm 322$
3/4 chambers	$32 \pm 7$	$73 \pm 11$	$187 \pm 16$	$718 \pm 31$
3/4 chambers (with BO)	$20 \pm 6$	$52 \pm 9$	$146 \pm 14$	$516 \pm 26$

0.4 ns for BI. In the standard ATLAS offline reconstruction, a resolution of less than 2 ns is routinely achieved for the BM and BO chambers after offline calibration and corrections for the signal propagation along the strips. The current experience with the ATLAS barrel trigger shows that a simple online calibration allows the use of a timing window  $|t_{\text{hit}} - t_0| < 6.25 \text{ ns}$ . A tighter window would require signal propagation corrections and finer calibrations at the Level-0 trigger.

### Endcap Trigger Scheme

The availability of the individual *TGC* hits in the Sector Logic makes it possible to reconstruct track segments in the endcap middle station with an average angular resolution of

4 mrad (Fig. 8.12). The track segments are reconstructed with the pattern matching algorithm, where the TGC hits are compared with predefined hit patterns. Each predefined hit pattern has associated angle and position of a track segment.

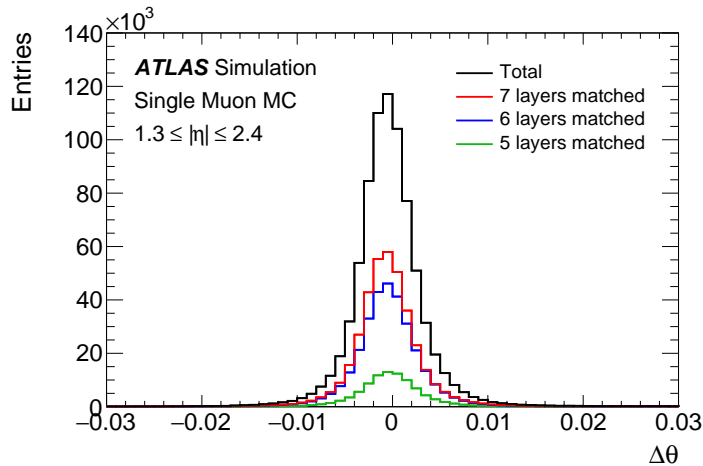


Figure 8.12: Distributions of the difference of the polar angle ( $\Delta\theta$ ) between the TGC track segment reconstructed by the pattern matching for the trigger threshold of 20 GeV and the track segment reconstructed by the ATLAS full offline analysis. MC sample is used, where a muon is produced with random  $p_T$  in a range 1–100 GeV. No pile-up is involved. The red, blue, and green histograms are for the TGC track segments reconstructed with seven, six, and five hits, respectively, over the seven layers. The black histogram shows the sum of them.

In the original Endcap Sector Logic, at least two (one) hits are required in the inner three (two) layers and at least three hits are required in the outer four layers for wires (strips) for the TGC hits in the endcap middle station. In the Endcap Sector Logic for HL-LHC, a looser coincidence of at least five (four) hits in seven (six) layers for wires (strips) will be used. The product of the acceptance and the efficiency is expected to be improved to more than 90%, a few per cent higher than the original system.

The muon track candidates are provided by the combination of the TGC track segments in the endcap middle station and the information from other detectors. Figure 8.13 shows the concept of the combination. The transverse momentum of a muon candidate in  $1.05 < |\eta| < 1.3$  is determined from a combination of the position and the angle of the TGC track segment and the positions of the hits in the TGCs in the endcap inner station, the RPCs in the BIS78 region, and the Tile calorimeter. The transverse momentum of a muon candidate in  $1.3 < |\eta| < 2.4$  is determined from the polar angle difference between the TGC track segment and the track segment reconstructed in NSW.

The baseline algorithm introduced above covers the region  $1.05 < |\eta| < 2.4$ . Although the number of TGC layers in  $2.4 < |\eta| < 2.7$  is limited (three wire layers and two strip layers), the coverage could, in principle, be extended to the region  $1.05 < |\eta| < 2.7$  by a coincidence between NSW track segments and the hits in the existing TGC layers.

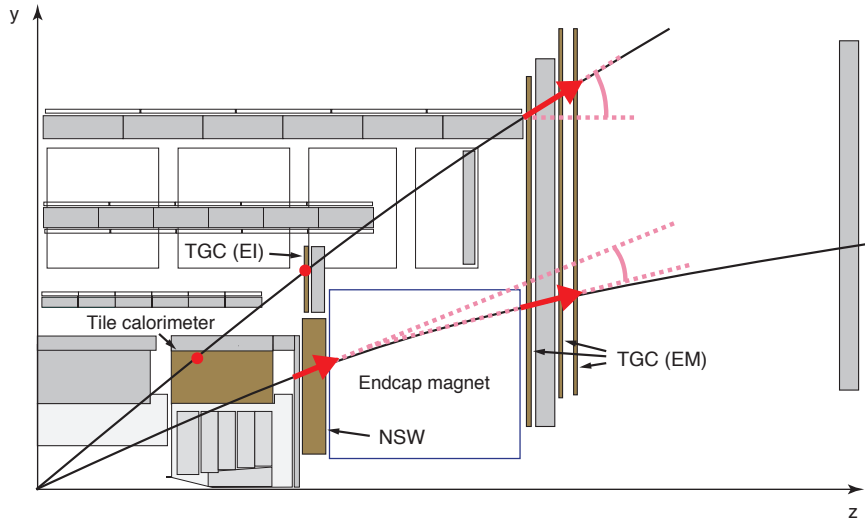


Figure 8.13: Concept of the Endcap Sector Logic. Track segment is reconstructed by the TGC hits in the endcap middle (EM) station. It is combined with the position of the TGC hits in the endcap inner (EI) station, the position of the Tile calorimeter hit, and the track segment reconstructed by NSW, depending on the region. In this diagram, a detector slice in a so-called large sector is shown. In the so-called small sectors, the RPCs in the BIS78 region replaces the TGC hits in the endcap inner (EI) station.

### Endcap Trigger Efficiency

The efficiency of the Level-0 muon trigger based on the loosened TGC coincidence for  $1.05 < |\eta| < 2.4$  has been studied with MC samples. A muon is produced in an event randomly in  $\eta$ ,  $\phi$ , and  $p_T$ . A track segment is reconstructed by a minimum  $\chi^2$  fit to the TGC hits which satisfy the coincidence requirement. The efficiency for  $p_T$  thresholds of 5, 10, 15, and 20 GeV have been studied. Requirements are applied for the polar angle of the segment direction depending on the  $p_T$  thresholds. Figure 8.14 shows the result of the efficiency estimation. The obtained efficiency in the plateau region is a few per cent higher than the original endcap system (see Fig. 8.2).

The efficiency has also been studied with the TGC segment reconstruction based on the pattern matching. The study focuses on the region covered by both TGC and NSW. A requirement on the polar angle difference between the TGC and NSW segments is applied depending on the  $p_T$  thresholds. Figure 8.15 shows an estimate based on the single muon MC samples, described in the previous paragraph. The obtained efficiency in the plateau region is similar to the one obtained for the minimum  $\chi^2$  method (see Fig. 8.14). The distribution around the threshold is steeper than the original system, indicating a better transverse momentum resolution.



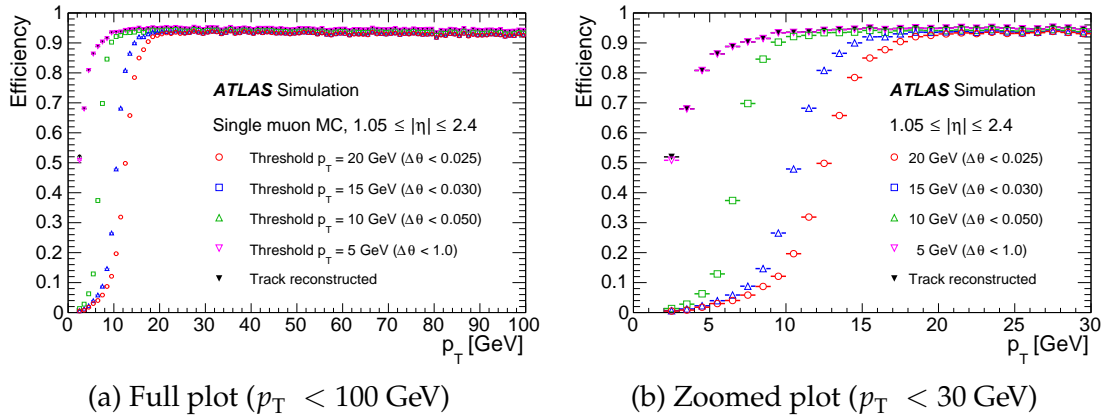


Figure 8.14: Expected efficiency for the Level-0 muon trigger based on the loosened TGC coincidence for  $1.05 < |\eta| < 2.4$ . The values are estimated with a single muon MC sample, where a muon is produced for an event randomly in  $\eta$ ,  $\phi$ , and  $p_T$ . No pile-up is involved. The red, blue, green, and magenta plots are for  $p_T$  thresholds of 20, 15, 10, and 5 GeV, respectively. The black plots show the efficiency for offline track reconstruction.

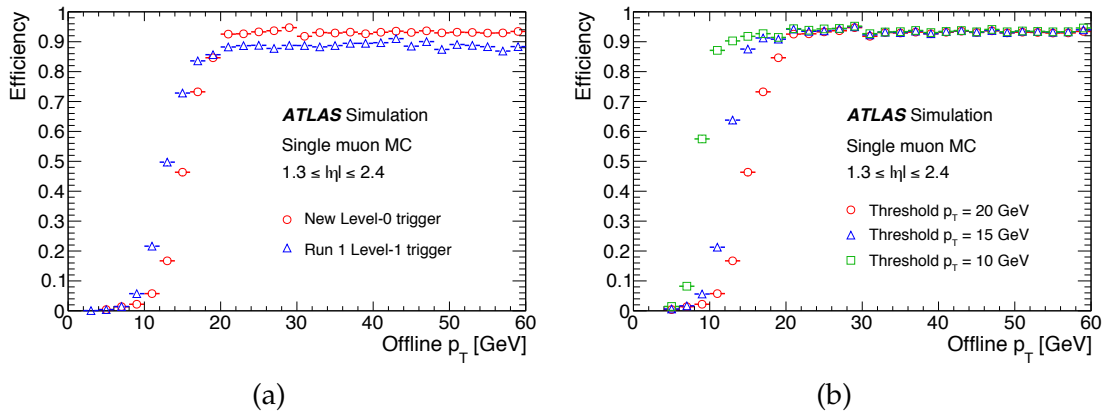


Figure 8.15: (a) Expected efficiency for the muon trigger based on TGC and NSW in the region  $1.3 < |\eta| < 2.4$  for a  $p_T$  threshold of 20 GeV. The plots for the Run 1 scheme and a HL-LHC scheme are shown. In the HL-LHC scheme, a looser coincidence, five (four) hits over seven (six) layers for wires (strips) is used. The HL-LHC scheme provides a higher efficiency in the plateau region with better rejection of low  $p_T$  muons. (b) Expected efficiency for the muon trigger based on a HL-LHC scheme with TGC and NSW for  $p_T$  thresholds of 10, 15, and 20 GeV. No pile-up is involved in the MC simulation.

### Endcap Trigger Rate

The Level-0 muon trigger rate for the endcap has been studied with the data samples taken in 2016. A zero-bias trigger is required [8.8]. The events have been overlaid to provide an expected environment at HL-LHC for the number of interactions per bunch crossing up to around 200 (Fig. 8.16). The distributions are shown for  $p_T$  thresholds of 5, 10, 15, and 20 GeV. The efficiency in the plateau region is obtained to be 92%.

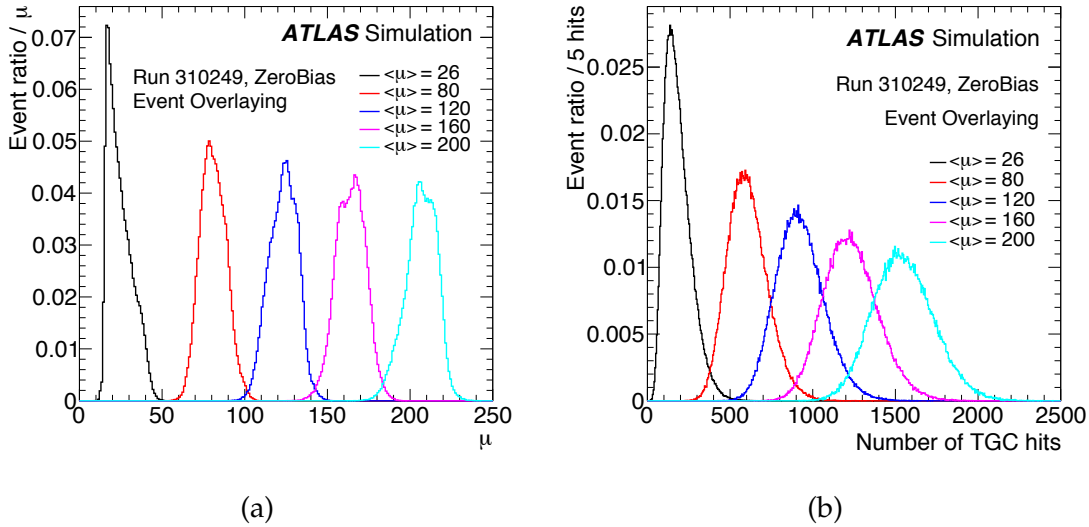


Figure 8.16: (a) Distributions of the number of interactions per bunch crossing ( $\mu$ ) for the overlaid data samples. The overlay was done with the target average  $\mu$  around 80, 120, 160, and 200. (b) Distributions of the number of TGC hits for the overlaid data samples. The number depends on the values of  $\mu$ .

Figure 8.17 shows the result of the estimation of the Level-0 single-muon trigger for  $1.05 < |\eta| < 2.4$  for  $p_T$  thresholds of 15 GeV and 20 GeV. Figure 8.18 shows the estimated rate depending on the  $p_T$  thresholds. Expected improvements, the looser TGC coincidence and the requirement based on the TGC and NSW segments, are both included. The TGC coincidence requires at least five (four) hits in seven (six) layers for wires (strips). The TGC segments are reconstructed by a minimum  $\chi^2$  fit on the TGC hits, and requirement is applied for the angles. The NSW segments are emulated by the offline MDT and Cathode Strip Chamber (CSC) segments in  $1.3 < |\eta| < 2.4$ , and requirement is applied for the deflection angle between the TGC and NSW segments. The coincidence with the Tile calorimeter is required for  $1.05 < |\eta| < 1.3$ . A nonlinear relation between the trigger rate and the luminosity is obtained, which is related with a larger number of TGC hits for higher luminosity (Fig. 8.16). A tuning of the trigger algorithm is ongoing to remove or suppress the nonlinearity, while the estimated rate is already smaller in the total Level-0 rate of 1 MHz.

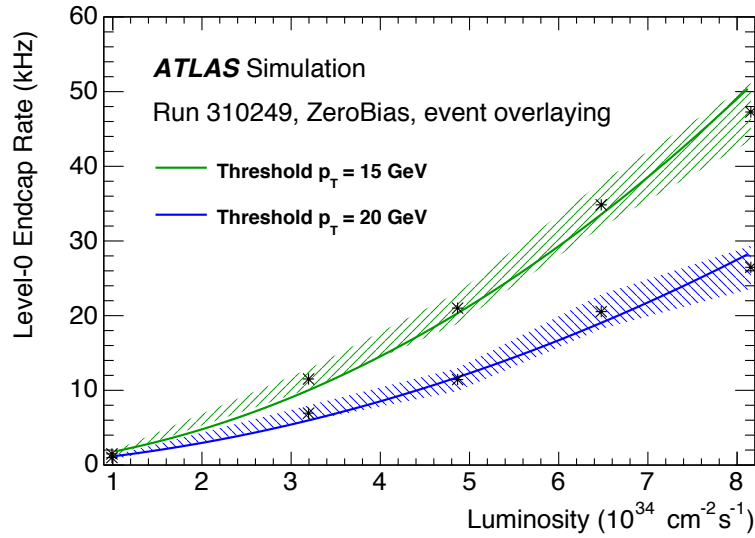


Figure 8.17: The estimated rate of the Level-0 single-muon trigger based on *TGC*, *Tile* calorimeter, and *NSW* for  $1.05 < |\eta| < 2.4$ . The curves show the result of the fit by a second-order polynomial. The luminosity has been extracted from the number of interactions per bunch crossing for the *HL-LHC* nominal colliding bunches.

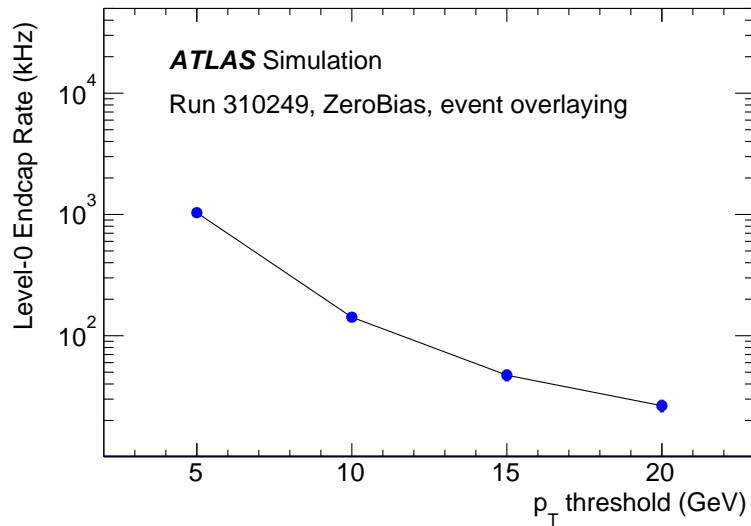


Figure 8.18: The estimated rate of the Level-0 single-muon trigger based on *TGC*, *Tile* calorimeter, and *NSW* for  $1.05 < |\eta| < 2.4$ . The values are evaluated from the overlaid data samples for a luminosity of  $7.5 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ . The values are estimated for  $p_T$  thresholds of 5, 10, 15, and 20 GeV.

### Trigger for Exotic Signatures

Several new physics models that extend the [Standard Model \(SM\)](#) predict exotic signatures such as the [Long-Lived Particle \(LLP\)](#) [8.9]. The sensitivity of the searches could be improved by specific trigger algorithms. An example is the algorithm with loosened interaction-point pointing constraint. The Sector Logic is designed not to kill the exotic signatures.

#### 8.4.2 Hardware Design

The Sector Logic is implemented in the off-detector boards housed in the [ATCA](#) shelves. The readout of the [RPC](#) and [TGC](#) hit signals and of the Level-0 trigger information is implemented in the same boards. Additionally, the control of the [RPC](#) and [TGC](#) on-detector boards is based on the same boards. The barrel and endcap systems are based on  $2 \times 16$  and  $2 \times 24$  boards (Fig. 8.19), where the number “2” represents the positive and negative sides of the beam axis.

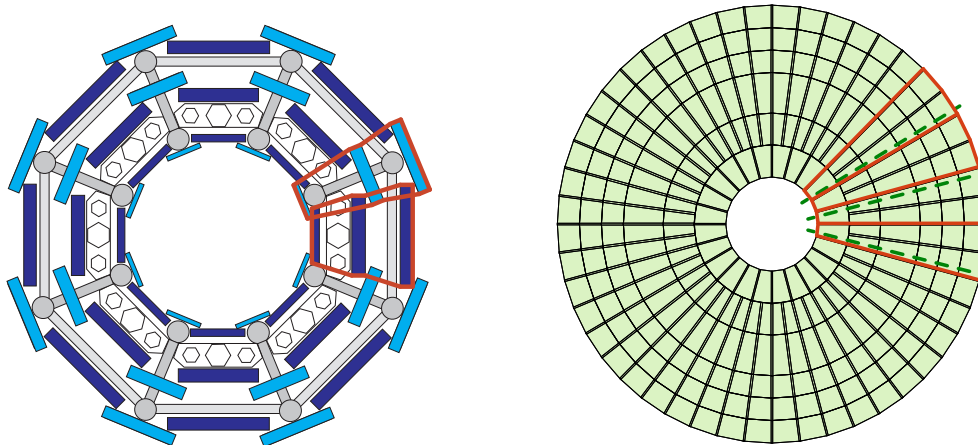


Figure 8.19: Structure of the detector and the coverage of the off-detector boards where the Sector Logic is implemented. The left and right figures are for the barrel and the endcap, respectively. The outermost layer of [TGC](#) is shown for the endcap. The regions surrounded by the red lines show the coverage of the boards. For the barrel, the coverage has two types, which is indicated by the red lines. For the endcap, the coverage for four boards are indicated by the red lines. The segmentation of the [MDT Trigger Processor](#) (green lines) is different, and a Sector Logic board is connected to three [MDT Trigger Processor](#) boards at maximum.

The requirement for the off-detector boards is similar between the barrel and the endcap. A common board design is proposed to minimise the resources required for the development. Figure 8.20 shows a block diagram. Each board has one [FPGA](#), in which the Sector

Logic and the readout logic are implemented. The Xilinx Virtex UltraScale or UltraScale+ family is proposed for the implementation. For the barrel, the resource of the Sector Logic has been estimated with an extrapolation from the current coincidence matrix ASIC implementation. The implementation of the logic in the proposed FPGA family seems feasible. For the endcap, the resource estimation for the Sector Logic is an R&D item. Given the total number of TGC channels similar to RPC, the implementation in the proposed FPGA family is considered to be feasible. For the cost estimation, the FPGA models XCVU13P and XCVU190 are assumed for the barrel and the endcap.

The FPGA on the off-detector board is required to have 104 pairs of optical receivers and transmitters. Most of the receivers and transmitters are used to receive the detector signals and to transmit the control signals for on-detector boards. Some of the receivers and transmitters are used for the connection with the MDT Trigger Processor and the Level-0 MUCTPI to receive and transmit the muon track candidates. Four transmitters are used for the connection with FELIX to transmit the readout data. One receiver is used to receive the signals of TTC. Tables 8.4 and 8.5 show the number of receivers and transmitters depending on the connected electronics.

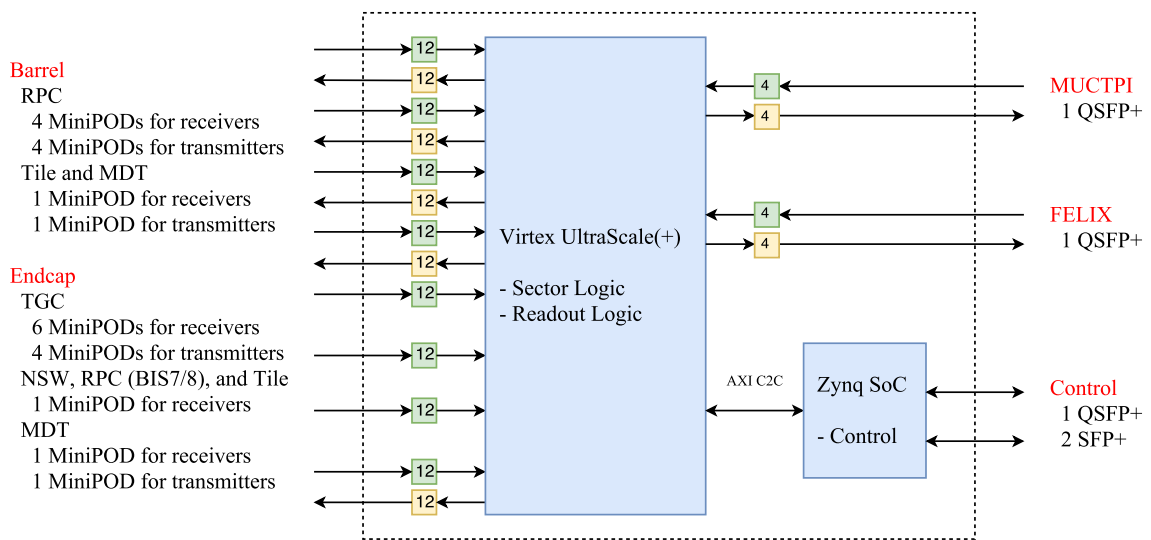


Figure 8.20: Block diagram of the off-detector board in which the Sector Logic is implemented. The readout logic and the control functions are also implemented. The blocks with the number “12” indicate the MiniPOD modules [8.10], each of which contains 12 channels. Two blocks with the number “4” constitute a QSFDP+ module [8.11], which is used for four transmitters and four receivers. The board is connected to the electronics of various detectors employed for the Level-0 muon trigger. There are connections with Level-0 MUCTPI and FELIX. The board is controlled with a Xilinx Zynq® device.

## 8.4 Sector Logic

---

Table 8.4: Number of receivers and transmitters per blade depending on the connected elements for the barrel off-detector boards.

Connected elements	Number of receivers	Number of transmitters
RPC on-detector boards	36	36
Tile calorimeter boards	3	0
MDT Trigger Processor	6	10
Level-0 MUCTPI	0	2
FELIX	1	2

Table 8.5: Number of receivers and transmitters per blade depending on the connected elements for the endcap off-detector boards.

Connected elements	Number of receivers	Number of transmitters
TGC on-detector boards	67	35
NSW Trigger Processor	6	0
RPC (BIS78) boards	2	0
Tile calorimeter boards	2	0
MDT Trigger Processor	6	10
Level-0 MUCTPI	0	3
FELIX	1	4

### 8.4.3 Firmware Design

The firmware of the FPGA on the off-detector board includes the Sector Logic, the readout logic, and the control functions (Fig. 8.21). In the Sector Logic, the trigger scheme described in Section 8.4.1 is implemented. Primitive track candidates are identified from the signals of RPC and TGC for the barrel and the endcap, respectively. The Sector Logic makes a spatial and time coincidence, with a time resolution of one bunch crossing or better, as in the case of the barrel system. Additional coincidence is taken with other inner detectors. The Sector Logic is designed to simultaneously process multiple closeby muons to retain the acceptance of the multi-muon trigger. The muon track candidates are transmitted to the MDT Trigger Processor. The Sector Logic receives the muon track candidates confirmed by the MDT Trigger Processor, and transmits the list of candidates to the Level-0 MUCTPI.

The RPC and TGC hit signals are separated before the Sector Logic, and transferred to the readout logic. The trigger information from the Sector Logic, including the NSW and Tile calorimeter data with limited granularity, is also transferred to the readout logic. The readout functionality for the full-granularity data of the NSW and the Tile calorimeter are provided separately from the off-detector boards where the Sector Logic is implemented. In

the readout logic, the data are stored in a derandomiser buffer during the Level-0 decision time, and transmitted to **FELIX** after the Level-0 trigger accept signals are received. The control of the **RPC** and **TGC** on-detector boards are originating from Xilinx **Zynq®** device.

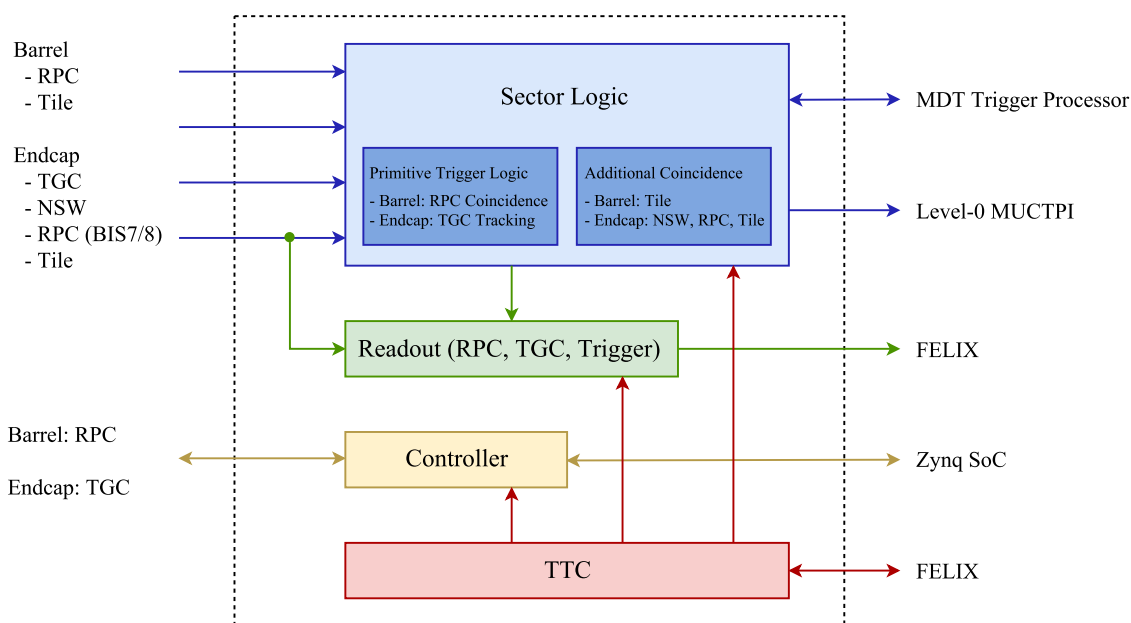


Figure 8.21: A simplified block diagram of the proposed firmware for the **FPGA** of the off-detector boards. The detector signals are transferred to the Sector Logic (blue), and muon track candidates are provided. The muon track candidates are transmitted to the **MDT Trigger Processor**. The Sector Logic receives the muon track candidates from the **MDT Trigger Processor**, and transmit to the Level-0 **MUCTPI**. The readout logic (green) handles the readout for **RPC** and **TGC** hit signals and trigger information. The **TTC** manager (red) and the control (yellow) blocks manage the control of the off-detector board and **RPC** and **TGC** on-detector board.

#### 8.4.4 Detector Signal Inputs

**Inputs from RPC** The hit data and time information for the **RPCs** in the barrel inner, middle, and outer stations are provided to the Barrel Sector Logic for each bunch crossing. The size of the transferred data is reduced by encoding and zero suppression. The optical links with a bandwidth of 9.6 Gbps per link are assumed, which manage the data transfer for possible highest hit rate evaluated from the actual measurement in Run 2 [8.1].

**Inputs from TGC** The hit data of the **TGCs** in the endcap inner station in  $1.05 < |\eta| < 1.3$  and the endcap middle station are provided to the Endcap Sector Logic for each bunch crossing. The hit data comprise one bit per channel per bunch crossing as an identifier of the existence of a hit. The data are transferred without zero suppression. The optical links with a bandwidth of 8.0 Gbps per link are assumed, which manage the data transfer independently of the hit rate.

**Inputs from NSW** The **NSW** Trigger Processor is designed to reconstruct the track segments with an angular resolution of 1 mrad [8.2]. The position and the angle of the segments are provided to the Endcap Sector Logic for each bunch crossing. Further explanation is given in Section 8.5.6. The segmentation of the **NSW** Trigger Processor is different from that of the Sector Logic, and hence the output of single **NSW** Trigger Processor is divided and transferred to multiple off-detector boards where the Sector Logic is implemented (Fig. 8.19).

**Inputs from Tile Calorimeter** The **Tile** calorimeter system is divided into three layers in depth. The last layer, known as the D-layer, is used for the Level-0 muon trigger, since most of the charged particles other than muons stop before this layer and only muons deposit energy consistent with minimum ionising particles. The energy deposit is estimated using an optimal filter algorithm. For each calorimeter cell, an identifier of the energy deposit to be greater than the threshold or not is passed to the Endcap Sector Logic for each bunch crossing.

### 8.4.5 Realtime Output Data Format

The data transfer between the Sector Logic, the **MDT** Trigger Processor, and the Level-0 **MUCTPI** is based on optical links. A maximum transfer rate of 9.6 Gb/s is assumed for each link, where six 32-bit words are transferred every 25 ns (Table 8.6). The data are encoded with a 8b10b scheme. The data are transferred regardless of the results containing muon candidates or not. The first and the last words are the header and trailer words. The contents of the header and trailer words are shown in Tables 8.7 and 8.8. The data transfer of two muon candidates is possible per link, with maximum 64 bits per candidate. Each barrel (endcap) Sector Logic can transfer 4 (6) muon candidates to the Level-0 **MUCTPI**.

Table 8.6: Contents of the six 32-bit words transferred by an optical link every 25 ns.

32-bit Word Identifier	0	1	2	3	4	5
Content	Header	Muon candidate 1		Muon candidate 2		Trailer



Table 8.7: Contents of the 32-bit header word.

Number of bits	Name	Comment
12	BCID	Identifier of the bunch crossing
20	Reserved	Reserved bits

Table 8.8: Contents of the 32-bit trailer word.

Number of bits	Name	Comment
8	Comma	K28.5, K character of the 8b10b encoding
6	Board ID	Identifier of the boards
4	Fibre ID	Identifier of the fibres
8	Parity	Parity bits, one bit corresponds to 16 bits in the data words
6	Reserved	Reserved bits

**Barrel Sector Logic to MDT Trigger Processor** One 64-bit word is used for a muon candidate selected by the Barrel Sector Logic. It includes the information of the muon track candidate as well as the RPC hit positions and the coincidence type. The contents of the data words are shown in Table 8.9.

Table 8.9: Contents of the 64-bit word for the data from the Barrel Sector Logic to the MDT Trigger Processor.

Number of bits	Name	Comment
10	Position $\eta^{\text{RPC0}}$	$\eta$ coordinate of the position on RPC0
10	Position $\eta^{\text{RPC1}}$	$\eta$ coordinate of the position on RPC1
10	Position $\eta^{\text{RPC2}}$	$\eta$ coordinate of the position on RPC2
10	Position $\eta^{\text{RPC3}}$	$\eta$ coordinate of the position on RPC3
6	Position $\phi$	$\phi$ coordinate of the position on the innermost layer used for the coincidence (inner or middle station)
2	Coincidence type	Identifier of the types of coincidence
4	$p_T$ threshold	Highest transverse momentum threshold satisfied
1	Charge	Muon candidate charge
11	Reserved	Reserved bits

## 8.5 NSW Trigger Processor

**Endcap Sector Logic to MDT Trigger Processor** One 64-bit word is used for a muon candidate selected by the Endcap Sector Logic. It includes the information of the muon track candidate and the track segment reconstructed by the TGC hits. It also includes the track segment reconstructed in the NSW Trigger Processor, which is used for the  $p_T$  evaluation in the MDT Trigger Processor. The contents of the data words are shown in Table 8.10.

Table 8.10: Contents of the 64-bit word for the data from the Endcap Sector Logic to the MDT Trigger Processor.

Number of bits	Name	Comment
10	Position $\eta$	$\eta$ coordinate of the position on the outermost TGC layer
6	Position $\phi$	$\phi$ coordinate of the position on the outermost TGC layer
6	Angle $\Delta\theta$	Segment polar angle with respect to the vector from the interaction point to the segment position, the last bit assigned for the muon candidate charge
4	Angle $\Delta\phi$	Segment azimuthal angle with respect to the vector from the interaction point to the segment position, the last bit assigned for the sign
4	$p_T$ threshold	Highest transverse momentum threshold satisfied
28	NSW segments	Identical to the NSW Trigger Processor output
3	Quality	Coincidence pattern (NSW, BIS78 RPC, etc.)
3	Reserved	Reserved bits

**Sector Logic to Level-0 MUCTPI** One 48-bit word is used for a muon candidate. The contents of the data words are shown in Table 8.11. The reserved bits can be used for the trigger information for exotic signatures.

## 8.5 NSW Trigger Processor

This section describes the motivation, the expected performance, the hardware and firmware design, and the interfaces of the NSW Trigger Processor upgrade.

### 8.5.1 Motivation for the Upgrade

The NSW project [8.2] is a Phase-I muon upgrade project slated for the installation during Long Shutdown 2. In the Phase-I system, signals from sTGC and MM are used separately

Table 8.11: Contents of the 48-bit word for the data from the Sector Logic to the Level-0 *MUCTPI*.

Number of bits	Name	Comment
14	Position $\eta$	$\eta$ coordinate of the segment position on the innermost <i>MDT</i> layer
6	Position $\phi$	$\phi$ coordinate of the position
8	$p_T$	Muon candidate transverse momentum
4	$p_T$ threshold	Highest transverse momentum threshold satisfied
1	<i>MDT</i> confirmation	Reconstructed by <i>MDT</i> or not
1	Charge	Muon candidate charge
2	0/1/2/3 stations	Number of <i>MDT</i> segments for the candidate
3	Quality	Average or worse segment quality
9	Reserved	Reserved bits

to form *NSW* track segments in the trigger processor electronics. A requirement is applied on the pointing of trigger segments back to the interaction region, and segment position information is transmitted to the Sector Logic, where a coincidence with *TGC* is taken. The latency budget for the Phase-I *NSW* Trigger Processor is extremely tight. The *NSW* trigger as designed for the installation after Run 2 already meets the Phase-II requirement of an angular resolution of 1 mrad. However, it is quite possible to further lower the thresholds for muon momentum and increase the efficiency by taking advantage of the increased latency to do a more refined calculation of muon pointing and momentum in the *NSW*, or to improve robustness and redundancy, for example in case of missing layers.

In the Phase-II Level-0 muon trigger, the *NSW* trigger segment reconstructed in the *NSW* Trigger Processor is used to determine not only the deflection angle with *TGC* in the Sector Logic but also the deflection angle and the sagitta (see Section 8.6.1) with *MDT* in the *MDT* Trigger Processor for a refined momentum resolution. Even though the resolution of the deflection angle measured with the *NSW* and *TGC* segments is dominated by the *TGC*, the deflection measurement using the more precise *MDT* information would benefit from a *NSW* segment with high precision. A combined segment fit with increased precision is possible if both *sTGC* and *MM* hits are combined into a single fit.

In the Phase-I *NSW* Trigger Processor, the *sTGC* and *MM* segments are derived separately and a procedure to remove duplicates is implemented before sending the segment information to the Sector Logic. A coincidence of hits is required in at least three out of four layers for each of the *sTGC* and *MM* quadruplets before fitting a segment. In the upgraded electronics, all *sTGC* and *MM* layers are available in the same trigger processor *FPGA*, allowing for improved background rejection and increased efficiency due to more flexible coincidence requirements.

Historically, triggering on forward muons can be challenging due to backgrounds from secondaries that can be difficult to eliminate and assess until the NSW detector is actually running. The trigger performance can be enhanced by improved granularity that can be achieved with more latency and increased resources.

### 8.5.2 Performance of the Segment Reconstruction

The performance of the track segment reconstruction by NSW could be affected by the background hits. Figure 8.22 shows the impact of large cavern background rates on the azimuthal precision of the MM-only segment. The current algorithm performance will be degraded, but the performance can be regained with a more sophisticated algorithm that takes more FPGA resources and more latency. This improvement in the azimuthal resolution will permit a finer tuned matching between segments in the TGC and the NSW.

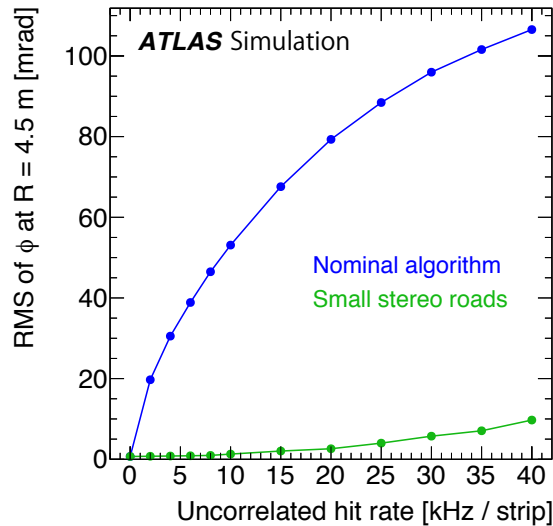


Figure 8.22: Azimuthal angle ( $\phi$ ) precision for different background rates for the standard algorithm (blue) and an improved algorithm (green) that uses higher granularity but requires longer latency and more FPGA resources.

### 8.5.3 Hardware Design

In this proposed upgrade, the Phase-I trigger processor hardware (ATCA mezzanines and blades), excluding the ATCA crates and optical fibres, are replaced. Figure 8.23 shows a diagram of the proposed hardware. The new NSW Trigger Processor is designed to receive both sTGC and MM inputs from a  $1/16^{\text{th}}$   $\phi$  sector in a single FPGA. Two sectors are

processed in each board, and the full system requires 16 ATCA boards to operate. In the proposed implementation, the trigger processor firmware is implemented in an FPGA on a mezzanine card of an ATCA carrier blade. The input fibres from the NSW detectors and the output fibres to the Sector Logic connect directly to the mezzanine card. Some services are implemented on the carrier card, which adhere to ATCA standards and provide all required interfaces to meet the specifications for Phase-II.

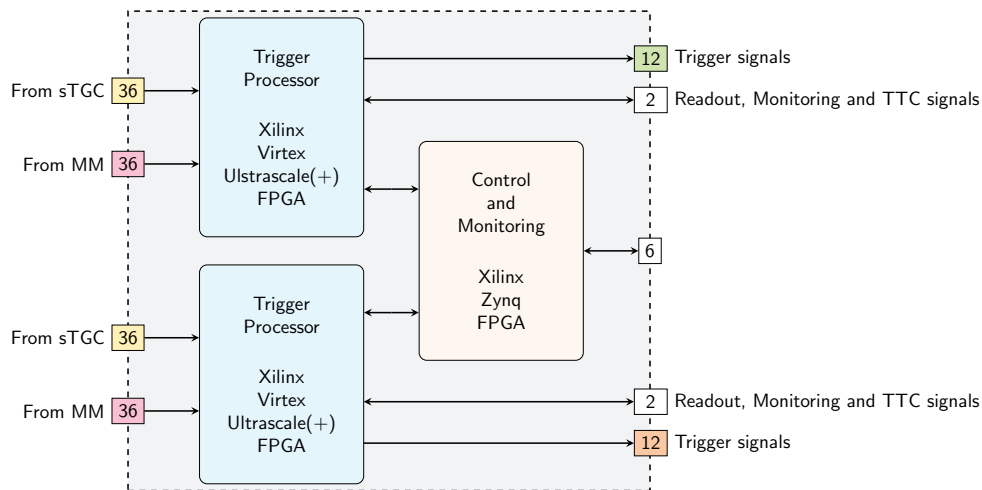


Figure 8.23: Block diagram of the off-detector board in which the NSW Trigger Processor is implemented. This ATCA blade includes two FPGAs that each process a  $1/16^{\text{th}}$   $\phi$  sector of the detector. There are connections with the Sector Logic boards and FELIX/TTC. The board is controlled with a Xilinx Zynq® device.

The trigger processor FPGA is required to have at least 72 high speed optical receivers for the NSW detector signals (36 for sTGC and 36 for MM), 12 transmitters to send the track segments to the Sector Logic, as well as additional receivers and transmitters for control signals. In particular, one additional receiver and one transmitter are used for the connection with FELIX to receive the TTC signals and to transmit the readout and monitoring signals (signal readout, statistics, sampled events, algorithm parameters, etc).

The Xilinx Virtex UltraScale+ family of FPGAs is proposed for the implementation, although the final decision will depend on the evolution of the technology in industry. The UltraScale+ FPGA includes devices that can process all NSW detector layers (sTGC and MM).

### 8.5.4 Firmware Design

The firmware includes the trigger processor and the control functions. Figure 8.24 shows a simplified block diagram of the proposed firmware. In the trigger processor, the sTGC

## 8.5 NSW Trigger Processor

and **MM** inputs are processed and reformatted into a position on each of the 16 detector planes. In the case of the **MM**, the position of the strip with the first hit in time within each group of 64 strips is used (Address in Real Time or ART, data). In the case of the **sTGC**, a cluster centroid is calculated using the charge information in neighbouring strips. Coincidences are then taken with the **sTGC** and **MM** signals from the different planes to suppress fake triggers. The **sTGC** and **MM** information is finally combined in a fit that is used to derive the segment parameters to be transmitted to the Sector Logic. Additional firmware blocks implement functionality supporting the **NSW** Trigger Processor algorithm such as the monitoring, the readout, the **TTC** signal handling, the configuration, control and monitoring of the hardware (including the **FPGAs**), etc.

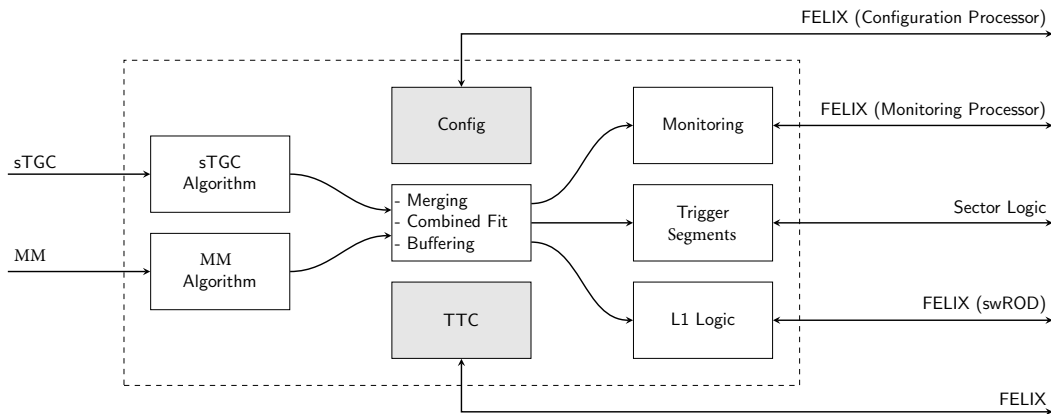


Figure 8.24: A simplified block diagram of the proposed firmware for the **NSW** Trigger Processor off-detector boards, indicating the input and output connections.

### 8.5.5 Detector Signal Inputs

The detector signal inputs are identical to those in the Phase-I **NSW** Trigger Processor.

**Inputs from sTGC** The **sTGC** strip information identified from up to four pad-trigger towers are provided to the trigger processor for each bunch crossing. The radial and azimuthal information of the pad-trigger roads are also provided, and the pad-trigger provides the azimuthal measurement for the **sTGC** inputs. Charge thresholds are applied to select up to five strips in a band and a layer centroid calculation is performed to derive the position coordinates for each **sTGC** plane.

**Inputs from MM** The trigger input data from the **MM** detector provides the address of the strip that first crosses the threshold. The address of the  $\sim 450 \mu\text{m}$  strip is used to derive

the **MM** hit position per layer. Stereo strips are used in the middle layers to derive the azimuthal information.

### 8.5.6 Realtime Output Data Format

Each track segment reconstructed by **NSW** Trigger Processors will be presented as 28-bit data with the format shown in Table 8.12. This format allocates more bits to the  $\eta$  and  $\phi$  information compared to the format used in Run 3 to account for the increased precision from a combined **sTGC** and **MM** fit.

The required resolutions (1 bit) are approximately 1 mrad ( $\pm 15$  mrad full scale) for  $\Delta\theta$ , 15 mrad for  $\phi$ , and 0.0001 in pseudo-rapidity  $\eta$ . The  $\Delta\theta$ ,  $\phi$ , and  $\eta$  position information are included in the output, in addition to one flag. The monitor flag is set when this bunch crossing has been identified for further monitoring by the **NSW** trigger.

Table 8.12: Contents of the data from the **NSW** Trigger Processor to the Endcap Sector Logic.

Number of bits	Name	Comment
14	Position $\eta$	$\eta$ coordinate of the <b>NSW</b> segment position
8	Position $\phi$	$\phi$ coordinate of the <b>NSW</b> segment position
5	$\Delta\theta$	Segment polar angle with respect to the vector from the interaction point to the segment position
1	Monitor	Monitoring bit

## 8.6 MDT Trigger Processor

The **MDT** chambers currently used for precision tracking will be included in the Level-0 muon trigger to improve the momentum resolution and the redundancy. This section describes the motivation, the expected performance, the hardware and firmware design, and the interfaces of the **MDT** Trigger Processor. The section also introduces the readout path for the **MDT** chambers and the resource estimation.

### 8.6.1 Motivation for the Implementation

The selectivity of the current Level-1 muon trigger is limited by the moderate spatial resolution of **RPC** and **TGC** (Fig. 8.3). It has been considered since the start of the R&D for **HL-LHC** to use the **MDT** chamber data in the Phase-II Level-0 muon trigger [8.12][8.13][8.14]

[8.15]. The **MDT** chambers provide a better spacial resolution than **RPC** and **TGC**, and a  $p_T$  resolution close to that of the offline reconstruction can be achieved.

The use of the **MDT** chambers improves the robustness of the Level-0 muon trigger. The rate of the accidental coincidences in the BI-BO trigger in the barrel is expected to have nonlinearity with the luminosity. This is suppressed by the requirement based on **MDT**, mitigating the risks on the unexpected rate increase.

Studies of the **MDT** trigger concept show that it is not needed to reconstruct the full muon track inside the muon spectrometer, but that it is sufficient to use the positions and the angles of track segments reconstructed in the muon chambers for a refined momentum measurement. This concept makes it possible to implement the hardware-based **MDT** trigger in the Phase-II Level-0 trigger.

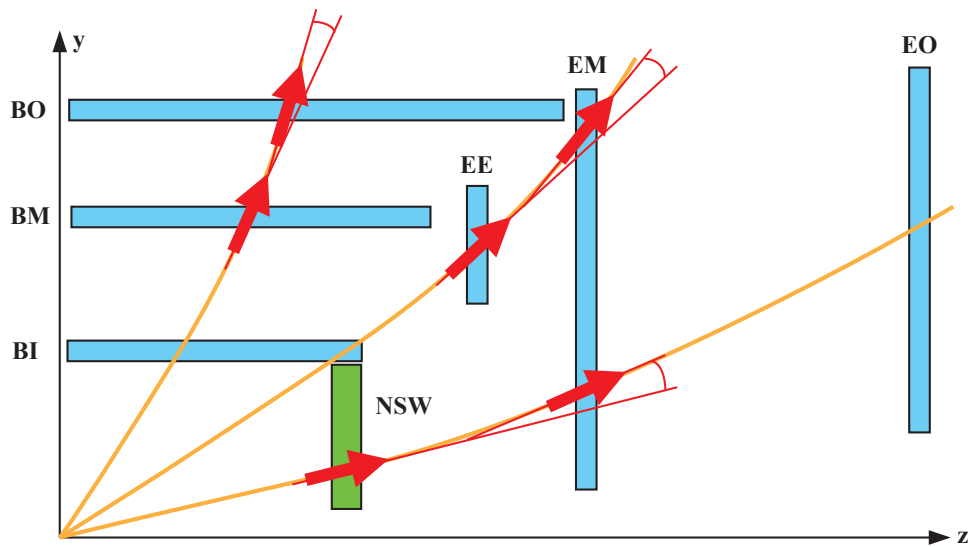
Figure 8.25 shows schematically two variables that can be used to measure the muon  $p_T$  in the **MDT** Trigger Processor. The first variable is the polar-angle difference between track segments in two chambers, referred to as the deflection angle hereafter. The second variable uses track segments in three chambers and is defined as the distance between the position of one segment from the straight line joining the other two, referred to as the “sagitta” hereafter. The two variables provide different rate rejections and coverages (see Sections 8.6.2 and 8.6.6), and the best performance is achieved by the combination.

### 8.6.2 Trigger Performance

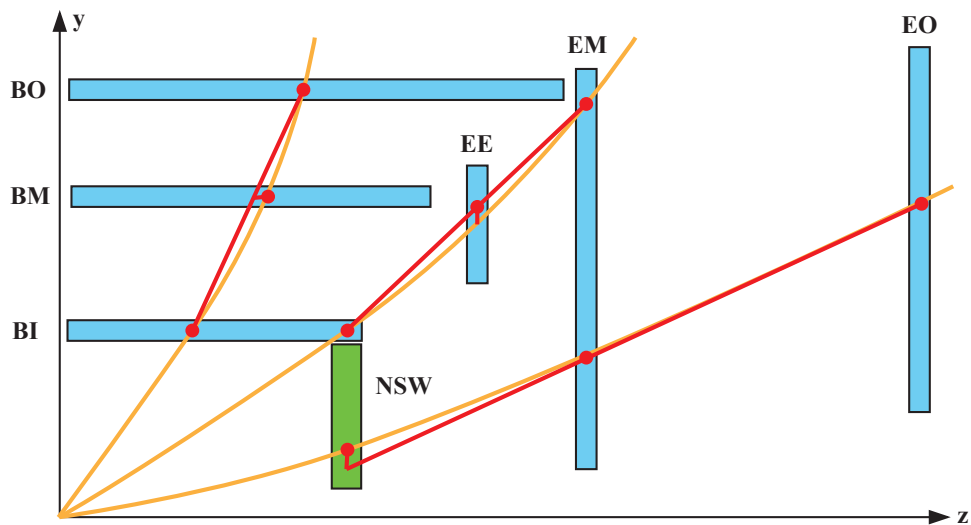
The performance of the **MDT** trigger has been studied with **MC** samples after the requirements expected for the Sector Logic. For the barrel, the **RPC** trigger scheme of “3/4 chambers + BI-BO” is employed. For the endcap, the **TGC** segment is reconstructed by a minimum  $\chi^2$  method. The track segments reconstructed by the ATLAS offline software are used to obtain the deflection angle and the sagitta for the **MDT** trigger.

The trigger efficiency has been studied with **MC** samples where a muon is produced in each event. The muon direction is random in the  $\eta$ - $\phi$  plane. No pile-up is involved. Figure 8.26 shows the efficiency for 20 GeV threshold as a function of the muon  $p_T$  evaluated by the ATLAS offline tool. Figure 8.27 shows the relative efficiency of the **MDT** trigger with respect to the **RPC** and **TGC** trigger. The **MDT** trigger provides an improved selectivity for the muons with  $p_T$  around the threshold, while keeping a high plateau efficiency. The same **MC** samples as in Section 8.4.1 are used for the study of the trigger rate of the **MDT** trigger in the barrel. Figure 8.28 shows the estimated Level-0 single muon rate depending on the luminosity for  $p_T$  thresholds of 10 GeV and 20 GeV. Figure 8.29 shows the rates as a function of  $|\eta|$  for the case of  $\langle\mu\rangle = 200$ . The distribution of offline reconstructed muon matching the **RPC** trigger with offline  $p_T$  above the Level-0 threshold is shown for comparison. For the sample with  $\langle\mu\rangle = 200$ , the **MDT** trigger reduces the Level-0 **RPC** trigger by approximately 55% and 70% for  $p_T$  thresholds of 10 GeV and 20 GeV, respectively. For the endcap, the trigger rate has been studied with Run 1 and Run 2 data samples, and the reduction by





(a) Two-stations: deflection angle



(b) Three-stations: sagitta

Figure 8.25: (a) Schematic of the concept of the polar angle difference between the two segments. The combination considered for the barrel is the middle and outer stations, but any other combination can in principle be used as well. (b) Schematic of the concept of the distance between the position of a segment and the position extrapolated from two segments. The figures show the cross-sectional view for so-called even sectors. The variable is defined similarly in so-called odd sectors. For  $1.05 < |\eta| < 1.3$ , the endcap inner station is assumed alternatively to the barrel inner station.

## 8.6 MDT Trigger Processor

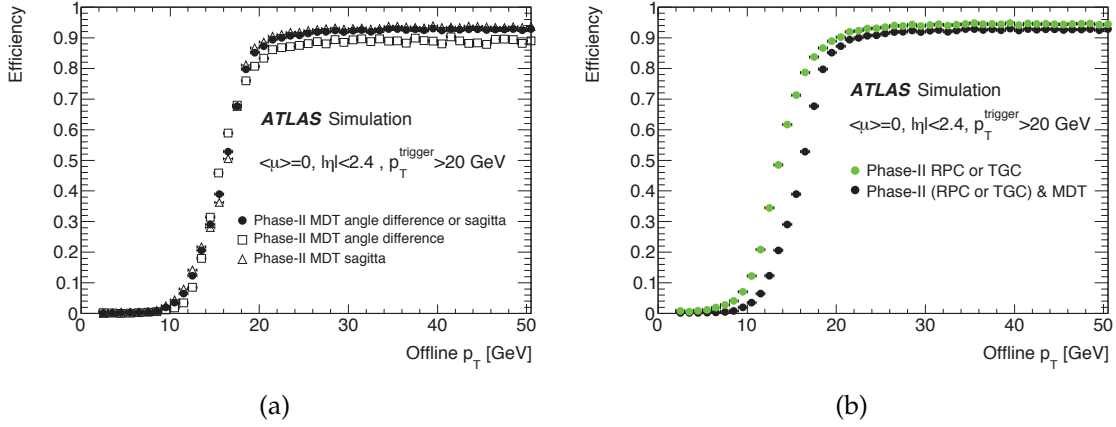


Figure 8.26: Trigger efficiency in  $|\eta| < 2.4$  as a function of  $p_T$  measured by the offline reconstruction. A  $p_T$  threshold of 20 GeV is assumed. The values are estimated with single-muon MC with no pile-up. Figure (a) shows the efficiency for the angle difference, the sagitta, and the combination. Figure (b) shows the efficiency with and without the MDT.

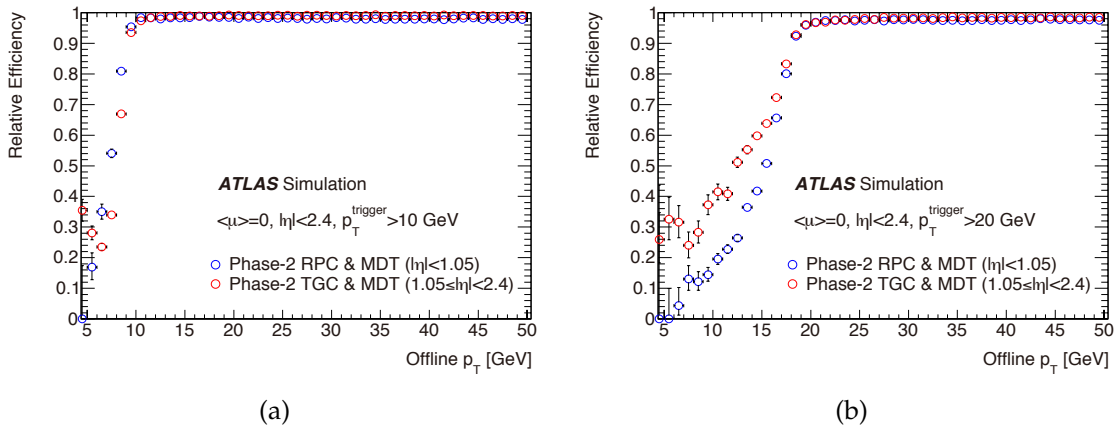


Figure 8.27: Relative efficiency of the MDT trigger with respect to the RPC and TGC trigger as a function of  $p_T$  measured by the offline reconstruction. The values are estimated with single-muon MC with no pile-up. The range  $p_T < 4$  GeV is not shown because of the large uncertainty due to low statistics. Figures (a) and (b) for  $p_T$  thresholds of 10 GeV and 20 GeV, respectively.

the **MDT** trigger with respect to the rate of **TGC** and **NSW** is found to be roughly 50% for  $p_T$  threshold of 20 GeV [8.1].

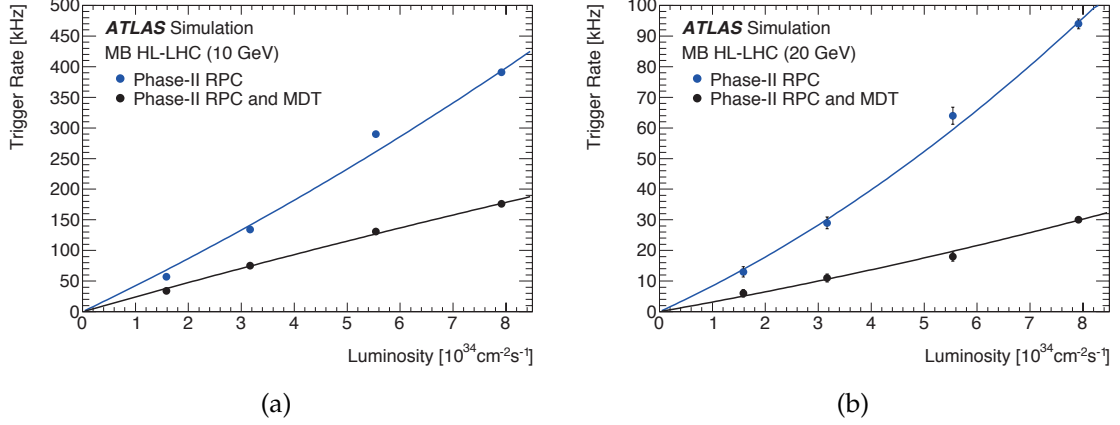


Figure 8.28: The estimated rate of the Level-0 single-muon trigger based on **RPC** only and **RPC** plus **MDT** for the barrel region  $|\eta| < 1.05$ . The dots with error bars show the values estimated from **MC** samples with the  $\langle \mu \rangle$  values corresponding to various luminosity scenarios. The curves show the result of the fit by a second-order polynomial.

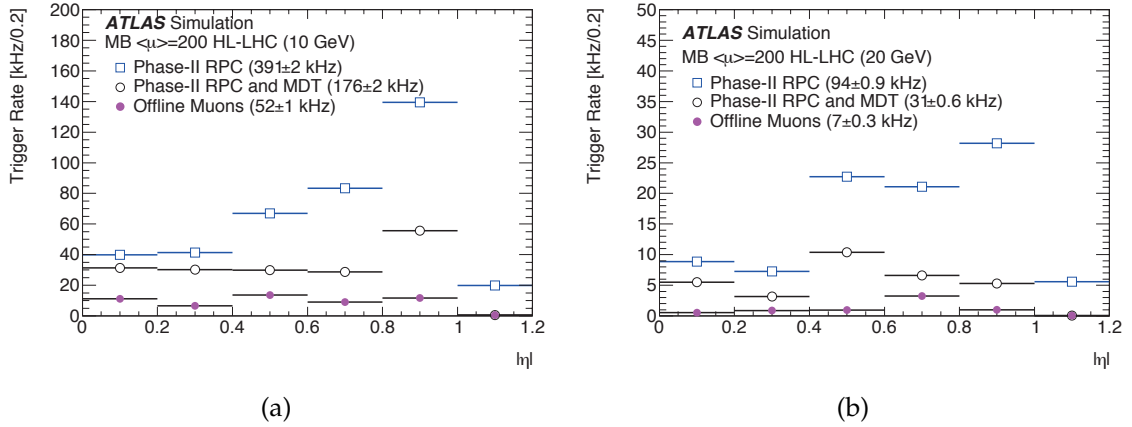


Figure 8.29: The estimated rate of the Level-0 single-muon trigger based on **RPC** only and **RPC** plus **MDT** for the barrel region  $|\eta| < 1.05$  in **MC** sample with  $\langle \mu \rangle = 200$ . The full dot distribution is obtained by requiring that the **RPC** trigger matches an offline reconstructed muon with  $p_T > 10$  (20) GeV.

As the deflection of the muon trajectory is greater for lower momentum, the Sector Logic already has a good selectivity and the relative reduction by the **MDT** trigger per muon is smaller for lower momentum. However, the lower  $p_T$  thresholds will be used mainly for multi-muon triggers, and the rate reduction by the **MDT** trigger will be multiplied. The

10 GeV threshold will be used mainly for dimuon trigger, and the rate reduction would be approximately twice as large as that for single muon for the same threshold.

### 8.6.3 Hardware Overview

The MDT chambers are divided in  $\phi$  into 16 sectors. They are further partitioned between barrel and endcap subsystems, A and C sides (divided at  $\eta = 0$ ), and three stations (inner, middle, and outer). The MDT trigger system is partitioned into 64 processing elements, each handling one sector, one side (A/C), and one subsystem (barrel/endcap).

Figure 8.30 illustrates an ATCA blade which handles one out of the 64 sectors. The blade contains a total of 144 fibre receivers and transmitters, with the exact complement of each to be determined. The fibre transmitters may be located on a rear transition module. The ATCA services are provided by a CERN-designed IPMC mezzanine. This device controls and monitors operation of the module, including voltage, current, and temperature monitoring. Access to the IPMC is provided through the Intelligent Platform Management Bus (IPMB) on the ATCA backplane and Gigabit Ethernet.

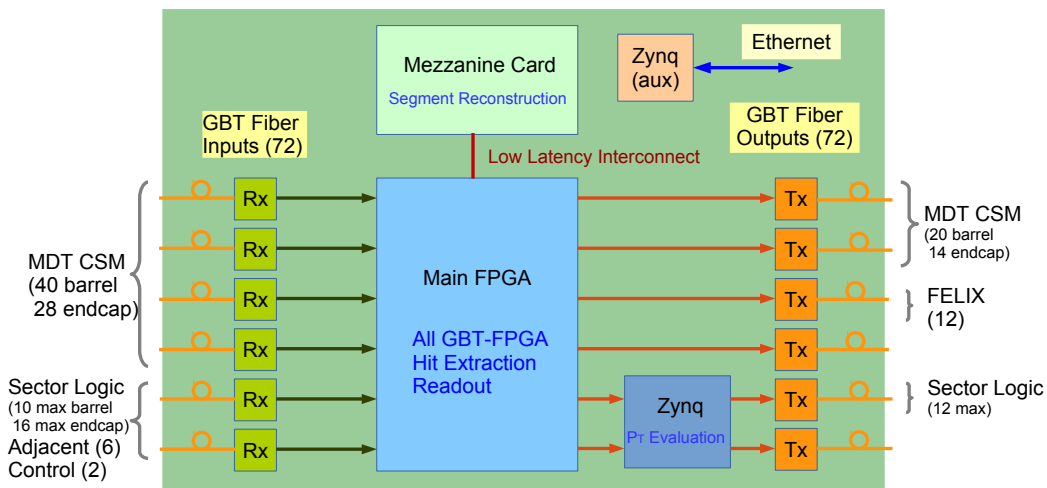


Figure 8.30: Layout of the ATCA blade where the MDT Trigger Processor is implemented.

The blade contains one Xilinx UltraScale FPGA, which handles receipt of the MDT hits, receipt of the Sector Logic track candidates, hit extraction, calibration, transmission of hits to the mezzanine boards, receipt of the segment data from the mezzanine boards, evaluation of  $p_T$ , and transmission of the muon track candidates to the Sector Logic. Each mezzanine board performs a segment reconstruction. The mezzanine boards are introduced in the design so that the segment reconstruction gets separate from the other functions and the task sharing of the development gets easier. A modular design allows to confine losses to

a given item during the production and maintenance phases. The blade [FPGA](#) also transfers the [MDT](#) hits to [FELIX](#). Connections are planned to receive and transmit control signal such as [TTC](#). The transmitters to the [Chamber Service Module \(CSM\)](#) send the [TTC](#) stream and the configuration to the electronics. An external DDR memory device is required for buffering of the data for readout.

The Xilinx [Zynq®](#) devices are anticipated for the system control functions including detector controls, monitoring, and local data acquisition. The Xilinx [Zynq®](#) device requires flash memory for programme storage, RAM, and an interface to a MicroSD card for removable storage. A Gigabit Ethernet switch provides access to the Xilinx [Zynq®](#) devices from the base fabric on the [ATCA](#) backplane, a front-panel RJ-45 connector or the [IPMC](#).

A block diagram of the trigger logic is shown in Fig. 8.31. Raw [MDT](#) hits are received on three groups of [Gigabit Bidirectional Trigger and Data Link \(GBT\)](#) (or [lpGBT](#)) links from the inner, middle, and outer [MDT](#) stations. Since some [MDT](#) hits can arrive before the Sector Logic track candidates, they will need to be buffered. [MDT](#) hits can be calibrated where the  $(r-t)$  relation is stored in [LUT](#). [MDT](#) hits in the [RoI](#) derived from the Sector Logic data are extracted by the hit extractor. Track segment is reconstructed and used for  $p_T$  evaluation. Muon track candidates as well as the evaluated  $p_T$  are transmitted to the Sector Logic. Additional links between adjacent sector [MDT](#) Trigger Processor blade are foreseen to pass segments found between sector so that to handle sector overlap as well as the transition region where muon can have segments from the barrel and the endcap.

An [MDT](#) Trigger Processor implements three separate processing pipelines for the hit extraction and the segment reconstruction. The processing for three simultaneous [RoIs](#) is possible. The estimated busy fraction is 0.1% (0.5%) for an [RoI](#) rate of 100 kHz (180 kHz) per [MDT](#) Trigger Processor. The [MDT](#) Trigger Processor is designed to be used for  $p_T$  thresholds down to around 5 GeV.

### 8.6.4 Hit Extraction

The hit extractor logic (Fig. 8.32) matches [MDT](#) hits with [RoIs](#) in space and time for further processing of the segment reconstruction. The “[RoI](#) processor” converts the [RoI](#) information from global coordinates to a range of tube ID numbers, which are passed to the “tube range [LUT](#)”, and passes a time and position to the “[MDT](#) hit calibration” block. The block “tube range [LUT](#)” converts the [MDT](#) ID number in the [TDC](#) data to a tube/layer number format.

In the barrel, due to deflection by the magnetic field, the [RoI](#) directions and positions, at the centre of the [MDT](#) chamber spacer frame, need to be determined separately for the inner, middle, and outer [MDT](#) precision chambers, as shown in Fig. 8.33a. The algorithm handles the different [RPC](#) coincidence requirements leading to distinct hit patterns. For example, in the case where a layer does not have [RPC](#) hits, the [RoI](#) position and direction at this layer can be determined from the extrapolation of the [RoI](#) positions and angles in the

## 8.6 MDT Trigger Processor

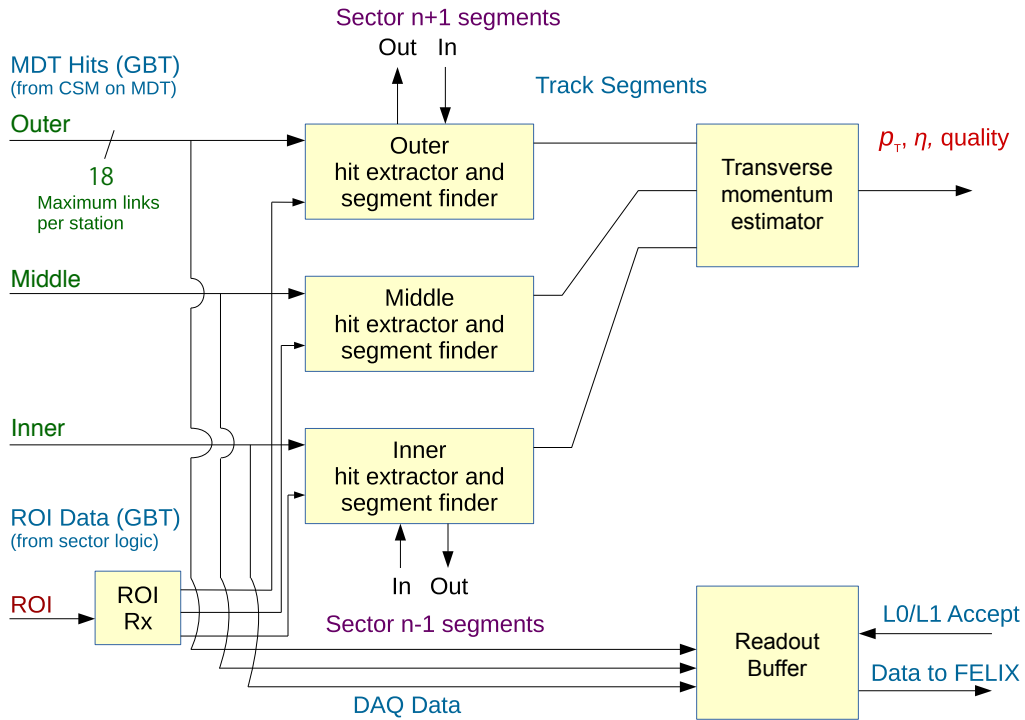


Figure 8.31: Block diagram of the *MDT* Trigger Processor.

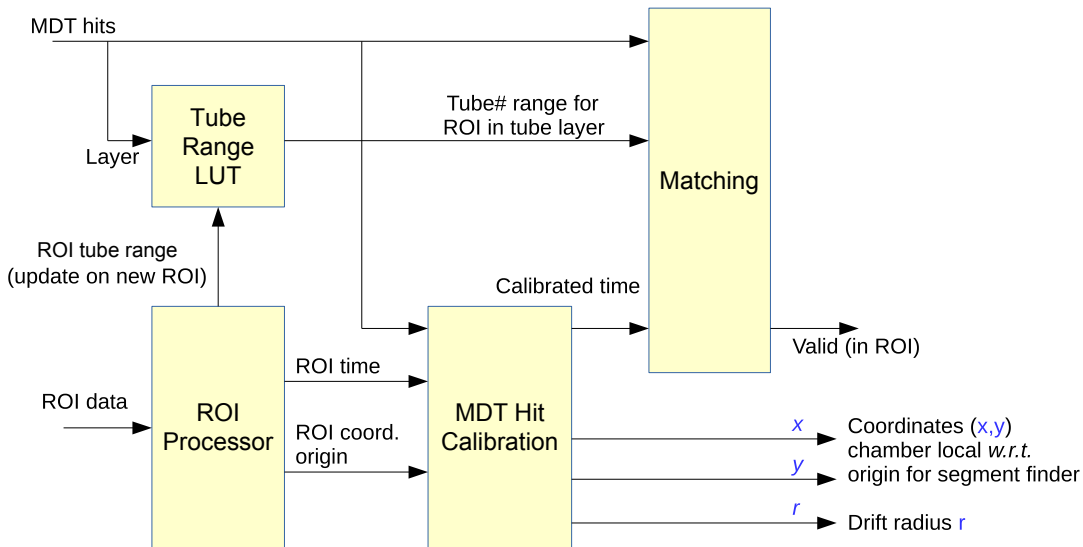


Figure 8.32: Block diagram of the hit extractor. The *MDT* hits are matched with the *RoIs* derived from the Sector Logic track candidates, and calibrated data are transferred to the following logic.

neighbouring layers, as seen in Fig. 8.33b. In the endcap, the **RoI** direction and position is determined from the **TGC** track segment information. Since there is no significant magnetic field between the endcap middle and outer stations, the direction of the **RoI** is defined to be the same for these stations. The position of the **RoI** in the middle of the spacer frame of these stations is determined by a linear extrapolation for the **TGC** track segment. For each station, the matching window is defined as a parallelogram centred on the **RoI** position as depicted in Fig. 8.33c.

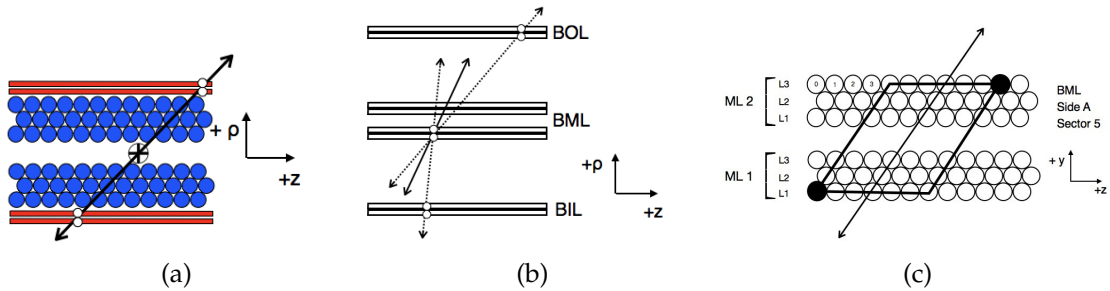


Figure 8.33: (a) Cross sectional view of the position and the direction reconstructed in a barrel middle station. The red rectangles represent RPC1 and RPC2. The blue circles represent the **MDT** tubes while the white circles represent **RPC**  $\eta$ -hits. The double-sided arrow indicates the fit line with the plus sign marking its intersection with the centre of the spacer frame. (b) Cross-sectional view of the averaging used in cases where there are  $\eta$ -hits in BI and BO but only one **RPC** layer of BM. The long rectangles represent **RPC** layers and the circles represent **RPC**  $\eta$ -hits. The dotted lines indicate the fits used to determine the position, which is shown as a solid line. (c) Illustration of an **RPC** **RoI** window definition in terms of **MDT** tube identification. The solid line represents an **RoI** and the parallelogram represents the window opened around it with a size of  $\pm 3$  **MDT** tubes. The solid circles represent the pair of **MDT** tubes used to identify the matching window. The numbers indicate the multilayer (ML), layer (L), and tube IDs used for each **MDT** chamber.

Studies have shown that a window width of seven **MDT** tubes for the inner and middle stations, and a wider window of 13 tubes for the outer stations, to account for the inaccuracy in the **RoI** extrapolation to the outer layers due to residual magnetic field, is sufficient to achieve matching efficiencies better than 95% for hits originating from real muons. The matching of the **MDT** drift time to the bunch crossing identifier is expected to reject at least  $\sim 50\%$  of background hits originating from pile-up and cavern background. With both the time and space matching applied, only  $\sim 3\%$  of the hits within an **RoI**, corresponding to a maximum of 100 hits per station, remain to perform the segment reconstruction.

The **TDC** value,  $t_{\text{TDC}}$ , recorded by the frontend electronics and transmitted to the **MDT** Trigger Processor, is the sum of the delays in the readout electronics ( $t_0$ ), the muon time of flight from the bunch crossing to the **MDT** chamber ( $t_{\text{TOF}}$ ), the drift time ( $t_{\text{Drift}}$ ) from the ionisation, and other smaller corrections. The  $t_0$  and  $t_{\text{TOF}}$  are the largest corrections, and a single pre-calculated time offset per **MDT** station is sufficient to extract **MDT** hit drift times.

The MDT hit drift time is converted into drift radius ( $r_{\text{Drift}}$ ) using a space-to-drift time relationship, also called ( $r-t$ ) relation. The ( $r-t$ ) relation is nonlinear in the MDT [8.16] due to the tube geometry and the non-uniform electric field inside the tube. As the muon spectrometer momentum resolution is dominated by multiple scattering in the interested momentum range, only a moderate knowledge of the space drift-time relationship is required. In order to achieve a  $p_T$  resolution better than 10% at 20 GeV, the sagitta has to be measured with a precision of 1 mm, which translates into a segment position resolution of 600  $\mu\text{m}$ . The requirement on the segment angular resolution is driven by the angular resolution of the NSW, which is 1 mrad. Single ( $r-t$ ) relation per MDT chamber is sufficient since the variation inside chamber due to temperature gradients, non-uniform magnetic field, and different  $\gamma$  background levels is small.

### 8.6.5 Segment Reconstruction

The track segments are reconstructed on the mezzanine boards. There are two options for the design of the mezzanine board, which share the design of the main ATCA blade. One design has FPGA as the main element for the segment reconstruction. The other design is based on AM chips. The former is the current baseline, but the final choice will be given after the review.

#### Segment Reconstruction with FPGA

The mezzanine board with FPGA provides the flexibility for the algorithm of the segment reconstruction. The baseline algorithm is the Legendre transform, while another algorithm based on a seeded pattern recognition and a segment fit is also proposed. One has the option to use a segment fit after the Legendre transform for ambiguity resolving and refined angle and position determination. The concept and the performance for each algorithm are introduced in the following.

**Segment Reconstruction with Legendre Transform** The Legendre transform makes it particularly convenient to reconstruct segments using pattern recognition in the MDT chambers, since it does not require to calculate the exact position of the hits along the perimeter of the drift circle. Additionally, the algorithm can be easily parallelised for an FPGA implementation, which means that the processing time does not scale up with the number of hits.

The implementation of the Legendre transform for segment finding is based on the transform of each MDT tube position and drift time,  $(R_0, z_0, r_{\text{Drift}})$ , to the polar coordinates Legendre space  $(r, \theta)$  [8.17]. The basic principle of the transform is described in Fig. 8.34, and uses

$$r = z_0 \cos \theta + R_0 \sin \theta \pm r_{\text{Drift}}. \quad (8.1)$$



Each drift circle is represented by two lines in the Legendre space corresponding to its concave and convex parts, respectively. Each  $(r, \theta)$  point in the Legendre coordinate space represents a line in the Cartesian coordinate system, tangent to the corresponding drift circle. If the Legendre transform is applied on all the drift circles of one chamber, the most populated  $(r, \theta)$  bin of the Legendre space will represent the tangent line that is common to most of the drift circles.

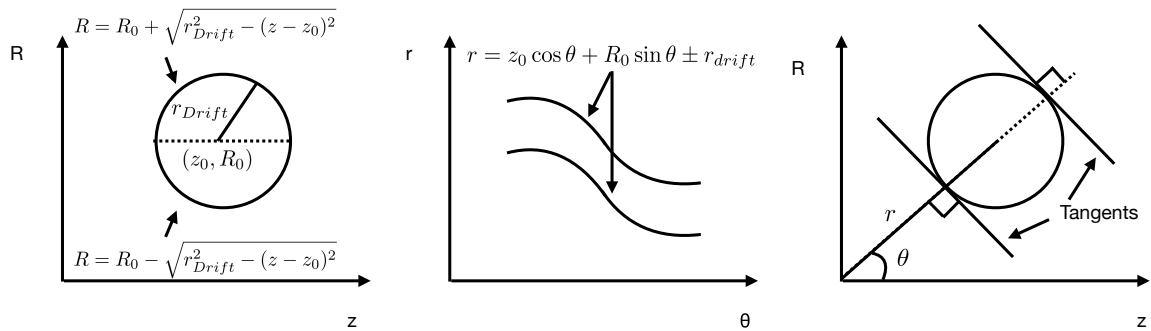


Figure 8.34: Principle of the Legendre transform. The drift circle in Cartesian (left) and Legendre (centre) space as well as the basic principle and mathematical properties (right) are shown.

Figure 8.35 shows a typical event display for the barrel middle station for  $Z \rightarrow \mu\mu$  MC with the background hits. The top plot shows the hits for an event. The full blue circles on the top plot correspond to hits that are attached to the offline reconstructed segment, while the empty ones are background hits not belonging to the offline segment. All the hits within 90 mm distance, corresponding to  $\pm 3$  tubes, are transformed in the Legendre space, that is shown on the bottom plot. The bin with maximum occupancy defines the slope and the intercept of the most popular tangent line.

Proposed granularity of the Legendre space is a bin size of 0.2 mm and 0.5 mrad for  $r$  and  $\theta$ , respectively [8.17]. The performance of the segment reconstruction with the Legendre transform has been studied for different bin sizes and the proposed bin size has proved to give the best results. The limits of the Legendre space have to be sufficiently large such that the common tangent line bin is always within range. Studies have shown that a Legendre space of  $\pm 64$  bins of 0.5 mrad along  $\theta$  and  $\pm 64$  bins of 0.2 mm along  $r$  are sufficient to ensure that the tangent line is reconstructed in more than 95% of the cases.

The performance in terms of the segment angle and position reconstruction has been evaluated using a single-muon MC sample without background hits. To calculate the accuracy of the segment reconstruction that can be achieved, the position and angle between the reconstructed segment and the segment from the offline reconstruction are compared. Figure 8.36 shows the position and angle difference distributions. The position and angle resolution of the reconstructed segments can be estimated by fitting the core distribution in the range  $\mu \pm 1.5\sigma$  with a Gaussian function. The Legendre transform algorithm is able to reconstruct segments with an accuracy for the core of the distributions of  $\sim 0.6$  mrad in

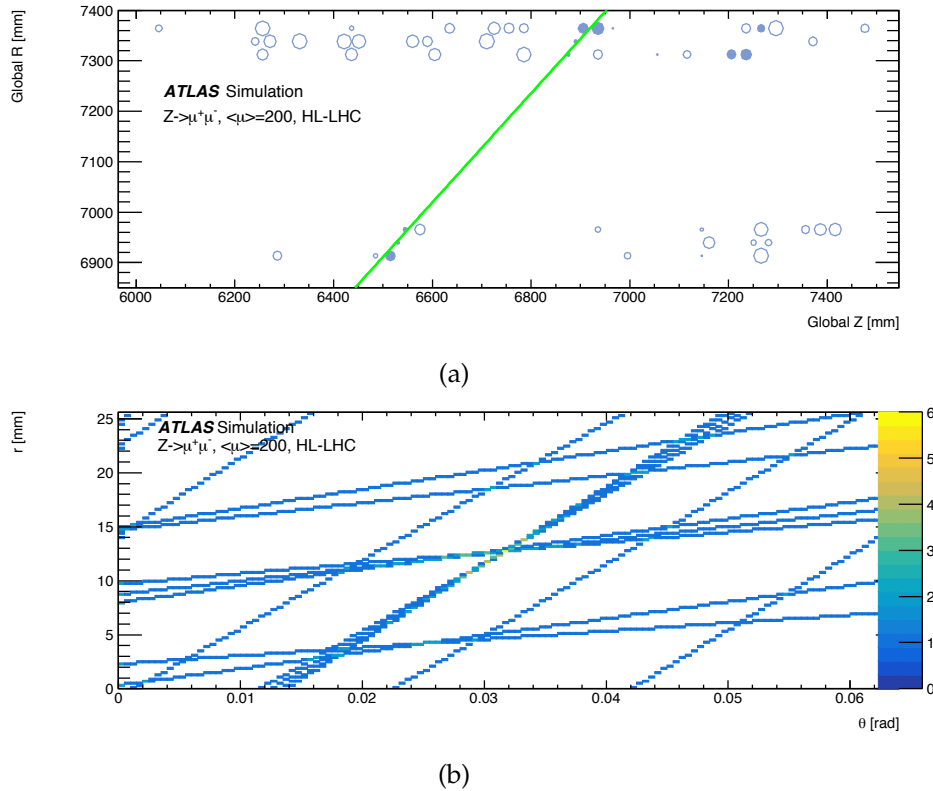


Figure 8.35: Event display (a) and the corresponding representation in the Legendre space (b) for a  $Z \rightarrow \mu\mu$  MC event with background hits. The event display is shown on a global plane, where  $R$  is the cylindrical radius with respect to the axis that passes through the middle of the sector in  $\phi$  and  $Z$  is the coordinate along the beam axis.

angle and  $\sim 90 \mu\text{m}$  in position over a  $p_T$  range of 1–100 GeV. They satisfy the requirements of 1 mrad and  $600 \mu\text{m}$  for the segment angle and position, respectively.

The hardware implementation of the segment reconstruction with Legendre transform fits the resources of modern FPGAs (see Section 8.6.9). In the initialisation stage, the sin and cos values are copied from a pre-computed LUT to local registers for a range of track angles around the RoI angle. The value of  $r = z_0 \cos \theta + R_0 \sin \theta \pm r_{\text{Drift}}$  is calculated simultaneously for all 128  $\theta$  values using DSP blocks in the FPGA fabric. Each  $r$  value is used to increment a histogram bin, while simultaneously keeping a record of the location of the most populated bin. Finally, a binary search is performed to find the most populated bin. From the most populated bin, the output of the segment is provided.

**Segment Reconstruction by Seeded Pattern Recognition and Fit** An alternative and resource-saving method of the segment reconstruction relies on the direction of the flight of the muon evaluated by the Sector Logic. An angular resolution expected is 15 mrad in

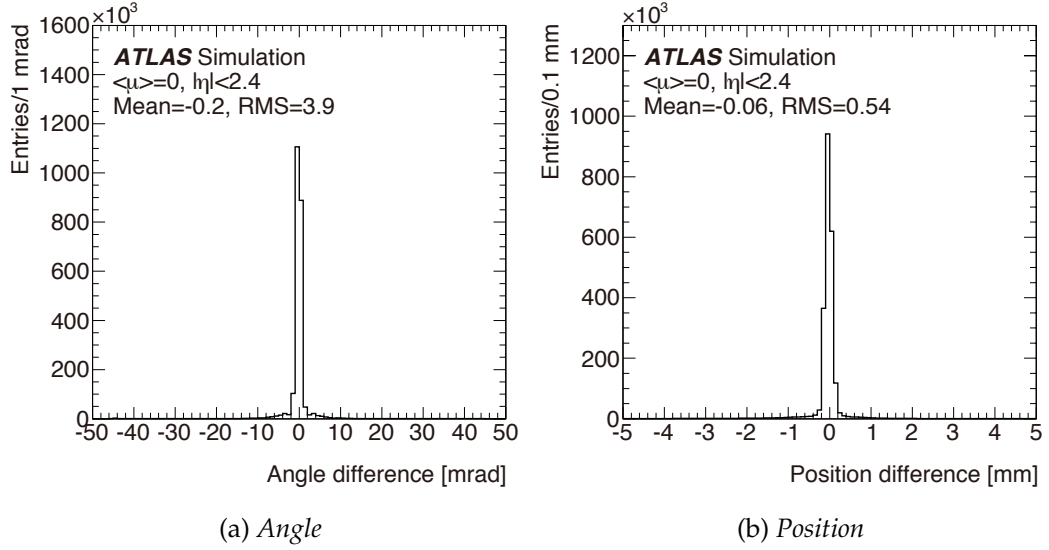


Figure 8.36: Distributions of the difference between reconstructed angle and position of the Legendre transform from the corresponding offline reconstructed segment in the same MDT station.

the barrel inner and middle stations, 50 mrad in the barrel outer station, and 4 mrad in the endcap region. This angle measurement is used in the pattern recognition step [8.18][8.19]. The segment parameters can be determined by the fit with the embedded CPU by using a SoC device alternatively to FPGA.

For the following discussion of the seeded pattern recognition and fit, we use a coordinate system in which the  $z$  axis is orthogonal to the drift tube layers and the  $y$  axis is parallel to the tube layers (see Fig. 8.37). If  $(y, z)$  denotes the position of the anode wire of a tube, the distance  $r$  of a muon track from the anode wire of a traversed drift tube is given by

$$r = \frac{|m \cdot z + b - y|}{\sqrt{1 + m^2}},$$

where  $m$  and  $b$  denote the slope and the intercept of the straight segment and  $y$  and  $z$  are the horizontal and vertical coordinates of the tube in the chamber coordinate frame. Using the measurement of the drift radius  $r$  and the value of the slope  $m$  as provided by the Sector Logic, one can transform this equation into an equation for the intercept  $b$ . There are two solutions  $b_{\pm}$  for the value of the intercepts:

$$b_{\pm} = \pm r \sqrt{1 + m^2} - (m \cdot z - y).$$

The values of  $b_{\pm}$  obtained from the hits in the drift tubes in the RoI can be used to find the right hit pattern. If one treats the hits in the two multilayers of an MDT chamber separately and computes  $b_{\pm}$  at the  $z$  position of the second layer of each multilayer, the limited resolution in the segment slope from the Sector Logic of up to 50 mrad translates into a  $b_{\pm}$

resolution of  $\lesssim 1$  mm, which is sufficient to identify the hits which line up on a straight segment.

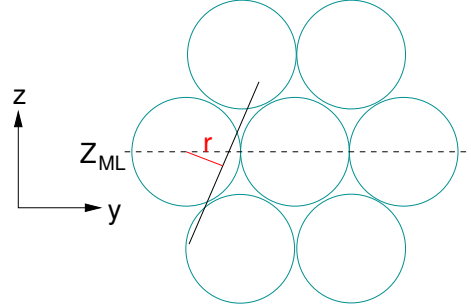


Figure 8.37: Schematic drawing illustrating the definition of the chamber coordinate system and the distance  $r$  of straight line from the anode wire of a drift tube.

The computation of  $b_{\pm}$  and the right MDT hits can be carried out on FPGA or programmable logic part of SoC. Choosing the centre of the  $y$  axis in the middle of the RoI, eight bits are enough to store  $b_{\pm}$  in steps of 0.82 mm. The values of  $b_{\pm}$  are filled in a histogram with 1 mm to 2 mm bin width. From the histogram, one finds the largest accumulation of equal  $b_{\pm}$  values. The hits which are associated to the largest accumulation are the hits which lie on a common straight line.

The MDT hits which are found to line up on a straight line are then transferred to the embedded CPU via a fast AXI link for the determination of the segment parameters and the  $\chi^2$  as goodness-of-fit estimator. For each of these hits, one knows the sign  $\alpha$  in front of the  $r\sqrt{1+m^2}$  term. By replacing the slope  $m$  by  $\bar{m} + \delta$  and linearising  $\sqrt{1+(\bar{m} + \delta)^2}$  in  $\delta$ , one gets a single equation with two free parameters,  $\delta$  and  $b$ , for each hit:

$$\alpha r \sqrt{1 + \bar{m}^2} - (\bar{m}z - y) + \left( \frac{\alpha r \bar{m}}{\sqrt{1 + \bar{m}^2}} - z \right) \delta - b = 0.$$

The parameters  $\delta$  and  $b$  can be determined by a least-square fit to the following  $n$  points:

$$\left( \alpha_k r_k \sqrt{1 + \bar{m}^2} - (\bar{m}z_k - y_k), \frac{\alpha_k r_k \bar{m}}{\sqrt{1 + \bar{m}^2}} - z_k \right)_{k \in \{1, \dots, n\}}.$$

The segment fitting procedure was tested by evaluating the deviations of the reconstructed segment slopes from the true segment slopes (Fig. 8.38). At a segment reconstruction efficiency close to 100%, this distribution shows negligible non-Gaussian tails. A similar distribution is obtained for the segment positions. Negligible non-Gaussian tails are important for a reliable momentum determination from the segment positions and segment slopes.

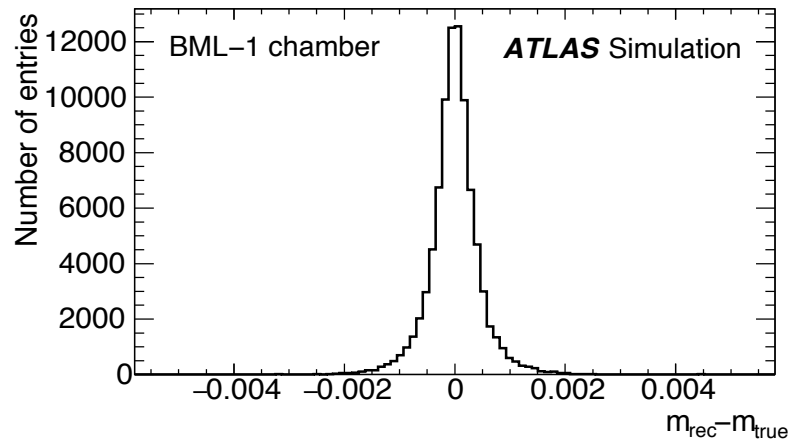


Figure 8.38: *Distribution of the deviation between the reconstructed and true slopes in simulated data for the chambers of the barrel middle station. The distribution has negligible non-Gaussian tails. Background hits corresponding to Run 2 are included in the simulation.*

### Segment Reconstruction with AM Chips

In the alternative design, the AM chip developed for FTK [8.20] and Phase-II HTT is employed for the segment finding. The AM chip is a powerful pattern recognition device, which is widely used for the trigger upgrade programme at ATLAS. The application for MDT is designed so that the system may share the exact hardware design of the full custom AM chips and the mezzanine card named Pattern Recognition Mezzanine (PRM) with the other AM-based hardware tracking projects.

In the MDT segment reconstruction with the AM chips, a segment candidate requires a coincidence of hits in the eight tube layers in the barrel inner station, and six tube layers everywhere else. A hit is represented with the tube channel number, which uniquely identifies the position of the tube centre and an unsigned radius. For each MDT station, a set of 800k most probable coincidences (“patterns”) is computed offline with a MC simulation for high  $p_T$  muons, typically for those with  $p_T > 6$  GeV. The unsigned drift radius is divided into bins in the pattern computation with 3 mm binning for the outermost two layers and 5 mm binning for the other layers in an initial study. The coincidence patterns are computed and loaded into AM chips. A pattern might be consistent with more than one set of sign assignments (i.e. the left and right solution for each drift radius) for the six or eight tube hits. All the possible sets of sign assignment per pattern are stored so that it can be mapped for found patterns in the segment reconstruction. If a muon trajectory fires two tubes per layer according to its incident angle, one of the tubes will be taken as the representative in the pattern computation, avoiding duplication of patterns.

The main functional blocks of the segment reconstruction with the **AM** chips are the hit extraction, the segment finding, the mapping for sign assignments, the road warrior, and the segment calculation. The **AM** chips and the **FPGAs** on the PRM are interfaced with the parallel fast DDR buses, constituting the segment reconstruction engine. Two PRM cards cover an **MDT** sector, and a PRM card covers two **MDT** stations. Eight **AM** chips will be available per station so that four different majority logic thresholds in the pattern matching will run simultaneously. Either of 3-out-of-6 or 5-out-of-8 will be taken as the lowest possible coincidence, which minimises the segment reconstruction inefficiency due to the intrinsic hit-level inefficiency or incomplete pattern lists. The positions of tube centre, the full resolution drift radius, and the sign assignment will be given in advance to the segment calculation for the **MDT** hits of the coincided patterns. In addition, an initial value for the angle for a found pattern will be available in the segment calculation. They allow the track parameters of angle and position to be extracted by an online fit technique of linear approximation, which has been used for the **FTK** system, or  $\chi^2$  minimisation for straight trajectories. The road warrior functionality followed by the segment calculation will discard patterns which share the same set of fired **MDT** tube hits for an event in order to minimise the segment calculations, which could be the case if more than one independent patterns share the subset of tubes hits.

Figure 8.39 shows performance plots for the segment finding with the **AM** approach at the endcap middle station, evaluated with the single muon **MC** sample without the background hits. The left figure shows the number of **AM** output after the sign assignment and before the road warrior algorithm runs, where the different colours correspond to different majority logic thresholds. The right figure shows the segment finding efficiency at the same station as a function of  $p_T$  of muon candidates reconstructed by the offline tool. The efficiency is defined by the fraction of finding a segment reconstructed by the **AM** pattern finding which shares a hit with the segment reconstructed by the offline tool. An efficiency of about 99% is obtained.

### Segment Reconstruction Algorithm Choice

The performance of the **MDT** trigger is sensitive to the segment reconstruction efficiency and resolution including the non-Gaussian components of the angle and the position. The choice of the segment reconstruction will be given after the careful review of the following items. The review is planned to take place in the end of 2019. For the evaluation of the performance, common **MC** samples are used between different algorithms.

1. Segment resolution: resolution of the position and the angle for the segments of the **MDT** trigger with respect to the segments reconstructed by the ATLAS offline analysis
2. Trigger efficiency: trigger efficiency depending on  $p_T$ ,  $\eta$ , and  $\phi$ , for single-muon and multi-muon **MC** samples with the background hits expected under the **HL-LHC** condition

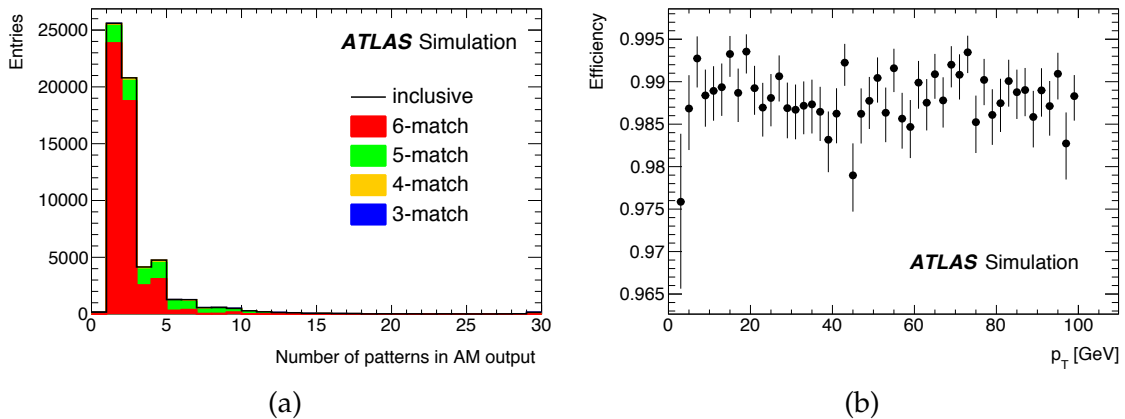


Figure 8.39: Performance plots for the segment finding with the *AM* approach at the endcap middle station, evaluated with the single muon *MC* sample without the background hits. The left figure shows the number of *AM* output before the road warrior algorithm runs, where different colours correspond to different majority logic thresholds. The right figure shows the segment finding efficiency at the same station as a function of  $p_T$  of muon candidates reconstructed by the offline tool.

3. Fake efficiency: efficiency of finding segments not matched with the offline segments, for single-muon *MC* samples with background hits expected under the *HL-LHC* condition
4. Algorithm flexibility: flexibility of the algorithm to retain the exotic signatures such as displaced vertices and slow particles
5. Latency: maximum latency of the *MDT* Trigger Processor depending on the hit rate
6. Capability to handle multiple *RoIs*: maximum number of *RoIs* that can be handled simultaneously, as well as the expected deadtime depending on the *RoI* rates
7. Firmware implementation feasibility: feasibility of the firmware implementation including the resource usage in *FPGA*
8. Cost: cost of chips, engineering, and algorithm development
9. Risk: risk on the availability of the chips and possible future change in the cost
10. Power consumption: power consumption depending on the *RoI* and *MDT* hit rates

## 8.6.6 Transverse Momentum Evaluation

For each *RoI*, the track fitter receives up to three segments from the three stations.<sup>1</sup> It evaluates the transverse momentum  $p_T$  with the variables described in Section 8.6.1.

The applicability of the different methods is mostly defined by the detector structure geometry and the limited detector coverage in specific areas of the muon spectrometer in addition to the muon trigger acceptance. Using offline reconstructed muons of medium quality

<sup>1</sup> In the transition region up to four stations may be crossed by a muon and can be handled.

from a single muon MC sample, the method applicability can be studied as a function of the  $\eta, \phi$  position along the spectrometer. The study shows that for 73% of muons the three-stations method is applicable, and for 21% only the two-stations method is applicable. The remaining 6% of the muons where neither methods is applicable are mostly concentrated in the very forward region of  $|\eta| > 2.6$  and the calorimeter service region at  $|\eta| \simeq 0$ .

In terms of performance, the sagitta method relies on the accuracy of the track segment position reconstruction while the accuracy of the method that uses the deflection angle is dependent on the angle resolution of the reconstructed track segments. Both the sagitta and the deflection angle are correlated with the muon  $p_T$  and the magnetic field. No precise magnetic field map is required because the momentum dependency on the deflection of the muon trajectory can be parameterised with simple parabolic functions depending on  $\eta$  and  $\phi$  [8.21]. Figure 8.40 shows the  $p_T$  resolutions of the reconstructed muon, using off-line reconstructed segments as input, and applying the sagitta or deflection angle and the parameterisation mentioned above to extract the  $p_T$ . The  $p_T$  resolution obtained for 20 GeV muons is  $\sim 6\%$ .

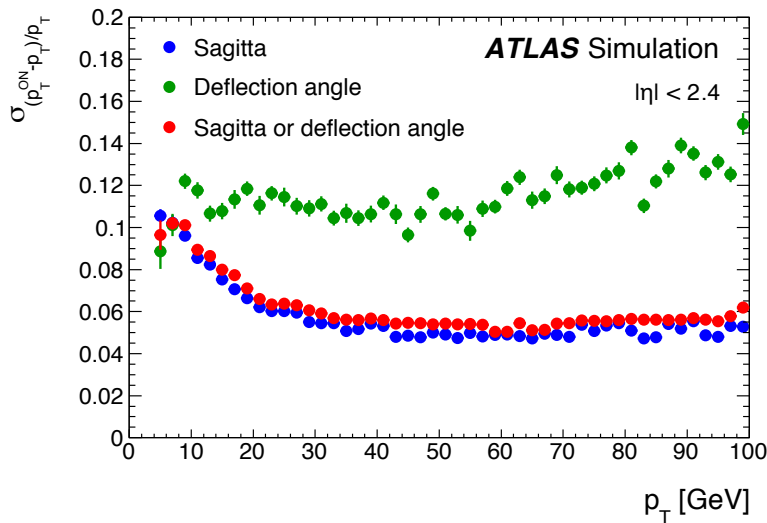


Figure 8.40: Resolution of  $p_T$  as a function of the offline muon  $p_T$ , defined as the standard deviation from a Gaussian fit to the  $p_T$  imbalance between the reconstructed muon momentum as described in the text and the offline muon momentum. The values are obtained from MC samples with no pile-up. For the “sagitta or deflection angle” method, when track segments in three stations are available the sagitta is used, otherwise the deflection angle, computed using track segments in two stations, is used.



## 8.6.7 Realtime Output Data Format

One 48-bit word is used for each muon candidate selected by the MDT Trigger Processor and sent back to the Sector Logic. The contents of the data word are the same as those for the data transfer from the Sector Logic to the Level-0 MUCTPI shown in Table 8.11.

## 8.6.8 Readout Path

In addition to the trigger processing described above, the MDT Trigger Processor will transfer to the DAQ all MDT hits matched to the L0A. Since the L0A rate foreseen is 1 MHz and the MDT maximum drift time is  $\sim 760$  ns, close to 100% of the MDT hits will be transferred to the DAQ. The system must be able to process  $\approx 1.2$  GHz MDT hit rate per sector, buffer them for  $10 \mu\text{s}$  L0A latency (or  $35 \mu\text{s}$  which is the L1A latency for the evolution scenario) and match the hits within  $\approx 760$  ns window to 1 MHz L0A. A block diagram of the proposed DAQ logic for one sector is shown in Fig. 8.41. In order to handle the expected MDT data volume, 16 window-match processors and six FELIX links at 100% utilisation would be adequate. For safety, 12 links are foreseen. The multiplexer should aim to keep groups of hits assigned to one L0A together and in time order. A large buffer, implemented in a separate memory bank, is required to store data before transferring to FELIX. The data is transmitted up to 12 FELIX links using load-balancing to send roughly equal bandwidth across all links. The data format of the data received from the CSM and transmitted to FELIX depends on the final technology choice of the CSM and final specifications of the TDC ASIC, and thus has not yet been defined.

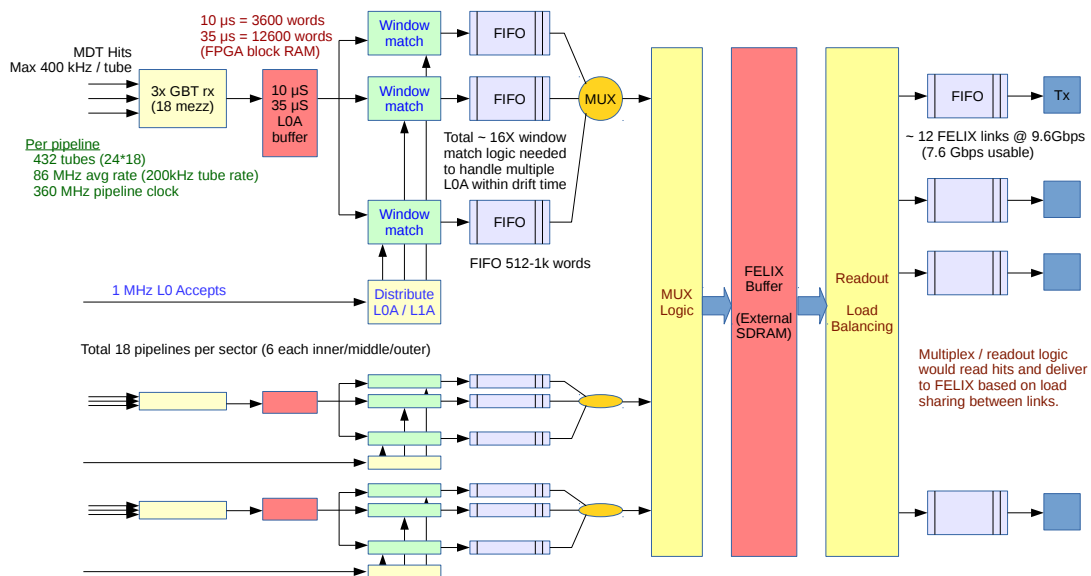


Figure 8.41: DAQ readout path for MDTs for one sector.

### 8.6.9 Resource Estimates

A preliminary estimate of **FPGA** resources required to implement the design with the segment reconstruction based on the Legendre transform is given in Tables 8.13 and 8.14. For the main blade **FPGA**, the resources given in Table 8.13 represent a small fraction of the available resources in *e.g.* a Virtex UltraScale XCVU125 device, which is one of the small available parts with the required number of high-speed links. For the mezzanine **FPGAs**, the best choice currently is the Kintex UltraScale XCKU115. A preliminary estimate of the power consumption for a blade results in  $\approx 250W$ , which is comfortably lower than the 350W limit.

Table 8.13: Preliminary **FPGA** resource estimate for **ATCA** blade. Note that this estimate includes resources which could be split between the main UltraScale-class **FPGA** and the logic fabric on the Zynq®.

<b>Main Blade FPGA</b>				
<b>Function</b>	<b>Number</b>	<b>LUTs</b>	<b>Regs</b>	<b>BRAMs</b>
GBT- <b>FPGA</b> Inputs	54	71,000	51,000	0
Tube Number <b>LUT</b>	18	150	150	220
<b>RoI</b> Processing	57	7800	7800	0
Hit Calibration	54	1000	1000	1000
Hit Extraction	54	1000	1000	0
<b>DAQ</b> ( <i>VERY</i> preliminary)	18	18,000	18,000	100
<b>Total</b>		99,000	79,000	1400
<b>%Use</b> (XCVU125)		14.0%	4.3%	28.0%

Table 8.14: Preliminary **FPGA** resource estimate for mezzanine board. This estimate is for three Legendre segment finder engines, handling inputs from three **MDT** stations. We expect these three to fit in one large **FPGA** (*e.g.* XCKU115). The mezzanine board will contain three such **FPGAs** which can operate in parallel to handle multiple simultaneous **RoIs**.

<b>Mezzanine FPGA (1 of 3)</b>				
	<b>Number</b>	<b>LUTs</b>	<b>Regs</b>	<b>BRAMs</b>
Legendre Engine	3	96,500	104,500	275
<b>Total</b>		289,000	313,500	825
<b>XCKU115</b>		663,360	1,326,700	4,320
<b>%Use</b>		44%	24%	19%

## 8.7 Latency Estimates

This section describes the latency estimates for the Level-0 Muon Trigger System. The values of **CBE** are shown throughout the section. The uncertainty and a margin are assigned

---

for each item, and [MEV](#) and [MPV](#) are obtained from the sum with [CBE](#) (Table 5.5). The latencies related to the optical links, including the serialisation, the optical transmission, and the deserialisation, are accounted on the transmitter side. This definition leads to slight difference from the values shown in other [TDRs](#) [8.1][8.22].

Estimations on the latency for the on-detector electronics of the new [RPC](#) and [TGC](#) systems are described in Ref. [8.1]. The latency from the [RPC](#) signal arrival at the on-detector boards named DCT to the [RPC](#) signal arrival at the barrel off-detector boards where the Sector Logic is implemented is shown as  $0.725 \mu\text{s}$ . By taking a sum with the latency from the bunch crossing to the signal arrival at DCT, a latency estimate from the bunch crossing to the [RPC](#) signal arrival at the barrel off-detector boards is obtained to be  $1.110 \mu\text{s}$ . The estimated latency from the bunch crossing to the [TGC](#) signal arrival at the endcap off-detector boards where the Sector Logic is implemented is  $0.888 \mu\text{s}$ .

An estimate on the latency for the on-detector electronics of the new [MDT](#) system is also described in Ref. [8.1]. The latency for each [MDT](#) hit to arrive at the [MDT](#) Trigger Processors depends on the hit rate and the drift time of the hits. The latency estimate from the bunch crossing to the arrival of the latest [MDT](#) hits at the off-detector boards for possible maximum hit rate and drift time is shown as  $2.137 \mu\text{s}$ . After the submission of Ref. [8.1], a further study of the latency was performed with a simulation based on the hardware description language. The maximum cable length has been updated to 90 m. The maximum value of the sum of the drift time and the propagation time has also been updated to 800 ns. The updated estimate on the latency from the bunch crossing to the arrival of the latest [MDT](#) hits at the off-detector boards is  $2.358 \mu\text{s}$ . The value for the earliest [MDT](#) hits is  $0.609 \mu\text{s}$ .

An estimate on the latency from the bunch crossing to the arrival of the [NSW](#) track segment at the Endcap Sector Logic for the Phase-I system is shown in Ref. [8.4] to be 41 clock ticks, corresponding to  $1.025 \mu\text{s}$ . After further studies, the originally estimated value is found to be small. The current best estimate for the Phase-I system is 47 clock ticks, corresponding to  $1.175 \mu\text{s}$ . For the Phase-II system, a value  $1.425 \mu\text{s}$  is assumed. The additional latency is exploited for the upgrade of the algorithm of the [NSW](#) segment reconstruction. An estimate on the latency from the bunch crossing to the arrival of the [Tile](#) calorimeter energy flags at the Sector Logic is  $1.425 \mu\text{s}$  [8.22].

Table 8.15 shows an overview of the estimate on the latency of the Level-0 Muon Trigger System for [HL-LHC](#), including the on-detector and off-detector parts. After receiving the [RPC](#) and [TGC](#) hit signals, the Sector Logic processes the signals and provides primitive muon track candidates. A selection is applied on the primitive muon track candidates based on the information from [NSW](#) and the [Tile](#) calorimeter, and the results are sent to the [MDT](#) Trigger Processors. A latency estimate for the Sector Logic track candidate arrival at the [MDT](#) Trigger Processor is  $1.675 \mu\text{s}$ . After that the [MDT](#) Trigger Processor starts further selection of the muon track candidates based on the [MDT](#) hits, and the results are sent to the Sector Logic. A latency estimate for the arrival of the muon track candidates at the

## 8.7 Latency Estimates

Sector Logic is  $3.810 \mu\text{s}$ . A latency estimate for the Level-0 muon track candidate arrival at the Level-0 **MUCTPI** is  $4.010 \mu\text{s}$ .

Table 8.15: Overview of the best estimate on the latency for the Level-0 Muon Trigger System for HL-LHC. The values are cumulative with the bunch crossing as the time origin.

Contents	Latency
<b>RPC/TGC</b> hit arrival at the Sector Logic	$1.110 \mu\text{s}/0.888 \mu\text{s}$
<b>NSW</b> Trigger Processor track segment arrival at the Sector Logic	$1.425 \mu\text{s}$
<b>Tile</b> calorimeter energy flag arrival at the Sector Logic	$1.425 \mu\text{s}$
Sector Logic track candidate arrival at the <b>MDT</b> Trigger Processor	$1.675 \mu\text{s}$
<b>MDT</b> hit signal arrival at the <b>MDT</b> Trigger Processor	$2.358 \mu\text{s}$
<b>MDT</b> Trigger Processor output arrival at the Sector Logic	$3.810 \mu\text{s}$
Level-0 muon track candidate arrival at the Level-0 <b>MUCTPI</b>	$4.010 \mu\text{s}$

Tables 8.16 and 8.17 show the estimates on the latency for the Barrel Sector Logic and the Endcap Sector Logic, respectively. After receiving the **RPC** hit signals, the Barrel Sector Logic processes the signals and provides primitive muon track candidates ( $1.360 \mu\text{s}$ ). After receiving the **TGC** hit signals, the Endcap Sector Logic processes the signals of the **TGCs** in the endcap middle station and provides primitive muon track candidates ( $1.013 \mu\text{s}$ ). The Endcap Sector Logic takes coincidence with the **TGCs** on the endcap inner station and also with the **RPCs** in the BIS78 region ( $1.113 \mu\text{s}$ ). The Endcap Sector Logic applies a further selection with the **NSW** segments for  $|\eta| > 1.3$  ( $1.450 \mu\text{s}$ ). The Sector Logic takes coincidence with the **Tile** calorimeter cells for  $|\eta| < 1.3$  ( $1.450 \mu\text{s}$ ). Transverse momentum is evaluated and the muon track candidate is transferred to the **MDT** Trigger Processor ( $1.675 \mu\text{s}$ ). The Sector Logic receives the muon track candidates from the **MDT** Trigger Processor ( $3.810 \mu\text{s}$ ), and transmits the muon track candidates to the Level-0 **MUCTPI** ( $4.010 \mu\text{s}$ ).

The processing time in the **MDT** Trigger Processor depends on the segment reconstruction algorithms. Table 8.18 shows the estimate on the latency for the baseline algorithm, where the segments are reconstructed by the Legendre transform. All **MDT** hit signals are transferred to a buffer with a depth of  $0.967 \mu\text{s}$  after the decoding and the domain crossing which take  $0.100 \mu\text{s}$ , to assure that the earliest **MDT** hit signals are stored until the arrival of the Sector Logic track candidates. The sum  $0.967 \mu\text{s} + 0.100 \mu\text{s} = 1.067 \mu\text{s}$  is equal to the difference between the earliest **MDT** hit signal arrival and the Sector Logic track candidate arrival. The **MDT** hits which are used for further processing are extracted with the **RoI** provided based on the Sector Logic output ( $3.432 \mu\text{s}$ ). The extracted hit signals are transferred to the mezzanine cards, the segments are reconstructed in the mezzanine cards, and the data of segments are transferred to the main board ( $3.560 \mu\text{s}$ ). The muon track candidates are selected with the transverse momentum ( $3.635 \mu\text{s}$ ). The data of selected muon track candidates are transferred to the Sector Logic ( $3.810 \mu\text{s}$ ). If a least-square fit is applied to refine the segment angle and position,  $0.500 \mu\text{s}$  is added. The estimate for the segment reconstruction based on the **AM** chips is  $0.301 \mu\text{s}$  larger than the baseline.

Table 8.16: Summary table of the best estimate on the latency for the Barrel Sector Logic for HL-LHC. The values are cumulative with the bunch crossing as the time origin.

Contents	Latency
RPC hit signal arrival	1.110 $\mu\text{s}$
Hit signal decoding and preprocessing	1.210 $\mu\text{s}$
Majority logic and path reconstruction	1.310 $\mu\text{s}$
Combination coincidence	1.360 $\mu\text{s}$
Tile calorimeter energy flag arrival	1.425 $\mu\text{s}$
Coincidence with the Tile calorimeter cells	1.450 $\mu\text{s}$
Trigger candidate selection	1.475 $\mu\text{s}$
Transverse momentum encoding	1.500 $\mu\text{s}$
Optical link to the MDT Trigger Processor (10 m)	1.675 $\mu\text{s}$
MDT Trigger Processor output arrival	3.810 $\mu\text{s}$
Final trigger processing	3.835 $\mu\text{s}$
Optical link to Level-0 MUCTPI (10 m)	4.010 $\mu\text{s}$

Table 8.17: Summary table of the best estimate on the latency for the Endcap Sector Logic for HL-LHC. The values are cumulative with the bunch crossing as the time origin.

Contents	Latency
TGC hit signal arrival	0.888 $\mu\text{s}$
Coincidence of the TGCs on the endcap middle station	1.013 $\mu\text{s}$
Coincidence with the TGCs on the endcap inner station	1.063 $\mu\text{s}$
Coincidence with the RPCs in the BIS78 region	1.113 $\mu\text{s}$
NSW track segment and Tile energy flag arrival	1.425 $\mu\text{s}$
Selection with the NSW and Tile data	1.450 $\mu\text{s}$
Trigger candidate selection	1.475 $\mu\text{s}$
Transverse momentum encoding	1.500 $\mu\text{s}$
Optical link to the MDT Trigger Processor (10 m)	1.675 $\mu\text{s}$
MDT Trigger Processor output arrival	3.810 $\mu\text{s}$
Final trigger processing	3.835 $\mu\text{s}$
Optical link to Level-0 MUCTPI (10 m)	4.010 $\mu\text{s}$

## 8.8 R&D Items

R&D is needed to establish the final design of the components of the Level-0 Muon Trigger System. The schedule of the designing, the prototype development, and the demonstration is summarised in Chapter 19. The R&D items which can be worked on with evaluation boards or custom test boards available before full prototypes should be basically completed

Table 8.18: Summary table of the best estimate on the latency for the *MDT Trigger Processor* for *HL-LHC*. Segment reconstruction based on the Legendre transform is assumed. The values are cumulative with the bunch crossing as the time origin.

Contents	Latency
Earliest <i>MDT</i> hit signal arrival	0.609 $\mu\text{s}$
Sector Logic track candidate arrival	1.785 $\mu\text{s}$
Latest <i>MDT</i> hit signal arrival	2.358 $\mu\text{s}$
Decoding and domain crossing	2.458 $\mu\text{s}$
Buffering	3.424 $\mu\text{s}$
Hit extraction	3.432 $\mu\text{s}$
Data transfer to the mezzanine cards	3.457 $\mu\text{s}$
Segment reconstruction	3.510 $\mu\text{s}$
Conversion to the detector coordinates	3.535 $\mu\text{s}$
Data transfer to the main board	3.560 $\mu\text{s}$
Transverse momentum evaluation and selection	3.635 $\mu\text{s}$
Optical link to the Sector Logic (10 m)	3.810 $\mu\text{s}$

before *PDR*, which will take place component by component from 2020. The R&D items which require prototype developments should be worked on with prototypes and the final design should be prepared around *FDR*, which will take place component by component from 2021. The main R&D issues are shown in the following.

- Detailed design of the Sector Logic should be provided and demonstrated, including the logic of the *RPC* coincidence, the *TGC* track segment reconstruction, the requirement on the angle difference between the *TGC* and *NSW* track segments, and the coincidence with the *Tile* calorimeter. For the barrel, possibilities of a three point sagitta measurement to improve the momentum resolution and a tighter timing cut to suppress the accidental coincidence will be studied. For the endcap, a possibility of the coverage extension from  $1.05 < |\eta| < 2.4$  to  $1.05 < |\eta| < 2.7$  by a coincidence between *NSW* track segments and *TGC* hits will be studied. The acceptance for the exotic signatures will be investigated, and dedicated logic will be prepared if appropriate.
- Detailed design of the *NSW* track segment reconstruction in the *NSW* Trigger Processor should be provided and demonstrated. A combined reconstruction with the *sTGC* and *MM* hits will be designed. The logic will be tuned with the efficiency, the angle and position resolution, and the fake segment reconstruction rate.
- Detailed design of the *MDT* Trigger Processor should be provided and demonstrated. The hit extraction, the track segment reconstruction, and the transverse momentum evaluation require the demonstration. For the track segment reconstruction, the performance of the different algorithms will be carefully studied and an algorithm will be selected based on the results (see Section 8.6.5).

- Trigger logic to suppress multiple track candidates for single muon should be designed. The chamber overlaps will be carefully taken into account in the logic.
- Dependence of the trigger logic on the pile-up should be studied, and the logic less sensitive to the pile-up should be established.
- Precision of the resource estimation for the Sector Logic, the [NSW](#) Trigger Processor, and the [MDT](#) Trigger Processor should be improved with detailed design of the trigger logic. The trigger logic will be optimised to minimise the required resource.
- Detailed design of the readout logic should be provided and demonstrated.
- Data flow including  $\sim 100$  input and output links and internal flow of [FPGA](#) should be demonstrated.
- Detailed design of the control functions for the on-detector boards of [RPC](#), [TGC](#), and [MDT](#) should be provided and demonstrated. The design of the control functions for the off-detector boards with [Zynq®](#) devices should also be provided and demonstrated.
- Latency estimation with the detailed trigger logic should be provided and demonstrated.
- Precision of the estimation for the power consumption should be improved with the detailed design of the trigger logic, the readout logic, and control functions.

## References

- [8.1] ATLAS Collaboration, *Technical Design Report for the Phase-II Upgrade of the ATLAS Muon Spectrometer*, Tech. Rep. CERN-LHCC-2017-017. ATLAS-TDR-026, CERN, Geneva, Sep, 2017. <https://cds.cern.ch/record/2285580>.
- [8.2] ATLAS Collaboration, *ATLAS New Small Wheel Technical Design Report*, Tech. Rep. CERN-LHCC-2013-006. ATLAS-TDR-020, CERN, Geneva, Jun, 2013. <https://cds.cern.ch/record/1552862>.
- [8.3] ATLAS Collaboration, *ATLAS First Level Trigger: Technical Design Report*, Tech. Rep. CERN-LHCC-98-14, ATLAS-TDR-12, CERN, Geneva, Jun, 1998. <https://cds.cern.ch/record/381429>.
- [8.4] ATLAS Collaboration, *Technical Design Report for the Phase-I Upgrade of the ATLAS TDAQ System*, Tech. Rep. CERN-LHCC-2013-018. ATLAS-TDR-023, CERN, Geneva, Sep, 2013. <https://cds.cern.ch/record/1602235>.
- [8.5] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, [JINST 3 \(2008\) S08003](#).
- [8.6] ATLAS Collaboration, *Measurements of the electron and muon inclusive cross-sections in proton-proton collisions at  $\sqrt{s} = 7$  TeV with the ATLAS detector*, [Phys. Lett. B 707 \(2012\) 438](#).



- [8.7] M. Cacciari, M. Greco, and P. Nason, *The  $p(T)$  spectrum in heavy-flavour hadroproduction*, **JHEP** **9805** (2001) 007.
- [8.8] ATLAS Collaboration, *Performance of the ATLAS Trigger System in 2015*, **Eur. Phys. J. C** **77** (2017) 317, [arXiv:1611.09661](https://arxiv.org/abs/1611.09661) [hep-ex].
- [8.9] A. Falkowski, J. T. Ruderman, T. Volansky, and J. Zupan, *Discovering Higgs Decays to Lepton Jets at Hadron Colliders*, **Phys. Rev. Lett.** **105** (2010) 241801, [arXiv:1007.3496](https://arxiv.org/abs/1007.3496) [hep-ph].
- [8.10] *MiniPOD AFBR-814VxyZ, AFBR-824VxyZ, 14 Gbps/Channel Twelve Channel, Parallel Fiber Optics Modules*, Tech. Rep. AV02-4039EN, Avago Technologies, 2013.
- [8.11] *Specifications for Enhanced Small Form Factor Pluggable Module SFP+, Revision 4.1*, Tech. Rep. SFF-8431, SFF Committee, July, 2009.
- [8.12] J. Dubbert, S. Horvat, O. Kortner, H. Kroha, and R. Richter, *Upgrade of the ATLAS Muon Trigger for the SLHC*, **JINST** **5** (2010) C12016, [arXiv:1011.0163](https://arxiv.org/abs/1011.0163) [physics.ins-det].
- [8.13] O. Sasaki, Y. Suzuki, K. Nagano, M. Ishino, M. Ikeno, K. Bendtz, T. Domaie, H. Sakamoto, T. Kawamoto, C. Fukunaga, H. Shichi, M. Tomoto, Y. Sugaya, K. Tokunaga, and H. Kurashige, *Design studies of the ATLAS muon Level-1 trigger based on the MDT detector for the LHC upgrade*, **JINST** **5** (2010) C12021.
- [8.14] ATLAS Collaboration, *Letter of Intent for the Phase-II Upgrade of the ATLAS Experiment*, Tech. Rep. CERN-LHCC-2012-022. LHCC-I-023, CERN, Geneva, Dec, 2012. <https://cds.cern.ch/record/1502664>.
- [8.15] Y. Horii, *A muon trigger upgrade with high transverse momentum resolution for the ATLAS detector at the High-Luminosity LHC*, **JINST** **10** (2015) C01005.
- [8.16] P. Bagnaia et al., *Calibration model for the MDT chambers of the ATLAS Muon Spectrometer*, Tech. Rep. ATL-MUON-PUB-2008-004, CERN, Geneva, Feb, 2008. <http://cds.cern.ch/record/1089868>.
- [8.17] T. Alexopoulos, M. Bachtis, E. Gazis, and G. Tsipolitis, *Implementation of the Legendre Transform for track segment reconstruction in drift tube chambers*, **Nucl. Instrum. Methods Phys. Res. Sect. A** **592** (2008) 456–462.
- [8.18] S. Abovyan, V. Danielyan, M. Fras, P. Gadow, O. Kortner, S. Kortner, H. Kroha, F. Müller, S. Nowak, R. Richter, and K. Schmidt-Sommerfeld, *Hardware Implementation of a Fast Algorithm for the Reconstruction of Muon Tracks in ATLAS Muon Drift-Tube Chambers for the First-Level Muon Trigger at the HL-LHC*, in *Proceedings for 2017 IEEE Nuclear Science Symposium and Medical Imaging Conference, Atlanta, Georgia, United States*. Oct, 2017.



- 
- [8.19] S. Abovyan, D. Cieri, V. Danielyan, M. Fras, O. Kortner, S. Kortner, H. Kroha, F. Müller, S. Nowak, R. Richter, and K. Schmidt-Sommerfeld, *Concept of a Cost-Effective First-Level Muon Track Trigger for Future Hadron Collider Experiments*, MPP-2018-37 (2018) .
- [8.20] ATLAS Collaboration, *Fast TracKer (FTK) Technical Design Report*, Tech. Rep. CERN-LHCC-2013-007. ATLAS-TDR-021, CERN, Geneva, Jun, 2013. <https://cds.cern.ch/record/1552953>.
- [8.21] P. Gadow, *Development of a Concept for the Muon Trigger of the ATLAS Detector at the HL-LHC*, Master thesis at Technical University of Munich, MPP-2016-086 (2016) . <https://cds.cern.ch/record/2162398>.
- [8.22] ATLAS Collaboration, *Technical Design Report for the Phase-II Upgrade of the ATLAS Tile Calorimeter*, Tech. Rep. CERN-LHCC-2017-019. ATLAS-TDR-028, CERN, Geneva, Sep, 2017. <http://cds.cern.ch/record/2285583>.



# 9 Global Trigger

## 9.1 Introduction

The design goal of the **Global Trigger** subsystem is to bring **Event Filter (Event Filter)**-like capability to the Level-0 trigger system. Unlike the other hardware trigger upgrades, the emphasis in the **Global Trigger** is overwhelmingly on firmware rather than on hardware. Topologically the **Global Trigger** consists of a layer of incoming multiplexing **MUX** nodes<sup>1</sup> feeding data into a layer of **GEP** processing nodes, where each **GEP** node receives the complete trigger data for an event, followed by a demultiplexing **CTP Interface** node providing the output to the **CTP**, illustrated in Fig. 9.1. Physically each node corresponds to one of the two large **FPGAs** on a **Global Common Module (GCM)**. This is a natural step in the evolution away from custom hardware towards functionality being provided in firmware or software.

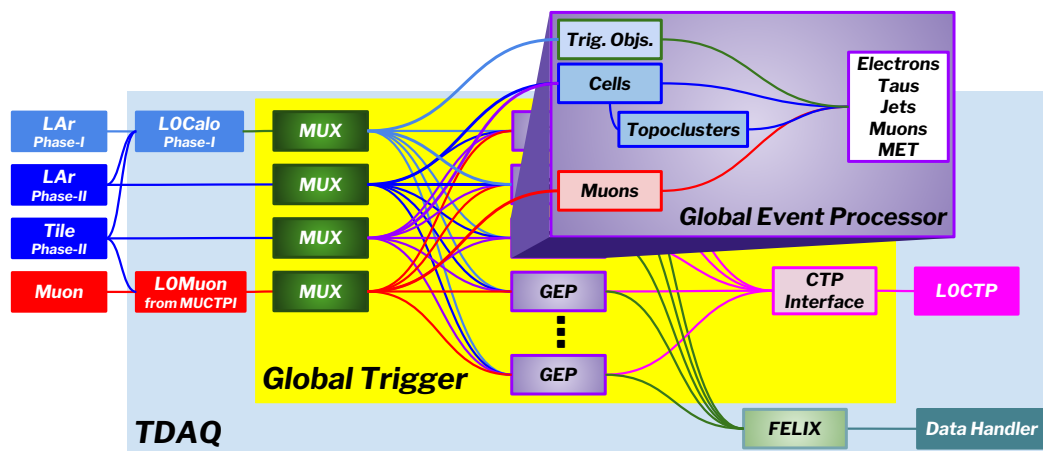


Figure 9.1: Illustration of the **Global Trigger** system processing and trigger decision flow.

The overall design of the ATLAS **Global Trigger** is analogous to that of the successful Run 2 CMS Calorimeter Trigger upgrade [9.1] and learns from the experience with the commissioning of the ATLAS **FTK** and **L1Topo** subsystems. The use of common hardware, the **GCM**, simplifies system design and long-term maintenance, and minimises the complexity

<sup>1</sup> Throughout this chapter the term node is used in this topological sense and corresponds to a single large **FPGA** running a common firmware infrastructure into which the specific firmware for a particular function is embedded.

## 9.1 Introduction

of the firmware. Employing a common firmware framework, illustrated in Fig. 9.2, with rigorous interfaces minimises interference between firmware for different trigger algorithms. This also allows the firmware to be factorised along the lines of the HLT Trigger Signatures which simplifies organisation and improves overall support for software development and performance studies.

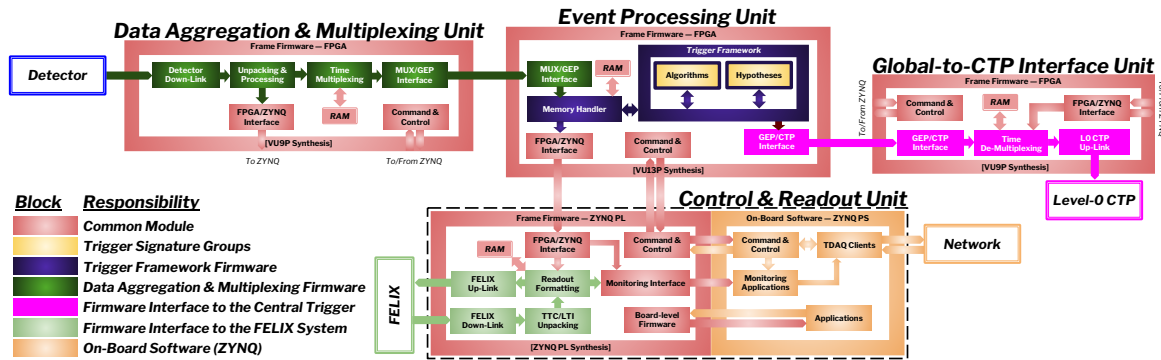


Figure 9.2: *Global Trigger* system firmware flow and interfaces.

The design amalgamates the complete trigger data for each event onto a separate GEP processing node. This concentration of data maximises flexibility by minimising the distributed processing of data within the system. As the number of input fibres into the system far exceeds the capability of any single device, the data are time-multiplexed such that data corresponding to a particular BC are received by an incoming multiplexing (MUX) layer and re-transmitted over a number of BC to the single GEP destination. Suitable ordering of the data by the MUX for this long transmission allows two-dimensional data to be processed by a one-dimensional algorithm on the GEP node as the data arrive, which can reduce the FPGA resource usage for sliding window algorithms by some two orders of magnitude over the traditional regional implementation. The relatively long time to transmit the data moreover decouples each node from the LHC BC rate, allowing asynchronous complex algorithms including those based on machine learning techniques. These can more closely emulate trigger processing in the Event Filter, and provide additional triggering flexibility beyond what is currently feasible with the Phase-I hardware trigger.

### 9.1.1 Requirements for the Global Trigger System

The baseline requirements for the Global Trigger can thus be defined:

- receive and process trigger primitives from calorimeter subsystems to produce trigger objects (TOBs);
- receive TOBs from the L0Calo FEX system and from L0Muon (via the MUCTPI);
- multiplex input data for serial transmission and processing;
- apply EF-like trigger algorithms and hypotheses on suitable selections of these TOBs;

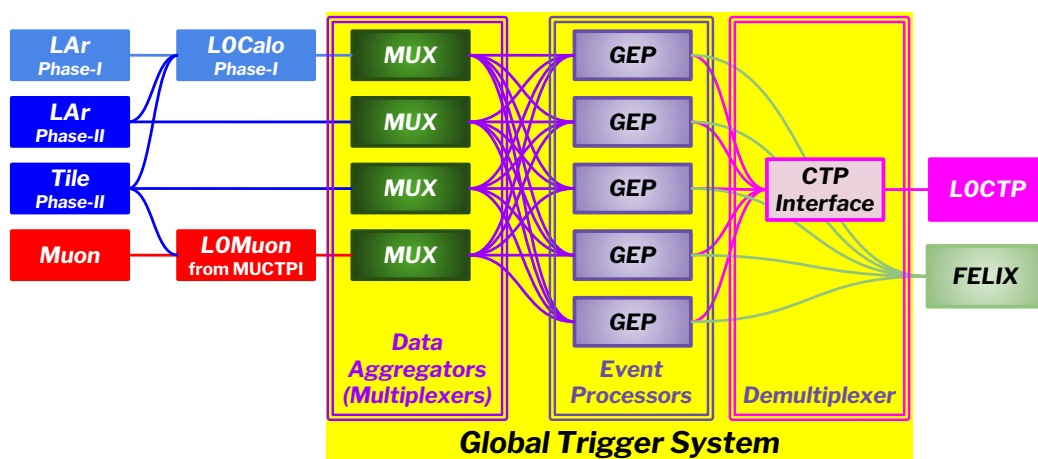


Figure 9.3: The functional design of the *Global Trigger* system, illustrating the detector inputs, multiplexing *MUX* layer, multiplexed event-processing *GEP* layer, demultiplexing *CTP Interface*, and connections to other systems.

- transmit the resultant *Global Trigger* trigger input bits (*TIP*) to the *CTP* for processing.

The trigger hypotheses to be constructed will include those based on object multiplicities, energy thresholds, and topological relationships of objects. The resultant *TIP* contains flag bits indicating which Global algorithm requirements have been satisfied and may also include multiplicities of identified objects.

This structure thus subsumes and surpasses the functionality of the Phase-I *L1Topo* modules. In addition to the minimal requirements, extended physics opportunities have also been identified:

- the internal time-multiplexed implementation removes the limitation on the number of *TOBs* from the *FEX* and *MUCTPI* systems present in the Phase-I hardware trigger;
- the access to cell level data from the calorimeters at 40 MHz allows for improved trigger object definitions using clustering algorithms that cannot be implemented in the *FEX* systems; and,
- the time-multiplexed architecture decouples the *GEP* node from the LHC *BC* rate thereby allowing the use of asynchronous and iterative high-level algorithms impossible in the Phase-I hardware trigger.

## 9.1.2 Overview of Global Trigger System Design

The architecture of the *Global Trigger* can be seen in Fig. 9.3, and consists of three primary components: a *MUX* system, a *GEP* system, and a demultiplexing *CTP Interface*.

## 9.1 Introduction

As illustrated in Fig. 9.4, each input detector and trigger system sends its data every BC with one packet per optical fibre to MUX nodes. The MUX receives these data packets from all the input optical fibres synchronously, collects and reorganises them, and then starts transmitting all of the event data from a specific BC to a GEP node. At the next BC the data for the next event are captured into the MUX buffers for the next GEP node and transmission of the data to that node begins.

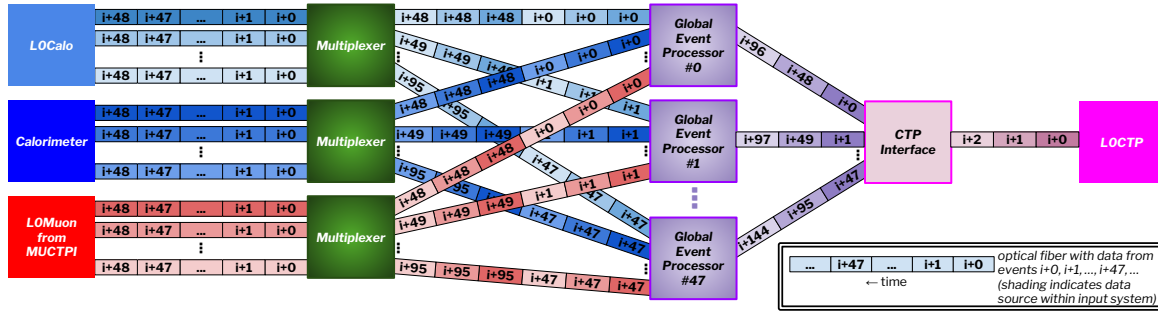


Figure 9.4: Multiplexing of incoming synchronous data into the GEP layer. Each MUX node receives data from all the input optical fibres synchronously every BC, collects and reorganises them, and then starts transmitting all of the event data from a specific BC to a GEP node. Each GEP node takes up to 48 BC to receive the event. Results from the GEP are sent to the CTP Interface where they are demultiplexed and sent to the CTP.

The time-multiplexed signals are transported to one of many GEP nodes over many BC where the inputs are processed by trigger algorithms, as illustrated in Fig. 9.4. The time required to transmit these data from the MUX to the GEP is a function of the number of optical fibres incident upon a given MUX, the number of GEP nodes, and the speeds of the data-sources-to-MUX and MUX-to-GEP links. With the 48 GEP nodes as described in Section 9.4.2, the inputs for the entire Global Trigger can then be mapped onto a single GEP node. The GEP nodes handle every BC in a round-robin fashion, with each one starting to receive a new event every 48 BC. The data for an event then takes up to 48 BC to be transferred from each MUX to the GEP node, before the data for the next event for that GEP node are captured. The processing on the GEP node can thus be decoupled from the LHC clock.

Suitable ordering of the data by the MUX for this long transmission can allow processing on the GEP node to begin as soon as first data arrive. Using the time of data transmission to scan across the full calorimeter allows two-dimensional data to be processed by a one-dimensional algorithm on arrival in the GEP. For example, having the MUX layer send data ordered in  $\eta$  means that the GEP node can process that data in complete  $\phi$  rings. This enables immediate regional processing as the data arrives, and the time for data transmission can then be exploited for regional data processing, including  $\phi$ -ring based pileup suppression, before starting iterative algorithms that require the full event to have been received. Ordering the data geometrically allows all algorithms to be fully pipelined which, in FPGAs, means that: processing is localised, fan-outs are reduced, routing delays are

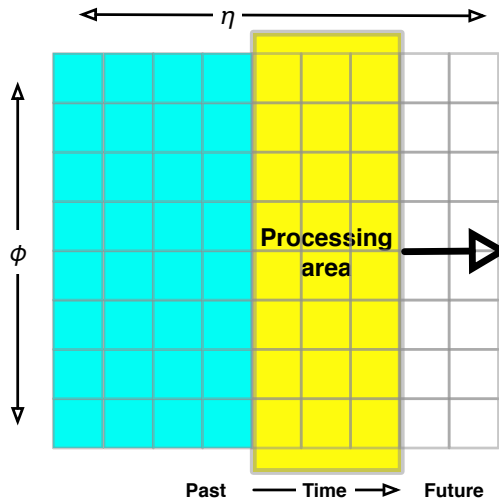


Figure 9.5: *Spatial pipelining of data into a GEP node. The 2D algorithm is reduced down to operations in a single dimension or localised 2D operations spread over time.*

minimised, and register duplication is eliminated. For any sliding-window algorithm this drastically reduces the FPGA resource usage as the 2D algorithm is reduced down to operations in a single dimension or localised 2D operations spread over time as illustrated in Fig. 9.5. The output from this stage can be pipelined into the next stage of the algorithm which can be similarly pipelined. This can reduce the FPGA resource usage for algorithms such as topoclustering described in Section 9.3.3 by some two orders of magnitude over the full-plane implementation.

Objects produced by the Global Trigger algorithms running on the GEP nodes are buffered for readout (via FELIX) while the resulting TIP are sent to the CTP Interface. The CTP Interface demultiplexes the TIP, re-synchronising the data stream to the LHC beam clock, so that it can be transmitted to the CTP on the correct BC.

Global Trigger thus builds upon the information available in the L0Calo and L0Muon subsystems and uses zero-suppressed full granularity cell information ( $|E_T| > 2\sigma$ ) from the calorimeters. It will perform event selection using a wide range of sub-detector trigger primitives (summary data on position and magnitude of energy depositions) allowing refined selections on electrons, taus, jets,  $E_T^{\text{miss}}$ , and muons.

## 9.2 Global Trigger Interfaces

The Global Trigger will receive inputs from the calorimeter and muon detector subsystems, as well as the L0Calo FEX trigger processors. The resultant Global Trigger TIP will be sent directly to the CTP system and to the Event Filter via FELIX. In addition to the TIP that

is sent to the CTP, FELIX has to receive all data necessary to support the EF processing, offline data analysis and monitoring. This includes a compact data structure that maps TIP items (flags or multiplicities) to the TOBs that contributed to these items. These interfaces are illustrated schematically in Fig. 9.3. All external interfaces on the realtime path are synchronous to the LHC clock, such that the MUX system and CTP Interface encapsulate the asynchronous time-multiplexed GEP system.

### 9.2.1 Calorimeter Subdetector Inputs

The Global Trigger will receive data from the LAr and Tile calorimeter subsystems every LHC bunch-crossing. The LAr Signal Processor (LASP) and Tile TDAQ Interface (TDAQi) will provide zero-suppressed cell information ( $|E_T| > 2\sigma$ ) that will further be processed by a clustering algorithm (described in Section 9.3.3) to provide extended calorimeter information to the GEP algorithms.

### 9.2.2 L0Calo Inputs

The Global Trigger will receive the TOBs from the eFEX, jFEX, and gFEX processors every BC, providing pre-processed calorimeter features that can be implemented directly as seeds in trigger algorithms. Each TOB provides an object type ( $e/\gamma$ ,  $\tau$ , jFEX jets, gFEX jets) and measurements of energy,  $\eta$  and  $\phi$ .

### 9.2.3 L0Muon Inputs

The Global Trigger will receive the muon-track candidate information (TOBs) from the L0Muon trigger via the MUCTPI every BC. The muon track-candidates will contain information on candidate  $p_T$ ,  $\eta$  and  $\phi$ .

### 9.2.4 Central Trigger Processor Outputs

The Global Trigger will produce a 1024-bit TIP representing the results of the Global Trigger algorithms every BC. The TIP is made up of flags and multiplicity values, the meaning of each bit field being defined by the configuration of the Global Trigger and the CTP. The flags may indicate the presence of an individual object passing the corresponding criteria (e.g.  $p_T$  threshold requirement), or more complex requirements on combinations of objects (e.g. invariant mass and charge-sign requirements on pairs of muons). Multiplicity values can be sent to the CTP counting objects passing criteria such as  $p_T$  thresholds and isolation cuts. The Global Trigger TIP is produced on the GEP node processing the event and then demultiplexed in the CTP Interface and sent to the CTP.



## 9.2.5 Front-End Link eXchange (FELIX) Inputs and Outputs

The [Global Trigger](#) will have an output interface to the [FELIX](#) system from every [GEP](#) module for the readout of the [TIP](#) and other trigger information. The [MUX](#) modules and [CTP Interface](#) will also have output links, but in the normal course of running these will only be used for reporting errors. Every module must also receive clock, configuration, and per-event [TTC](#) information from [FELIX](#), as well as the Level-0 (and in the evolved scheme Level-1) trigger-accept signals.

For every [LOA](#) the [Global Trigger](#) will produce a compact data structure mapping the [TIP](#) onto the [TOBs](#) which caused those [TIP](#) bits to be fired as part of the readout data. This will allow the [Event Filter](#) to quickly identify the [TOBs](#) which caused the [LOA](#) and hence the [RoI](#) requiring regional tracking. All intermediate objects produced by the algorithms will also be read out, including the topoclusters, jets, and other derived [TOBs](#). However, as with the [LOCalo FEX](#) modules, the full input data will in general not be read out from the [Global Trigger](#); rather the data will be read out at their respective sources and the integrity of the data will be confirmed by suitable checksums.

## 9.2.6 Configuration and Control, Interface to Detector Control System

On power up each module in the [Global Trigger](#) will be configured from local SPI flash. Each module will have an Ethernet interface and IPbus will be used to control the operation of all FPGAs, allowing steering of the system, parametric control of the algorithms and control of the high-speed links. This interface will further support remote programming of the FPGA and SPI flash along with access to playback and spy memories attached to the real-time path and environmental monitoring such as currents, temperatures and voltages. Low-level, ATCA-compliant control will be achieved by an ATLAS standard IPMC module communicating with [DCS](#) via the ATCA Shelf Manager.

## 9.3 Trigger Strategy and Algorithms

The trigger strategy for the [Global Trigger](#) will use a two stage processing scheme, as illustrated in Fig. 9.1. In the first stage, the [GEP](#) node receives the fine-granularity calorimeter data and processes them in algorithms that produce inputs to the second stage of the trigger processing. The first stage also receives the [TOBs](#) from the [LOCalo](#) and [LOMuon](#) systems. In the second stage the Level-0 [TOBs](#) provide a seed for the calculation of electron/photon, tau, anti- $k_t$  jet, and muon trigger candidates, which are refined using topological clusters and isolation information and combined in more complicated hypotheses such as topological relationships.

More specifically, **TOBs** from the **FEX** and **MUCTPI** systems can be further processed with the topocluster and LAr first-layer strip information to produce refined trigger objects. Using **jFEX** and **gFEX** jets as **RoI** seeds, jets can be clustered using topoclusters with an anti- $k_t$  algorithm (as described in Section 9.3.4). Event-by-event local pile-up suppression for jets using baseline subtraction techniques comparable to those developed for offline analyses can be applied. At this stage of processing, boosted-jet and jet substructure information can be extracted for use in the **Global Trigger** calculations. Similarly, the **eFEX TOBs** provide seeds for electron, photon, and tau identification wherein topoclusters and LAr first-layer strip information can be used to calculate longitudinal and transverse energy distribution with improved resolution and for particle isolation. The topoclusters can further be used to extend electron identification to forward regions ( $|\eta| > 2.4$ ). With access to the full list of muon trigger information from the **MUCTPI**, additional processing of muon candidates can proceed at the **GEP**. Muon isolation with topocluster and strip information will further improve background rejection. Global event observables will be calculated using topoclusters and muons, including  $H_T$  and  $E_T^{\text{miss}}$ .

Each of these refined objects and global observables will provide input to the extended trigger hypotheses evaluated by the **GEP**. Thresholds will be defined for each object or combinations of objects (multiplicity and  $E_T$  thresholds), corresponding to a series of hypotheses. Topological conditions based on the spatial relationship and energies of objects can also be defined, for example (n-)jet+ $E_T^{\text{miss}}$  triggers with specific  $\Delta\phi(\text{jet}, E_T^{\text{miss}})$  requirements. Specialised algorithms to support B physics and heavy ion physics are also anticipated to be included for the **GEP**. The results of these hypotheses are passed to the **CTP Interface**.

### 9.3.1 Algorithm Performance

Preliminary studies of the **Global Trigger** system algorithm performance have been performed. Descriptions of the preliminary algorithms are included below in this section, and a description of the resultant trigger performance is given previously in this document for topological clusters (Section 6.2), electrons and photons (Section 6.3), taus (Section 6.5) and jets (Section 6.6).

### 9.3.2 Algorithm Implementation

These algorithms will be implemented on **FPGAs**, which will require significant modifications compared to any "normal" offline algorithms that have iterative sequences of steps. The original iterative processing steps can be redesigned as a single processing step that can be pipelined in parallel threads on the **FPGA**. The **FPGA** implementations of these algorithms will be constrained by the **FPGA** logic resources and by the maximum allowed latency. As these **FPGA** algorithms are developed, the number of clock cycles required per

iteration and the [FPGA](#) resources required by the algorithm are metrics for the overall performance of the algorithm. Optimising on these metrics allows the [Global Trigger](#) to maximise the number of algorithms that can operate simultaneously. Prototype algorithms for the [GEP](#) are expected to have a strategy for processing data (parallel, sequential or mixed). Bit-wise simulation of each algorithm to reproduce online results will be required to study and optimise algorithm physics performance and precision.

In this section, two proposed algorithms are described, which are expected to be the most FPGA resource intensive: topological clustering of calorimeter cell information and the “anti- $k_t$ ” jet-finding algorithm (see also Section 6). These algorithms represent important algorithms that are desired for the [Global Trigger](#) and provide benchmarks for the implementation metrics mentioned above. The topoclustering algorithm is a prerequisite for many of the other algorithms anticipated to be utilised in the [GEP](#). The anti- $k_t$  jet-finding algorithm is an example of a challenging case of translating a highly-iterative algorithm to the [FPGA](#) environment. A third algorithm, which uses  $E_{\text{ratio}}$  calculated using the [LAR](#) strips to form isolation variables, yields large factors of background rejection and is readily adaptable to implementation in FPGA. This algorithm’s physics performance potential is described in Section 6.3, but the firmware implementation is not yet ready for a detailed description here.

The topoclustering and strip isolation algorithms are targeted to achieve a small utilisation of 1-2% of the proposed [FPGA](#) logic resources with a low algorithm latency ( $<1 \mu\text{s}$ ). The anti- $k_t$  jet-finding algorithm is expected to have a larger resource utilisation (10-20% of the proposed [FPGA](#) logic resources) and a longer latency. The total latency goal for producing topoclusters and processing with the anti- $k_t$  jet-finding algorithm is  $<2.5 \mu\text{s}$ , leaving a margin of roughly  $1 \mu\text{s}$  on further trigger processing. The description of the algorithms below includes estimates of the performance metrics, which indicate that it will be possible to implement the suite of required algorithms together with the additional necessary firmware (control logic, data flow, readout, etc.) while respecting the constraints on latency and [FPGA](#) logic/memory resources.

### 9.3.3 Topological Clustering

ATLAS uses clusters of topologically connected calorimeter-cell signals (topoclusters) as its principle signal definition for use in the reconstruction of isolated hadrons, jets, and hadronically-decaying  $\tau$  leptons [9.2]. In addition, topoclusters are used to represent the energy flow from soft particles which is needed for the reconstruction of  $E_T^{\text{miss}}$  and for establishing the isolation of leptons. Individual topoclusters are not expected to contain the entire response to a single particle. Rather, depending on the incoming particle types, energies, spatial separation, and cell signal formation, individual topoclusters represent the full or fractional response to a single particle, the merged response of several particles, or a combination of merged full and partial showers. Topoclusters, which are at the electromagnetic

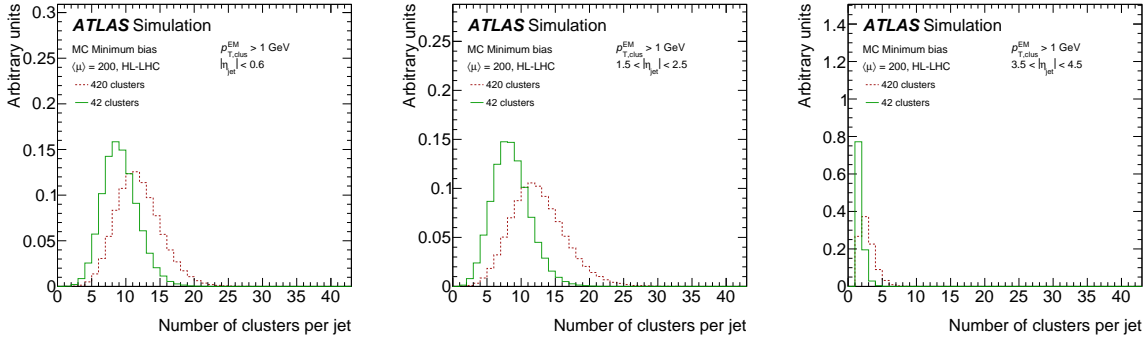


Figure 9.6: Distribution of the number of topoclusters inside anti- $k_T$  jets formed with  $R = 0.4$  in the (left) barrel ( $|\eta| < 0.6$ ), (central) endcap ( $2.0 < |\eta| < 2.5$ ) and (right) *F*Cal ( $3.5 < |\eta| < 4.5$ ) regions of the ATLAS calorimeter, generated with a minimum  $E_T$  threshold of 1 GeV for all topoclusters.

energy scale, can be further calibrated using a local cluster weighting scheme (LCW) [9.3] that relies on the shape of the topocluster that reflects the type of incident particles.

The algorithm to build topoclusters uses the spatial distribution of cell signals in all three dimensions to establish connections between neighbouring cells to reconstruct the energy and directions of the incoming particles. Calorimeter cells with insignificant signals, not connected to neighbouring cells with significant signals, are considered noise and discarded from further use in jet, particle, and  $E_T^{\text{miss}}$  reconstruction. A detailed description of the algorithm together with its performance can be found in Ref. [9.2]. Essentially the same algorithm is used to reconstruct topoclusters as in the *EF* and offline, albeit most likely based on  $E_T$  rather than  $E$ . A first pass selects contiguous groups of cells with energy measurement exceeding two times the expected noise ( $|E_T| > 2\sigma$ ) containing at least one seed cell with energy measurement exceeding four times the expected noise ( $|E_T| > 4\sigma$ ). Such clusters are referred to 422 topoclusters (occasionally as 42 topoclusters). In the *EF* and offline all cells on the periphery of this topocluster are then added regardless of their energy significance (these topoclusters are referred to as 420 topoclusters). This provides a first mitigation of the impact of pileup at the cluster level. However, the resulting topoclusters might be too large to properly reflect local energy flow. In such cases, topoclusters will be split between the various local energy maxima above threshold in a second splitting step. The number of topoclusters reconstructed per event can be seen in Fig. 9.8 for both 420 and 422 algorithms, and for various  $E_T$  thresholds on the clustered cells. This is particularly problematic in the *F*Cal where a large fraction of jets are built from a single topocluster (Figures 9.6 and 9.7). Producing useful jets in this environment will be very expensive in terms of *FPGA* resources and latency. For this reason, the *F*Cal detector region is treated by a dedicated electronics module, the *ffEX*, which is described further in Section 7.3.4.

The first stage of the *GEP* processing includes algorithms that build topoclusters. However, the offline topoclustering algorithm is an iterative algorithm in which calorimeter seeds are

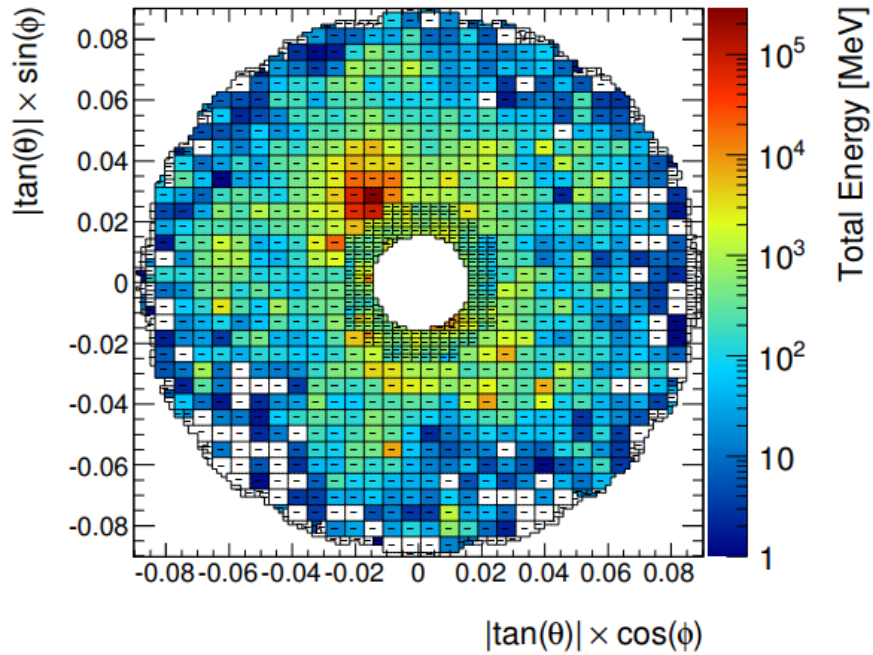


Figure 9.7: Event display of a typical VBF quark jet in FCAL1 at  $\mu = 200$ .

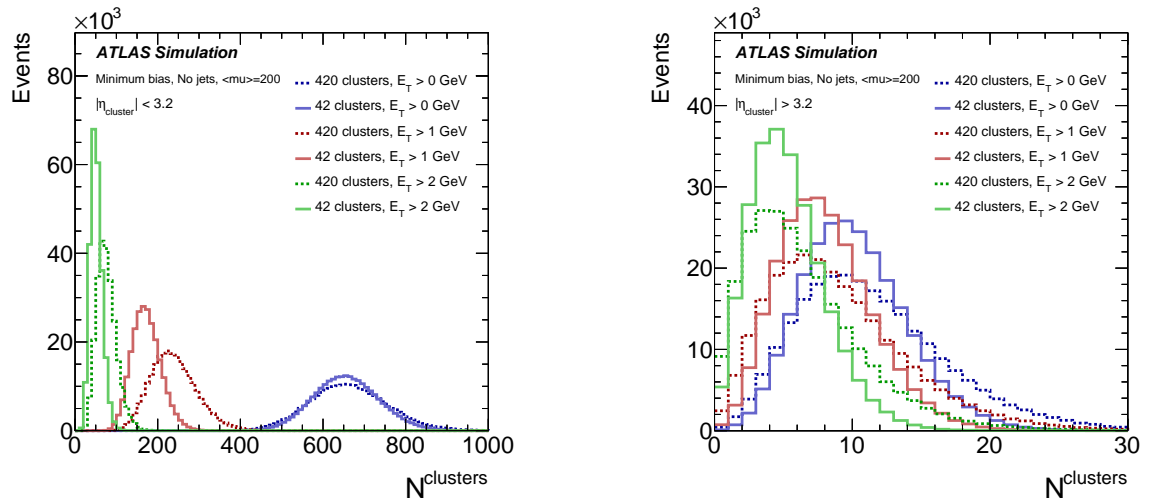


Figure 9.8: Multiplicity of topoclusters for 420 and 422 threshold algorithms (the latter labeled as 42), including variations of  $E_T$  threshold requirements on the clustered cells. The left plot is for topoclusters within  $|\eta| < 3.2$  and the right plot is for forward topoclusters ( $|\eta| > 3.2$ ).

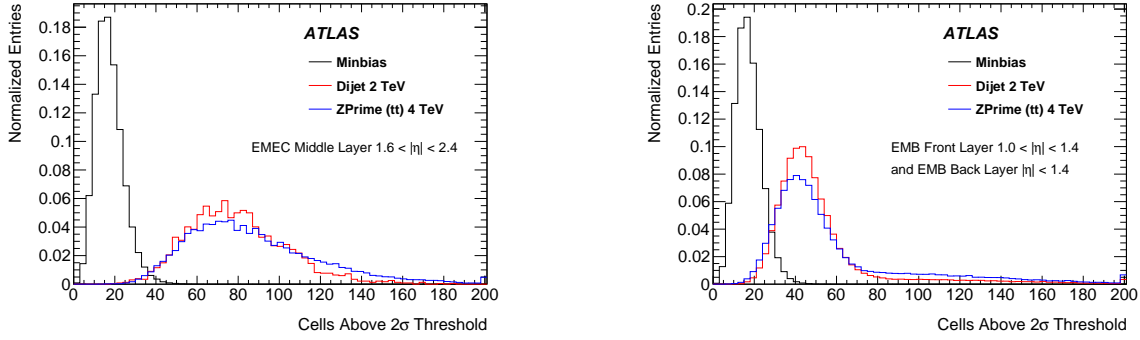


Figure 9.9: Channel occupancy per *LASP* unit for minimum bias events, high-energy dijet events, and high-mass  $Z^0 \rightarrow t\bar{t}$  events at  $\mu = 200$ . (left) *LASP* occupancy in the EMEC region  $1.6 < |\eta| < 2.4$ . (right) occupancy of a *LASP* unit covering  $1.0 < |\eta| < 1.4$  of the EMB front layer and  $|\eta| < 1.4$  of the EMB back layer. Each *LASP* unit is planned to process up to 512 cells. Figure extracted from [9.4].

chosen, then additional cells are iteratively associated with the seeds based on a series of energy thresholds. To adapt such an algorithm to the *FPGA* environment, two primary challenges must be overcome: first, the large calorimeter data volume must be managed such that the algorithm can be parallelised to the greatest extent possible, and second, a non-iterative algorithm can be designed that satisfies *FPGA* resource and latency restrictions. The large volume of the fine-granularity calorimeter data represents a significant challenge in meeting the *FPGA* resource usage target. One important reduction in the data volume is the decision to exclude cells with  $|E_T| > 2\sigma$  so that only 422-type topoclusters can be built on the *GEP*. On average, this corresponds to a fraction of cells of  $\approx 5.5\%$  [9.4] or  $\approx 30$  cells out of the 512 cells processed per *LASP*. However, for some events — especially events with noise bursts or other abnormal conditions — the fraction of high-energy significance cells can easily exceed this average value (Fig. 9.9). A contingency of 30% (153 out of 512 cells) has been assumed for bandwidth estimates. Two solutions are proposed in case this fraction is exceeded: sending cells ordered in energy until output bandwidth saturates, or sending large event fragments in an asynchronous mode to the *MUX*. If the number of cells is too large, then a forced accept should be issued and a mitigation of accept signals caused by regional calorimeter noise bursts is under investigation.

Processing challenges remain even with this significant reduction in the number of calorimeter cells. To address this, a dual serial-parallel data processing technique has been designed, which dramatically reduces the *FPGA* resources required. This implementation takes advantage of the serial delivery of calorimeter-cell input data to the *GEP*. Thus the topological clustering algorithm would receive one  $\eta$  data slice at a time. In addition, various cluster granularities have been studied to understand the balance between the algorithm's physics performance and *FPGA* resource usage. Granularities at the level of calorimeter cells, super-cells ( $\Delta\eta \times \Delta\phi = 0.025 - 0.1 \times 0.1$ , varying by longitudinal layer), *jFEX* towers ( $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ , summed longitudinally), and *gFEX* towers ( $\Delta\eta \times \Delta\phi = 0.2 \times 0.2$ ,



summed longitudinally) have been considered. In the following, we use the terminology "cell" to describe the smallest granularity of unique calorimeter data inputs used in the topoclustering algorithm, though the actual implementation could be a more coarse granularity. The calorimeter inputs are pre-processed to identify cells above a seed threshold energy, and all seeds are processed in parallel. Each longitudinal calorimeter layer is also processed in parallel, and the final output is formed by summing cluster energies from each layer. The output is thus the energy of the clustered object around each seed. In the event that two or more neighbouring seeds could be included in each other's clustered region, all clusters are generated and output to the next stage of GEP processing.

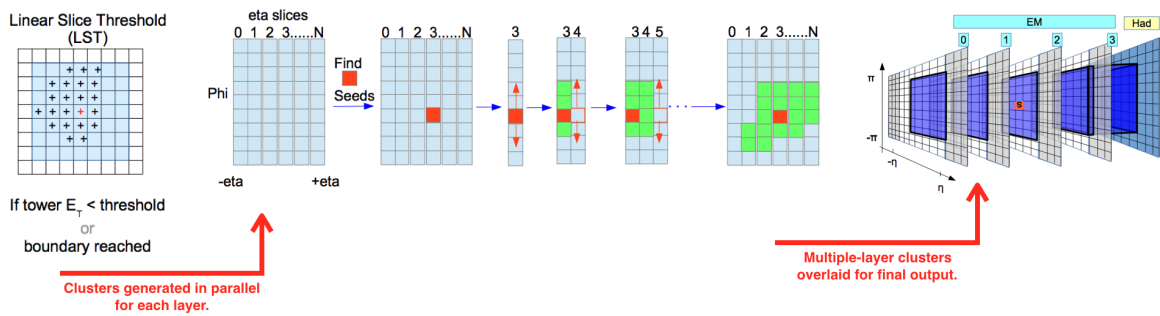


Figure 9.10: Example of the pseudo Linear Slice Thresholding (LST) topological clustering algorithm that could be implemented in the GEP.

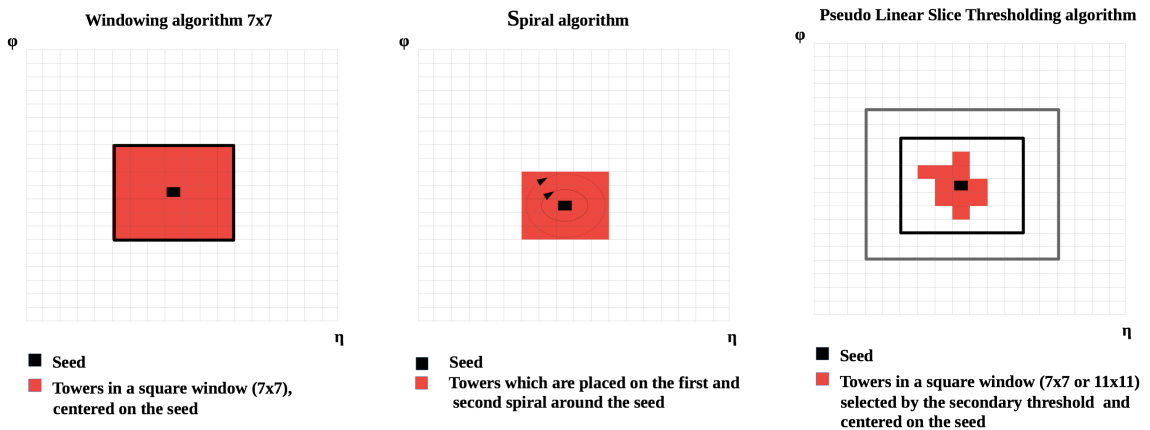


Figure 9.11: Illustration of the three topoclustering algorithm variants currently under study.

Three versions of the topoclustering algorithm have been developed for the GEP: a Windowing algorithm, a pseudo Linear Slice Threshold (LST) algorithm, and a Spiral algorithm. Each algorithm variant is being studied for resource/latency requirements as well as physics performance. All three variants meet the target resource and latency goals. To illustrate the topoclustering algorithms, the LST algorithm is shown in Figures 9.10 and 9.11. The LST algorithm is a seeded algorithm that sums contiguous calorimeter cells above pre-defined

thresholds. Seeds are pre-identified in the input data for each layer of the calorimeter. For each seed, cells are associated if the energy is above a fixed clustering threshold. Starting at the seed's eta slice, the algorithm sequentially tests cells in  $\eta$  slices in the direction of positive  $\eta$  until a slice is found that doesn't have a contiguous cell above threshold or if a predefined boundary is reached. This process is repeated in the direction of negative  $\eta$  starting from the seed. To improve performance, only the cells in a square window centred on the seed are considered in the algorithm ( $7 \times 7$  or  $11 \times 11$  cells), which would define a predefined boundary as described above. Furthermore, the algorithm is typically limited to a central eta region, which defines an additional boundary for the clustering. The two alternative versions are also illustrated in Fig. 9.11. The Windowing algorithm is a simpler variant based on the assumption that the cluster energy deposition region can be approximated by the cells in a square window, centred on the seed. The spiral algorithm is also a seed-based implementation, but contrary to Windowing and LST, the cluster comprises only those cells which are on the first and, optionally, on the second spiral of contiguous cells around the seeds. The spiral algorithm is easily adoptable to the [FPGA](#) architecture and yields a very small latency.

Each algorithm version comprises several components, and each component is designed to process data within a strict number of clock cycles. The overall algorithm latency is determined by the total number of clock cycles used by all algorithm components. All algorithm versions, except the spiral algorithm, use a rising/falling edge clocking technique which reduces algorithm latency. The alternative method is to use only the rising edge and a higher system clock frequency. Thus the spiral algorithm components must run at a higher clock frequency compared to the Windowing and LST versions. In all algorithm instances, an adder tree is used to calculate the cluster energy deposition, and is implemented using the [FPGA LUT](#) resources. However, the latest Xilinx UltraScale+ products (e.g., the VU13P [FPGA](#)) are equipped with a large number of [DSP](#) slices, and thus a method to implement [DSPs](#) has been studied as an alternative solution to implementing adder trees. No algorithm variants have yet been through a full latency (or speed) optimisation at the firmware synthesis level. A summary of the algorithm performance on various [FPGA](#) targets is given in Table 9.1, in which the [FPGA](#) resource usage is given as a percentage of the total available and the latency is given in [FPGA](#) clock cycles (ticks). Even without optimisation, the topoclustering algorithms have been shown to be achievable within the target performance metrics. The maximum algorithm latency is found to be 6 clock ticks (Windowing, with 5 longitudinal layers) and can be as low as 3 clock ticks (Spiral, one longitudinal layer). For a 160 MHz reference clock these correspond to latencies below 50 ns, for a 320 MHz reference clock below 25 ns.

In the tested implementation using Xilinx development boards, the serial data transmission has not yet been included in firmware because it cannot be properly implemented on the development boards. Thus, the utilisation of the [FPGA](#) logic resources is typically more intensive for [LUTs](#), than for flip-flops, due to the large amount of data being loaded on the [FPGA](#). Therefore the anticipated [FPGA](#) logic resource budget for serial data processing can



be found by dividing the presented usage by the number of  $\eta$  slices loaded. The numbers reported in Table 9.1 for the logic resource usage have been corrected for this difference. We find that the logic resources required are at or below 1% on the medium-sized Xilinx Ultrascale chips (VU095 and VU125), and at or below 0.1% for the Xilinx Ultrascale+ chips (VU9P and VU13P). The expected physics performance of the topoclustering is discussed in Section 6.2 and the final choice of algorithm will be made by considering criteria on the processing performance as well as physics performance, emphasising improvement to the reconstruction of jets, electrons, taus and  $E_T^{\text{miss}}$ .

Table 9.1: Summary of *FPGA* performance for various topological clustering algorithm instances. All *FPGA* targets are Xilinx chips from the Ultrascale and Ultrascale+ families. The *FPGA* resource utilisation is given as the fraction of total resources on the chip. The algorithm latency is given in *FPGA* clock cycles (ticks), algorithms have been run with both 160 MHz and 320 MHz reference clocks.

FPGA Family:	V-US	V-US	V-US	V-US	V-US+	V-US+	V-US+
Part Number:	XCVU095	XCVU095	XCVU095	XCVU095	VU9P	VU13P	VU13P
Algorithm:	LST 11x11	Win 11x11	Win 7x7	Win 7x7	Win 7x7	Win 7x7	1-Spiral
Granularity :	gTowers	gTowers	gTowers	gTowers	gTowers	gTowers	SuperCells
Flip-Flops (%):	1.4	0.83	0.56	0.54	0.24	0.17	0.13
LUTs (%):	2.1	1.5	0.88	0.83	0.37	0.25	0.21
Latency (ticks):	6	6	6	6	6	6	3

### 9.3.4 Anti- $k_t$ Jet Clustering

An important goal for the *Global Trigger* is to identify jets with algorithms similar to those used in *EF* and offline. In the following, studies of possible implementations of the standard anti- $k_t$  jet-finding algorithm [9.5] are summarised to highlight issues in implementing an iterative jet-finding algorithm in the *Global Trigger*. While we expect that the computational complexity of this algorithm makes it unsuitable for implementation in the *GEP*, we include here a pedagogical description to set up the following description of algorithm variants.

The *Exact anti- $k_t - E$ -merging* algorithm [9.5] is a sequential recombination jet clustering algorithm. It takes as input a list of  $N$  objects (the *Object List*), which may be clusters, calorimeter cells, trigger towers, etc. and returns a list of jets (the *Jet List*), which correspond to clusters of objects. Objects are clustered iteratively using the weighted separation between two objects (labeled  $i$  and  $j$ ) compared to each object's  $p_{T,i}^{-2}$ :

$$d_{ij} = \min(p_{T,i}^{-2}, p_{T,j}^{-2}) \frac{\Delta R_{ij}^2}{R^2} ; d_{iB} = p_{T,i}^{-2} \quad (9.1)$$

where  $\Delta R_{ij}^2 = (\phi_i - \phi_j)^2 + (\eta_i - \eta_j)^2$  and  $R^2$  is a parameter of the algorithm that sets the size scale of the jet clusters. At each iteration of the algorithm, the minimum of the *Separation List* ( $d_{ij}, d_{iB}$ ) is found. If the minimum corresponds to  $d_{aB}$  then object  $a$  is moved to the *Jet List* and removed from the *Object* and *Separation* lists. Otherwise, if the minimum is  $d_{ab}$  then objects  $a$  and  $b$  are merged to form a new object ( $k$ ) which is added to the *Object* and *Separation* lists while objects  $a$  and  $b$  are removed. This process continues until there are no objects left in the *Object List*. Merging is performed in the so-called ‘‘E-Scheme’’ by adding the 3-momenta of objects  $a$  and  $b$  to define  $p_T, \phi, \eta$  for the new object,  $k$ . At most,  $N$  iterations are required to finish the algorithm and after each iteration, the number of objects is reduced by one. Minimum-finding at each iteration scales like the number of objects present at that iteration squared. Thus, the simplest version of the anti- $k_t$  algorithm scales with the number of objects as  $N^3$ .

Implementation of the algorithm in an **FPGA** takes advantage of parallelism to reduce the latency required at each iteration. For example, a binary tree minimum finder run on an  $N$  object list requires  $\log_2 N$  steps to complete if overhead and timing delays are ignored. Additionally, the full *Separation List* ( $d_{ij}, d_{iB}$ ) is not calculated. Rather, separate lists of object  $p_T^{-2}$  or  $p_T$  (the  *$p_T$  List*) and the distance between object  $i$  and all other objects ( $\Delta R_{ij}^2$ ) (the *Distance Lists*) are used. Calculating the  $N$  elements  $p_{T,i}^{-2}$  requires a **LUT** and is probably most efficiently done during the clustering stage, which would also save on latency. The distances,  $\Delta R_{ij}^2$ , can be calculated in parallel but require  $N(N - 1)$  multiply operations. For  $N$  approaching 100 this would saturate the **DSP** units available on the target Xilinx UltraScale+ **FPGAs**. These calculations therefore have to be done in stages, increasing latency.

Minimisation of the *Separation List* is done in two stages in this scheme to avoid large  $\mathcal{O}(N^2)$  binary minimisation trees. In the first stage, the minima of each the *Distance Lists* ( $\Delta R_{ij}^2$ ) are found in parallel. Then, these distance minima are multiplied by the relevant  $p_{T,i}^{-2}$  and the minimum of that list is found. This identifies the two objects to be merged, or the jet to be removed. A large **LUT** must be used to transform the merged object 3-momentum into  $p_T^{-2}, \phi, \eta$ , and  $2(N - 2)$  multipliers are required to calculate the distances  $\Delta R_{kj}^2$  between the new object  $k$  and the remaining objects in the list. This algorithm therefore uses a significant amount of **LUT** and **DSP** resources for even moderate values of  $N$ . Assuming realistic synchronisation/timing requirements, the latency per iteration in this scheme is:  $\log_2(N - 1) + \log_2(N + 1) + 11$ . At a 320 MHz clock speed the total latency exceeds  $2.5 \mu\text{s}$  for the very modest value of  $N \sim 32$ . Taken together with the significant resources required to implement the algorithm, this disqualifies the *E-merging* version of the *Exact anti- $k_t$*  algorithm as a viable option.

Thus, the goal becomes identifying tradeoffs in precision that could balance the latency restriction. Several modifications to the basic anti- $k_t$  algorithm have been proposed to improve its computational efficiency. Two that are of particular relevance to the anti- $k_t$  implementation on the **GEP** are: the **Winner-Take-All (WTA)** merging scheme [9.6], and the use of **Tiling** [9.7]. In the **WTA** merging scheme computational resources are saved by simply

summing the  $p_T$ 's of the two objects and assigning the  $\phi, \eta$  position of the initial object with the larger  $p_T$  to the merged object. Tiling effectively reduces the number of objects to consider ( $N$ ) by dividing the area over which jet-finding is performed into smaller, overlapping tiles. In the following, we explore the use of the [WTA](#)-merging scheme, though Tiling is not ruled out as a future option.

The fact that object positions are not changed in the [WTA](#) merging scheme makes this method much more amenable to implementation in the [GEP](#) than the [E](#)-scheme. This has three consequences. First, if the *Distance Lists* ( $\Delta R_{ij}^2$ ) are sorted before the first iteration, these sorted lists can be used unchanged (simply disabling removed objects) for all iterations. Using a parallel sorting algorithm (e.g., bitonic mergesort),  $(\log_2 N)(\log_2(N) + 1)/2$  stages + timing/synchronisation overhead are required to do each  $N$ -element sort, all of which can be done in parallel. This removes one of the two minimisation stages per iteration required in the *Exact anti- $k_t$  - E-merging* scheme. Second, the [LUT](#) needed to calculate  $p_T^{-2}$  of a merged object is reduced in size compared to that needed in the *Exact anti- $k_t$  - E-merging* scheme. Finally,  $\Delta R_{kj}^2$  does not have to be calculated for a new merged object at each iteration because the position of the merged object is the same as one of its constituents. Thus both latency per iteration and [FPGA](#) resources required are reduced significantly in this scheme with respect to the *Exact anti- $k_t$  - E-merging* scheme. An estimate of the latency per iteration in the *Exact anti- $k_t$  - WTA* scheme has been made using known latencies for each of its component blocks (minimisation, [LUT](#), etc) and including conservative assumptions of additional delays required to meet timing and synchronisation requirements. The dependence of this latency on the number of input clusters ( $N$ ) is found to be  $\log_2 N + \log_2(N + 1) + 8$  per iteration, and is illustrated in [Fig. 9.12](#).

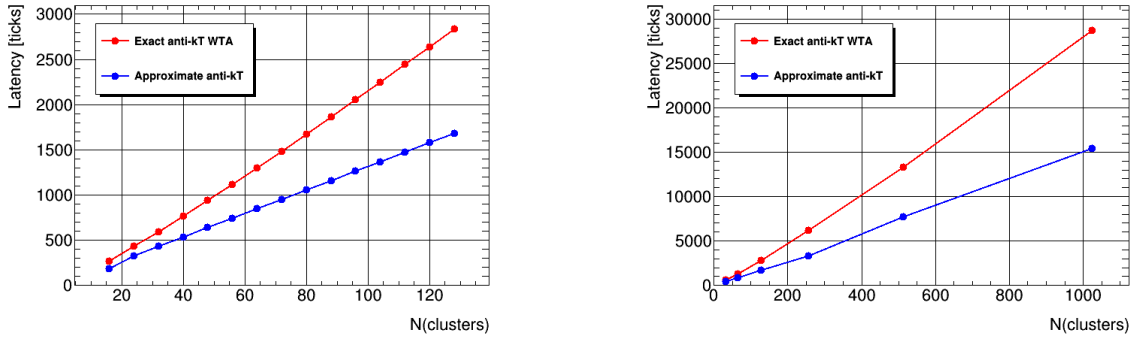


Figure 9.12: Latency of the prototype anti- $k_t$  algorithm variants in units of [FPGA](#) clock cycles (ticks) for a nominal 320 MHz reference clock as a function of the number of clusters being processed ( $N$ ). The right plot is a zoomed out version of the left plot.

Significant latency and resource savings can be realised in an *Approximate anti- $k_T$*  scheme where the criterion for clustering is modified from that specified for anti- $k_t$ . In this case, the list of object momenta ( $p_{T,i}$ ) as well as the *Distance Lists* are sorted before iterations start, and the closest object to the highest  $p_T$  object is merged with it unless  $\Delta R_{ij}^2 > R^2$  in which

case the highest  $p_T$  object is moved to the *Jet List*. The *WTA* scheme is used for merging. At each iteration the sorted  $p_T$  *List* is used to reference the *Distance List* corresponding to the maximum  $p_T$  object remaining. The minimum element of that (sorted) *Distance List* is then used to define the object to merge with the highest  $p_T$  object (or to promote the highest  $p_T$  object to the *Jet List*). No minimisations are required, and  $p_{T,i}^{-2}$  does not need to be calculated in this scheme, meaning that no *LUTs* are needed. The latency per iteration is therefore approximately independent of  $N$ . It is estimated to be  $2 \text{ ceil}[\log_2(\log_2 N)] + 7$  per iteration based on the mathematical and list-management operations that are required at each iteration.

An estimate of the latency per iteration in each of the above schemes has been made by summing the known latencies for individual element blocks in each scheme (minimisation, sorting, multiplies, look-ups, etc) and adding synchronisation/timing delays between blocks. A comparison of input and latency estimates for the various jet-finding schemes is summarised in Table 9.2. *FPGA* resource requirements, particular the number of *LUTs* and *DSP* units needed, decrease substantially when moving from the *Exact anti- $k_t$  – E-merging* to the *Exact anti- $k_t$  – WTA* to the *Approximate anti- $k_t$*  implementations. Although precise numerical estimates of resource usage are not yet available, it is apparent that the *Exact anti- $k_t$  – E-merging* variant is not a viable option due to latency considerations given the number of input clusters expected even in the most aggressive Tiling schemes. However, the *Approximate anti- $k_t$*  implementation appears to be feasible in terms of latency and *FPGA* resources, and assuming Tiling schemes will allow a reduction of  $\langle N \rangle$  to less than roughly 100. Studies of further optimisations in jet finding algorithms and their physics performance are ongoing.

Table 9.2: Comparison of the parameters used and latencies (in *FPGA* clock cycles for a nominal 320 MHz reference clock) for the initial sorting and per iteration estimated for various jet-finding schemes as a function of the number of clusters ( $N$ ) sent from the clustering step. Index  $i$  refers to individual cluster objects and runs from 1 to  $N$ .

Scheme	From clustering	Latency for initial sort	Latency per iteration
Exact anti- $k_t$ – E-merging	$p_{x,i}, p_{y,i}, p_{z,i}$ $p_{T,i}, p_{T,i}^{-2}$ $\phi_i, \eta_i$	–	$\log_2(N - 1) + \log_2(N + 1) + 11$
Exact anti- $k_t$ – <i>WTA</i>	$p_{T,i}, p_{T,i}^{-2}$ $\phi_i, \eta_i$	$\log_2[N(N + 1)/2]$ $+ \text{ceil}[\log_2(\log_2 N)] + 5$	$\log_2 N + \log_2(N + 1) + 8$
Approximate anti- $k_t$	$p_{T,i}$ $\phi_i, \eta_i$	$\log_2[N(N + 1)/2] + 5$	$2 \text{ ceil}[\log_2(\log_2 N)] + 7$

The inputs to these algorithm are not yet considered by this study. The list of topoclusters described in the previous subsection is a natural choice. However, there are potentially many topoclusters in each event which has a significant impact on the overall latency as shown in Fig. 9.12. Several methods are being studied to reduce the overall multiplicity of topoclusters including implementation of a Tiling procedure, applying minimum  $p_T$

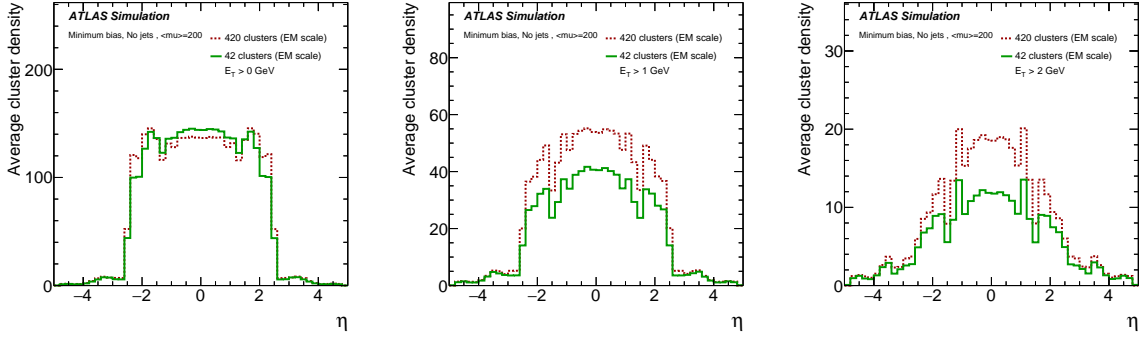


Figure 9.13: Average topocluster number density  $\langle \delta N_{clus} / \delta \eta \rangle$  as a function of  $\eta_{clus}$ , for clusters with  $p_{T,clus} > p_{T,min}$ , for various  $p_{T,min}$  values. Results are obtained using QCD dijet samples with  $\langle \mu \rangle = 200$ .

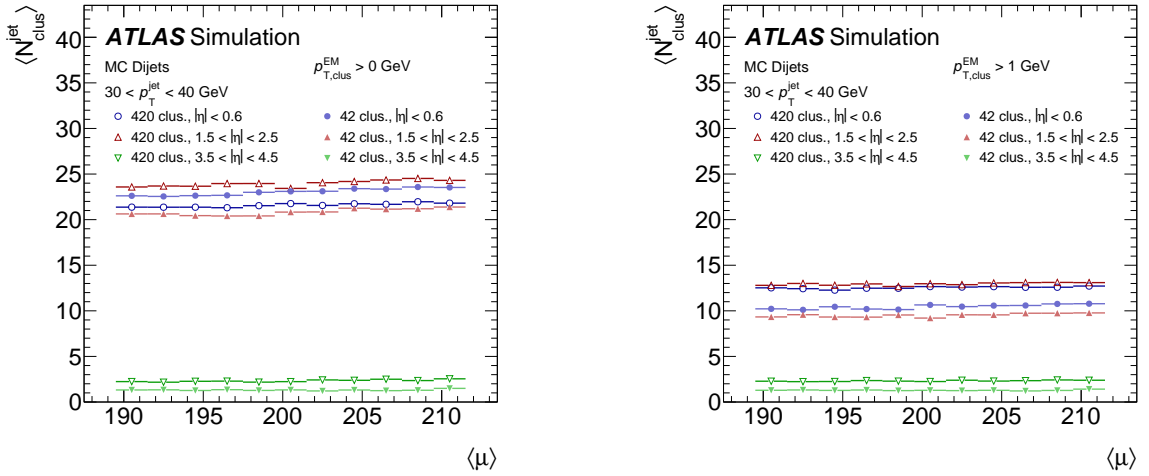


Figure 9.14: The average number of topo-clusters  $\langle N_{clus}^{jet} \rangle$  in anti- $k_T$  jets reconstructed with  $R = 0.4$  within  $30 < p_{T,jet} < 40$  GeV as a function of  $\mu$  using QCD dijet samples with  $\langle \mu \rangle = 200$  (left), and when selecting topo-clusters by  $E_{T,clus} > 1$  GeV inside jets (right).

requirements to the topoclusters, only running jet algorithms on topoclusters within **RoI** defined by the jets identified in the **jFEX** or the **gFEX**, or by building towers from the topoclusters (so called **TopoTowers**). All of these options will be studied as part of the optimisation of the algorithm performance and resource usage. As an initial indication, studies of the impact of a minimum  $p_T$  requirement were carried out for the offline [9.2]. Even a moderate  $p_T$  requirement of 1 or 2 GeV can reduce the overall number of topoclusters by a factor of 50 or more as shown in Fig. 9.13. Such cuts have benefit by reducing low-energy topoclusters that can increase sensitivity to pileup as shown, for example in Fig. 9.14.

Such low-energy topoclusters are typically removed during studies of the substructure of large-radius jets including those that contain the decays of highly Lorentz-boosted particles

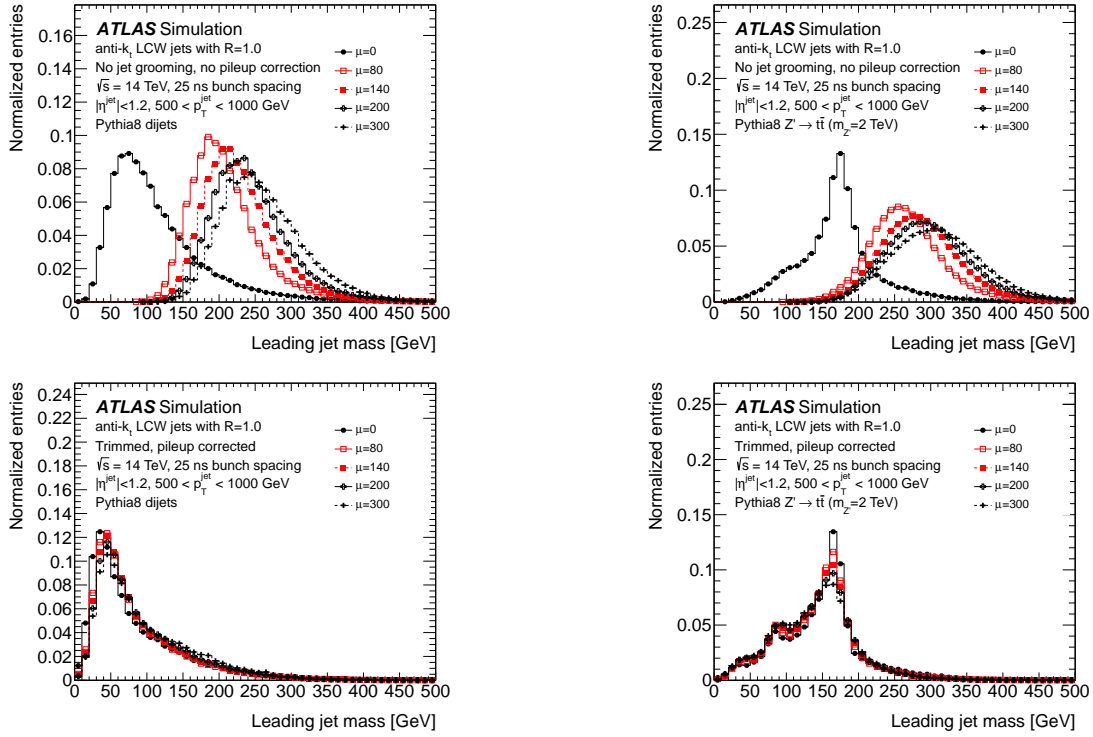


Figure 9.15: The anti- $k_t$   $R = 1.0$  leading jet mass distribution before (top) and after (bottom) jet trimming and pileup subtraction in dijet events and  $Z' \rightarrow t\bar{t}$  (right) events with mean number of interactions per bunch crossing ( $\mu$ ) of 0 (black closed circles), 80 (red closed squares), 140 (red open squares), 200 (black open crosses), and 300 (black closed crosses). Jets with  $|\eta| < 1.2$  and  $0.5 < p_T < 1$  TeV are used. In the bottom plots, jets are trimmed with parameters  $f_{\text{cut}} > 5\%$  and  $R_{\text{subjet}} = 0.3$  and are corrected for pileup using area-based pileup subtraction.

such as  $W$ ,  $Z$ ,  $H$  bosons, top quarks, and various exotica. One popular pileup suppression technique in such jets is called *trimming* where low- $p_T$  topoclusters in small jets (subjet radius  $R_{\text{subjet}} = 0.3$ ) that make up less than 5% of the energy of the large-radius jet ( $f_{\text{cut}} > 5\%$ ) are removed. An example of the signal improvements from such a technique is shown in Fig. 9.15. On the left are results from QCD dijets and on the right signals from  $Z' \rightarrow t\bar{t}$ . The upper plots are without any trimming while the lower plots include it. Trimming low  $p_T$  contributions is effective in removing the impact of pileup and improving the separation between signal and background. Further physics performance is discussed in Section 6.6.



---

## 9.4 Physical Realisation

### 9.4.1 Architecture

This section describes the design of the [Global Trigger](#) system, a multi-layer hardware scheme to multiplex, process, and demultiplex the input data and output trigger information. Use of a common hardware design throughout minimises the complexity of the firmware and eases long-term maintenance; this design being the Global Common Module ([GCM](#)). This section motivates the required number of [GCMs](#) at each layer of the global processing chain and describes the requirements to support the outside interfaces to the [Global Trigger](#), as well as internal interfaces between [GCMs](#), concluding with a description of a proof-of-principle implementation of the [GCM](#) design.

Each layer in the system consists of a number of processing nodes, where each node is based on a large [FPGA](#) with extensive input and output capabilities. Previous experience on ATLAS designing ATCA modules and on CMS designing  $\mu$ TCA modules, along with consideration of the power and cooling environment within the USA15 racks, suggests ATCA as a suitable form factor with a single ATCA Front Board supporting two independent nodes along with shared infrastructure. The [GCM](#) design is therefore an ATCA Front Board with two large [FPGAs](#), where each [FPGA](#) can be configured as a [MUX](#), [GEP](#), or [CTP Interface](#) node as required, along with a central processing chip (such as a ZYNQ MPSoC) to allow for monitoring, control and readout. A schematic block-diagram view of the [Global Trigger](#) is given in Fig. 9.16. This figure demonstrates conceptually how the I/O interfaces, multiplexing, event processing, and demultiplexing can be achieved with a single-module design. Each step requires a significant number of optical links, onboard processing capabilities and, central control.

### 9.4.2 Required Number of Global Common Modules (GCMs)

There are two drivers on the number of nodes in the [GCM](#) design and its implementation in the [Global Trigger](#): first, the number of [MUX](#) inputs (both the number of inputs per [MUX](#) and the number of [MUX](#) modules) and second, the number of required [GEP](#) modules, which defines the event pipeline. The number of real-time input fibres to the system add up to 2312 in the baseline design as summarised in Table 9.3. For symmetric input and output link speeds the number of input links on a [MUX](#) node should match the number of outputs, this latter being the number of [GEP](#) nodes. Balancing the number of [MUX](#) nodes and [GEP](#) nodes to keep the data flow in the multiplexing mesh network symmetric then gives 48 [MUX](#) nodes and [GEP](#) nodes to symmetrically distribute the 2312 input links.

High density optical connectors, a prerequisite for the design, target 12-fibre ribbons so there is good sense in keeping either the number of [MUX](#) nodes or [GEP](#) nodes a multiple

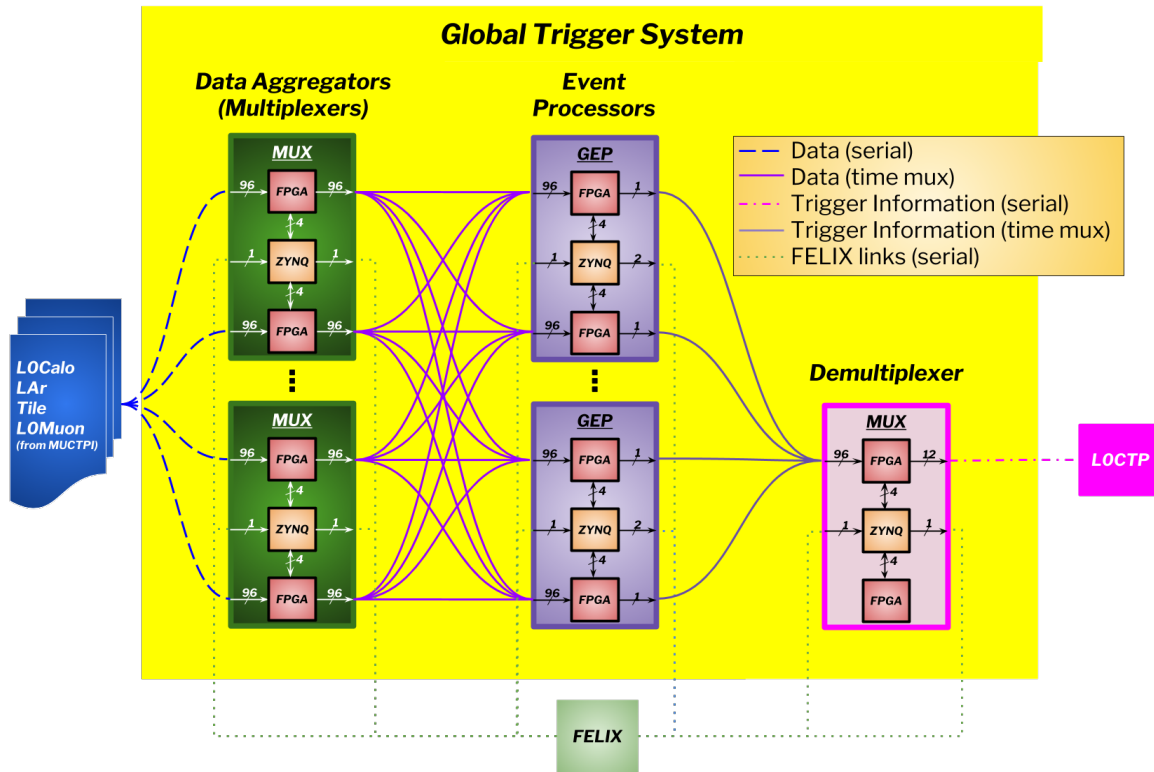


Figure 9.16: Schematic diagram of the Global Common Module physical design and interfaces.

of 12. Allowing the possibility of additional inputs, most particularly to allow the evolution described in Chapter 14 leads to the baseline design having 48 GEP nodes (24 GEP modules). With a large fraction of the input links running at speeds lower than the internal multiplexing links, this results in the 47 MUX nodes (24 MUX modules) listed in Table 9.3, leaving 1 MUX node spare in the baseline fibre plant. Further inputs can be accommodated by increasing the number of MUX nodes above 48 and keeping the number of GEP nodes at 48. To this end the ribbons into the GEP nodes foresee a maximum 72 MUX nodes.

### 9.4.3 Required Number of Links on each GCM

The three use cases of the GCM, as MUX, GEP and as CTP Interface, each bring different requirements on the number of input and output links.

For the MUX the maximum input loading on a node in Table 9.3 is 64. Mapping this onto 12-fibre ribbons suggests up to 72 as the number of inputs per node. Each node also requires an output fibre to each GEP node, which would give 48 output fibres per node in the baseline design. With muon information being on the critical path for latency (see Section 9.5 below) there is mileage in supporting double links from the MUCTPI MUX nodes,



Table 9.3: Summary of the real-time optical fibre interfaces to the *Global Trigger* from input data sources and output to *CTP*.

Input source	Link Speed (Gb/s)	Expected fibres	fibres/node	Nodes
eFEX	11.2	384	64	6
jFEX	12.8	192	48	4
fFEX	12.8	48	48	1
gFEX	11.2	24	24	1
MUCTPI	9.6 - 12.8	96	48	2
LASP	25.8	1484	48	31
TPPr	11.2	64	64	1
ZDC	11.2	20	20	1
TOTAL		2312		47
Output destination				
CTP	9.6	12	12	1

or alternatively sharing the *MUCTPI* links over more nodes to reduce the multiplexed data transfer time. The minimum number of outputs per node is then 48, with the possibility of 96 being desirable.

For the *GEP* the number of inputs on a node in the baseline design is 47. In the two-level Level-0/Level-1 system there will be additional inputs from *L1Track* and from the Level-0 *CTP* which will exhaust the number of links on 4 12-fibre ribbons. Future-proofing then gives 60 or 72 links as the minimum number of inputs. Outputs per node are minimal, in the baseline a single fibre to the *CTP Interface*, possibly two to reduce latency on the critical path. Even in the two-level Level-0/Level-1 system the number of outputs will still fit on a single 12-fibre ribbon.

For the *CTP Interface* there is a single input from each *GEP* node, so 48 in total. Again there could be mileage in having dual inputs from each *GEP* node to reduce latency, so minimum number of inputs is 48 with the possibility of 96 being desirable. Output is a single 12-fibre ribbon to the *CTP*. The second half of the *CTP Interface* module could be idle, or latency considerations might split this functionality across both nodes, but in any case in the two-level Level-0/Level-1 system this second node could drive the outputs to the Level-1 *CTP*.

In addition to the above inputs and outputs per node, each *GCM* module has inputs from the *TTC* system via *FELIX* and readout outputs to *FELIX*. This amounts to an additional input ribbon and output ribbon per module. The number of fibres per module is summarised in Table 9.4. The minimum number of inputs that the generic *GCM* module needs to support is therefore 144+12 and outputs 96+12, in both cases with 192+12 a desirable maximum. Whilst the *GCM* module would be built with all the necessary sockets for the optical I/O only those needed would be populated on each actual module.

The same type of fibre management as on the Phase-I FEX modules is proposed. This consists of ribbons on-board bundled together and routed out through optical backplane connectors in ATCA Zone 3, thereby affording an element of protection to the delicate fibre infrastructure. The connectors used on the FEX modules will allow up to 24 ribbons, 288 fibres, to be routed in this manner, more than enough for any of the combinations of inputs and outputs foreseen.

Table 9.4: Summary of the maximum number of optical fibre interfaces for the GCM modules.

Module	Fibre	Size	Number
<b>Multiplexer</b>			
	Input from Source	$2 \times 72$	144
	Output to GEP	$2 \times 48(96)$	96(192)
	FELIX	$2 \times 12$	24
<b>Event Processor</b>			
	Input from MUX	$2 \times 60$	120
	Output to CTP Interface	$2 \times 12$	24
	FELIX I/O	$2 \times 12$	24
<b>CTP Interface</b>			
	Input from GEP	$2 \times 48(96)$	96(192)
	Output to CTP	$1 \times 12$	12
	FELIX I/O	$2 \times 12$	24

#### 9.4.4 GCM Requirements and Preliminary Design

Figure 9.17 contains a preliminary block diagram for the GCM. There are two primary features. First, each board would host two large FPGAs that are capable of providing a sufficient number of MGTs to support the optical interfaces as well as the required event processing/multiplexing demands. And second, each board would host a central processing chip (such as a ZYNQ MPSoC) to allow for monitoring, control and readout. In addition to the internal Random Access Memory (RAM) each of these chips would be interfaced with large RAM storage to allow for long-latency buffering of the input and output data. The strategy, as shown in Fig. 9.16, is that each FPGA would serve as one MUX or one GEP node, or as the CTP Interface. This choice defines the requirements for FPGAs to be utilised.

The proof of concept architecture sketched in Fig. 9.17 assumes 2017 technology to allow for realistic costing and design strategy. To satisfy the required number of optical links and on-board MGT interfaces, two candidate FPGAs were identified: the Xilinx Ultrascale+ VU9P and VU13P (each provides up to 104 MGTs capable of 30+ Gb/s serial transmission in the C2104 variant). The candidate for the control/readout chip was a Xilinx UltraScale+ ZYNQ (e.g. ZU17EG). The GCM module would require maximally 204 optical inputs and

204 optical outputs. These requirements can be achieved using 12-channel optical modules (e.g. Avago MiniPODs), which would require the board to host 34 optical modules. This figure further illustrates that an optional mezzanine could be included, taking advantage of 28 MGT links from the ZYNQ chip and allowing additional processing or readout capacity in specialised use cases. However, in the baseline design there is no use case for such a mezzanine.

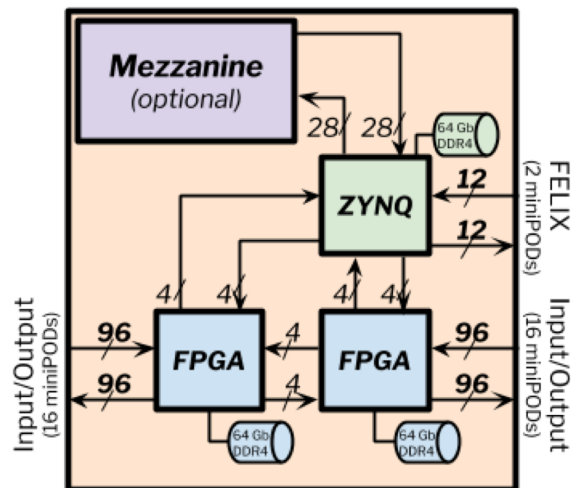


Figure 9.17: Block diagram for the Global Common Module board.

The GCM design is a natural evolution from the Phase-I gFEX hardware, and a proof of principle layout is shown in Fig. 9.18. This example design is based on a standard ATCA form factor and includes realistic footprints for the ATCA power entry and conversion, UltraScale+ and ZYNQ chips, IPMC mezzanine, miniPODs, GbE interfaces and other required components. Each of the chips are highlighted along with their primary board resources. This exercise demonstrates the basic feasibility of such a design within the ATCA form factor.

Building on the experience with the Phase-I eFEX and gFEX hardware the MPV power estimate is 350 W for the MUX variant and 375 W for the GEP variant, well within the 400 W slot limit (see Section 4.2). Experience with the eFEX cooling suggests that maintaining sensible FPGA temperatures in a single slot will not be difficult with suitable airflow management and heatsink design.

### 9.4.5 Fibre Plant

The same modular scheme as already successfully implemented in the CMS trigger, shown in Fig. 9.19, is proposed for the fibre plant. The 48-fibre bundle output from a MUX module is split out into 4 12-fibre ribbons and each ribbon is plugged into a separate  $12 \times 12$  fibre

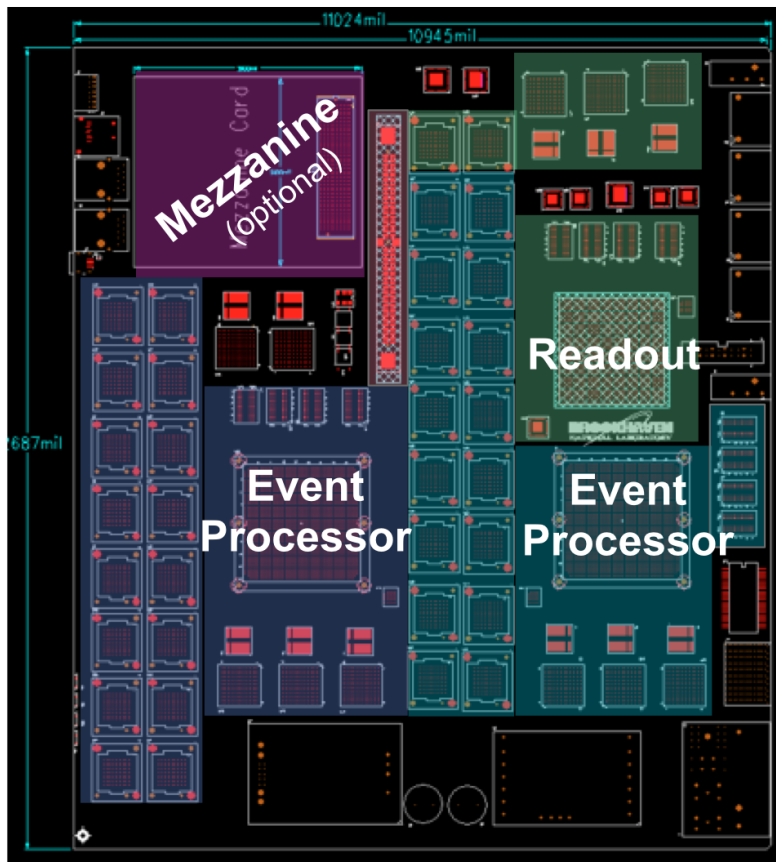


Figure 9.18: Example layout for the *GCM* board using the standard *ATCA* form factor.

ribbon remapping module. The 72-fibre bundle input to a *GEP* module is similarly split into 6 12-fibre ribbons (2 additional ribbons to allow expansion in number of *MUX* nodes above 48) and plugged into the outputs from the appropriate remapping modules, with  $4 \times 4$  remapping modules providing the full  $48 \times 48$  mesh between *MUX* and *GEP* systems. A separate custom module can provide the concentrator mapping between *GEP* and *CTP Interface*.

#### 9.4.6 Production Firmware Deployment Module

Modern firmware simulation tools are essential but not sufficient to guarantee fail-safe performance of the firmware on the target hardware system. Therefore, in order to evaluate various firmware blocks of the *Global Trigger* system a dedicated hardware platform is required by each group involved in the firmware development. For this purpose a number of Production Firmware Deployment Modules (*PFM*) will be produced and distributed to firmware developers. The *PFM* represents a slice of the *GCM*, which includes a processing unit, a control *FPGA* and a number of optical modules. Such a configuration of the *PFM*



Figure 9.19: Dual  $12 \times 12$  fibre ribbon remapping modules in a 2U enclosure installed in a standard LHC rack in Point-5. 8 such dual modules will be required on ATLAS.

allows testing and debugging algorithm and infrastructure firmware together with the corresponding software, thus minimising the risk of firmware failures on the final system. In order to allow for early firmware development and evaluation, the [PFM](#) needs to be available well before the production [GCM](#). With the production system in place, the [PFM](#) system will continue to be maintained for pre-commissioning of regular firmware upgrades during future operations.

## 9.5 Latency

The latency of the [Global Trigger](#) system has two primary components: the latency for the delivery of data inputs to the [MUX](#) system and that of processing and transmission of the [Global Trigger](#) decisions. As the arrival time of the muon information is significantly later than all other inputs this can be further factorised into the latency path for trigger signatures not depending on muons and those involving muon information. For the former these can be broken down as:

- arrival of data at [MUX](#)
- sorting and transmission of data from [MUX](#) to [GEP](#)
- contemporaneous spatially pipelined processing of incoming data, including topoclustering
- processing of trigger signatures which do not depend on jets or muons
- asynchronous/iterative jet algorithms
- processing of jet-based trigger signatures which do not depend on muons

In addition preprocessing can occur for those signatures which do require muon information, ready for merging with the muon information on its arrival, an obvious candidate being  $E_T^{\text{miss}}$  where the bulk of the vector sums can be formed in advance.

The latency critical path involving muons through to transmission to the **CTP** can be similarly broken down as:

- arrival of muon information at **MUX**
- sorting and transmission from **MUX** to **GEP**
- contemporaneous processing of incoming data, such as muon isolation
- completion of trigger signatures depending on muons
- assemble **Global Trigger TIP**
- serial transmission of **Global Trigger TIP** from **GEP** to **CTP Interface**
- serial transmission of **Global Trigger TIP** from **CTP Interface** to **CTP**

The best estimates of each of these components is summarised in Table 9.5, illustrating both the individual components and also the total latency for each step following the same prescription for **CBE** and **MPV** as in Section 5.2.8 and Table 5.5.

### 9.5.1 Data arrival time at Multiplexer Processor

The **LAr** [9.4] and **Tile** [9.8] full-granularity inputs are anticipated to arrive in  $1.425 - 1.7 \mu\text{s}$  **CBE** to **MPV**, followed by the **L0Calo** inputs ( $2.075 - 2.575 \mu\text{s}$ , see Section 7) and the **L0Muon** inputs at  $4.46 - 5.26 \mu\text{s}$  (see Section 8). The difference in arrival times between the full-granularity calorimeter inputs and the **L0Muon** inputs provides a significant period for calorimeter pre-processing in the shadow of the muon latency.

### 9.5.2 Processing latency and budget

A key driver in the **Global Trigger** latency is the time taken to transfer data from the **MUX** to the **GEP**. In the case of the **LAr** **LASP** the link bandwidths are matched, so the transfer time is **48 BC** ( $1.25 \mu\text{s}$ ) in addition to the latency of transmission and encoding. In the case of the muon information the input bandwidth is significantly less, so this transfer time is just over  $0.5 \mu\text{s}$ .

The contemporaneous processing of incoming data adds  $0.1$  to  $0.2 \mu\text{s}$  in each case. The time for processing of trigger signatures which do not depend on jets or muons, along with the time for iterative or asynchronous jet processing, remains hidden in the shadow of the muon latency.

A window of  $1.25 \mu\text{s}$  **CBE** ( $2.50 \mu\text{s}$  **MPV**) on jet processing processing leaves a safety margin of around  $1 \mu\text{s}$  on the processing of trigger signatures which do depend on jets before there is any impact on overall latency, so the overall latency is always dictated by the muon

arrival time. Here there is an overall margin of  $0.9 \mu\text{s}$  MPV in the Level-0 latency as shown in Table 5.5.

Table 9.5: *Global Trigger* latency estimate, which begins with the latency of input signals and steps through processing to the delivery of *TIP* to the *CTP*.

Category	Description	CBE ( $\mu\text{s}$ )	$\int$ CBE	MPV ( $\mu\text{s}$ )	$\int$ MPV
Inputs					
	Calorimeter	1.425		1.70	
	L0Calo	2.08		2.58	
	L0Muon	4.46		5.26	
Calo path					
	signals present	1.425	-	1.70	-
	MUX	1.400	2.83	1.500	3.20
	pipelined processing	0.100	2.93	0.200	3.40
	non-jet triggers	0.200	3.13	0.250	3.65
	jet processing (in parallel)	1.250	4.18	2.500	5.90
	non-muon triggers	0.200	4.38	0.250	6.15
Muon path					
	signals present	4.46	-	5.26	-
	MUX	0.650	5.11	1.250	6.51
	pipelined processing	0.100	5.21	0.200	6.71
	muon triggers	0.200	5.41	0.250	6.96
<i>TIP</i>					
	<i>TIP</i> assembly	0.050	5.46	0.100	7.06
	serial transmission	0.250	5.71	0.300	7.36
	link to <i>CTP</i>	0.150	5.86	0.200	7.56

## 9.6 Firmware

As described above in this Chapter, *Global Trigger* system will consist of a *GCM* being used in three distinct instances: as a multiplexing module, as an event processing module and as a demultiplexing module. These three instances will require dedicated firmware to enact the required functionality. Figure 9.2 illustrates a conceptual view of the overall data flow, firmware blocks and firmware interfaces for the *Global Trigger*. Each of the three use case instances require common firmware for the basic low-level frame, command/control and monitoring. Each component will also require an individual configuration interface and management system. In addition, the ZYNQ command/control chip will require specialised firmware and software to support board control and readout.



### 9.6.1 Firmware Blocks

There are several features of the [Global Trigger](#) system firmware architecture that naturally align to firmware blocks, and are described herein below.

**Common Module Firmware:** There are several common firmware aspects required for the [Global Trigger](#).

1. Each [GCM](#) requires a low-level frame firmware that provides a base environment for the management of FPGA interfaces: reference clocks, I/O interfaces, and general infrastructure. This firmware would be common to all FPGAs and ZYNQ, but would require a dedicated synthesis for each case.
2. The ZYNQ chip on each board will provide command, control, configuration and synchronisation signals for each processing FPGA. A firmware communication block must be developed to support these functions.
3. The ZYNQ chip will provide monitoring access for the data sent from the processing FPGAs. Thus firmware to provide an interface to the monitoring data is required to accept data from the FPGA/ZYNQ interface, process the data and transfer data to the ZYNQ's software applications.
4. The ZYNQ chip will also be required to manage and monitor the [GCM](#) board resources. Thus, firmware to manage board-level functions will be required.
5. All [FELIX](#) and monitoring interfaces for the [GCM](#) are handled on the ZYNQ chip, so a data flow interface between each processor FPGA and the ZYNQ is required. This interface provides data transmission from the FPGAs (Tx) to the ZYNQ (Rx) using common protocol, which must be adapted for each [GCM](#) use case.
6. The [GCM](#) will have significant internal and external [RAM](#) resources to allow buffering in the multiplexing, data handling and readout steps. The interface to this [RAM](#) will be common for each [GCM](#) use case, though the details of how data will be utilised is specific to each use case. Thus the firmware for the [RAM](#) interface will be developed as a common module.

**Data Aggregation and Multiplexing Firmware:** The [MUX](#) module will require firmware to receive input data from the detector and trigger sources, unpack the data, perform the time multiplexing and to transmit the multiplexed data to the [GEP](#) modules. The [GEP](#) will further require a version of the [MUX/GEP](#) interface to receive and decode the data.

**Trigger Signature Firmware:** The [Global Trigger TIP](#) will be derived from a series of trigger algorithms and hypotheses processed by the [GEP](#). As described in Section 9.3, these algorithms will be firmware-based with performance requirements that allow a series of algorithms to run simultaneously in parallel. The trigger firmware will be divided into two classes: algorithms and hypotheses. The algorithms will process input data to refine trigger objects and signatures. The hypotheses will group these objects and signatures to be tested against pre-defined thresholds.



**Trigger Framework Firmware:** The trigger signature firmware will require a coordinated scheme to deliver input data, sequence algorithms, collect and package algorithm outputs, and collate hypothesis results. These needs will be provided via a trigger framework that also incorporates a memory handling block. The memory handler must interface with the MUX, the core trigger framework and the ZYNQ firmware. Furthermore, the trigger framework must provide trigger information to the GEP/CTP interface firmware.

**Central Trigger Interface Firmware:** The [CTP Interface](#) must receive trigger information from the [GEP](#), time-demultiplex the data and transmit the demultiplexed results to the [CTP](#). Firmware to support these functions will be implemented on the [CTP Interface](#) and the [GEP](#), which requires firmware to package and transmit the [TIP](#) information from the [GEP](#).

**FELIX Interface Firmware:** The communications with [FELIX](#) are handled by the ZYNQ chip. The [FELIX](#) interface must provide firmware to transmit formatted trigger data to [FELIX](#) and to receive trigger control ([TTC/LTI](#)) information to be transmitted to all [Global Trigger](#) system components. The [FELIX](#) interface will also require firmware to unpack the trigger control information. Finally, there must be a firmware block that provides readout formatting, interfaces with data I/O ([FELIX Rx/Tx](#), monitoring, FPGA input, and the [RAM](#)).

**On-Board Software:** The ZYNQ chip will be responsible for communications to the outside general network, which will allow for monitoring and command/control processes. Thus there will be software required for IPbus network communication, management and distribution of command/control signals, and monitoring applications.

## 9.6.2 FPGA Resource Usage

The critical [FPGA](#) resource usage is on the [GEP](#). As illustrated in [Table 9.1](#) the resource usage for topoclustering will be well below 1% of the current Xilinx VU13P, with a similar number for the  $e/\gamma$  and  $\tau$  spatially pipelined algorithms. The resource usage will therefore be dominated by the jet finding and trigger algorithms and hypotheses.

An estimate for the latter can be made from the existing [L1Topo](#) usage, where the worst case usage for each of the 4 Virtex7 FPGA is between 36.5% and 68.3%. Allowing for double counting of common logic (each [L1Topo](#) FPGA performs the same sorting) and scaling for the different size of the FPGA gives an estimate of some 30% usage for this logic in the Xilinx VU13P device in the proof of concept architecture.

This choice of FPGA would therefore allow some 20% of the device for the jet finding algorithm, a realistic target. It is reasonable to expect that a new FPGA generation will be available at the time of the final board design, where historically the maximum FPGA logic availability doubles between generations, allowing significantly more resource for the jet finding.

### 9.6.3 Firmware Development

Preliminary studies of the algorithm and framework firmware have begun (e.g., Section 9.3) and will continue during the system design process. Initial development and offline testing of new algorithms will take place on commercial development boards and the Production Firmware Deployment Module described in Section 9.4.6. Alongside the firmware development will be both algorithmic and bitwise simulation.

Initial validation will be in the Surface Test Facility with final validation in Point-1. It will be possible to deploy additional GEP modules receiving duplicates of events in the live system to allow parasitic testing of new firmware alongside the production system. Evolution of the firmware will be controlled by a rigorous change control mechanism, such as that already used for the Run 2 L1Topo [9.9].

Steering of thresholds and selection of algorithms within each firmware build for the Global Trigger will be defined in the trigger menu and the change control will be defined within that context.

## References

- [9.1] A. Zabi et al., *The CMS Level-1 Calorimeter Trigger for the LHC Run II*, Journal of Instrumentation **12** (2017) no. 01, C01065.  
<http://stacks.iop.org/1748-0221/12/i=01/a=C01065>.
- [9.2] ATLAS Collaboration, G. Aad et al., *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, *Eur. Phys. J.* **C77** (2017) 490, [arXiv:1603.02934](https://arxiv.org/abs/1603.02934) [hep-ex].
- [9.3] ATLAS Collaboration, G. Aad et al., *Jet energy measurement with the ATLAS detector in proton-proton collisions at  $\sqrt{s} = 7$  TeV*, *Eur. Phys. J.* **C73** (2013) no. 3, 2304, [arXiv:1112.6426](https://arxiv.org/abs/1112.6426) [hep-ex].
- [9.4] ATLAS Collaboration, *Technical Design Report for the Phase-II Upgrade of the ATLAS LAr Calorimeter*, Tech. Rep. CERN-LHCC-2017-018. ATLAS-TDR-027, CERN, Geneva, Sep, 2017. <https://cds.cern.ch/record/2285582>.
- [9.5] M. Cacciari, G. P. Salam, and G. Soyez, *The Anti- $k(t)$  jet clustering algorithm*, *JHEP* **04** (2008) 063, [arXiv:0802.1189](https://arxiv.org/abs/0802.1189) [hep-ph].
- [9.6] A. J. Larkoski and J. Thaler, *Aspects of jets at 100 TeV*, *Phys. Rev.* **D90** (2014) no. 3, 034010, [arXiv:1406.7011](https://arxiv.org/abs/1406.7011) [hep-ph].
- [9.7] M. Cacciari, G. P. Salam, and G. Soyez, *FastJet User Manual*, *Eur. Phys. J.* **C72** (2012) 1896, [arXiv:1111.6097](https://arxiv.org/abs/1111.6097) [hep-ph].

- [9.8] ATLAS Collaboration, *Technical Design Report for the Phase-II Upgrade of the ATLAS Tile Calorimeter*, Tech. Rep. CERN-LHCC-2017-019. ATLAS-TDR-028, CERN, Geneva, Sep, 2017. <http://cds.cern.ch/record/2285583>.
- [9.9] K. Nagano and H. J. Stelzer, *L1Topo Menu Update Procedure*, Tech. Rep. ATL-COM-DAQ-2017-100, CERN, Geneva, Aug, 2017. <https://cds.cern.ch/record/2280269>.



## 10 Central Trigger System

The Central Trigger System consists of the Central Trigger Processor (CTP), the Muon-to-CTP-Interface (MUCTPI), and the distribution of Trigger, Timing, and Controls signals (TTC) to the sub-detector readout systems, of which the Local Trigger Interface modules (LTI) are an integral part. Figure 10.1 illustrates its various components in context. While

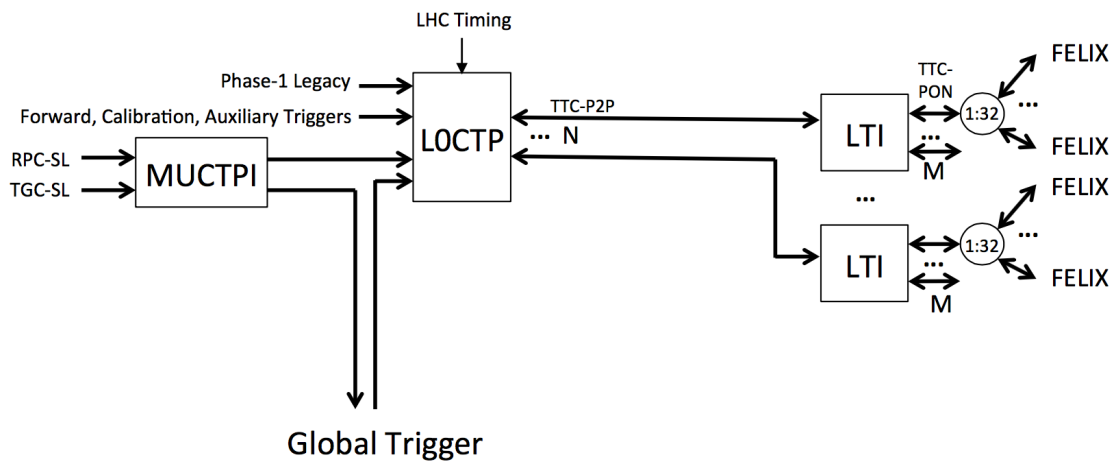


Figure 10.1: Context diagram showing the Level-0 CTP system interfacing to the Global Trigger and FELIX systems (via the LTI modules). Only the trigger path is depicted; control, monitoring, and readout are not shown.

the CTP and LTI hardware is newly developed and built for the Phase-II upgrade, the baseline for the MUCTPI is to use the Phase-I MUCTPI hardware [10.1] with upgraded firmware which takes into account the new requirements for Phase-II. Two MUCTPI modules may be used, compared to one in the Phase-I system, to handle increased input bandwidth requirements from the Phase-II L0Muon Sector Logic.

The CTP needs to be capable of supporting bigger trigger menus than in Phase-I, exploiting the ability of the Global Trigger to evaluate a large set of trigger criteria to address the physics challenges of the HL-LHC. This is reflected in increased input bandwidth and more logic to evaluate complex selection criteria.

The TTC system, incorporating the LTI modules, makes use of modern high-bandwidth optical-link technologies, replacing the original optical broadcast system that was developed

in the 1990s and incorporating the detector BUSY collection function that is currently performed electrically. The proposed **TTC** system for Phase-II builds on the experience that ATLAS has built up in Run 1 and Run 2, including both the central **TTC** system and related detector-specific electronics. The **LTI** together with **FELIX** will incorporate functionality that could not previously be supported in a common way for the detector systems, e.g. in connection with calibration procedures.

The main functions of the Central Trigger System are described in the following.

The **CTP** will have optical serial inputs to receive trigger information from the **Global Trigger**, the Phase-I legacy system and the **MUCTPI**. While the Phase-I **CTP** can form 512 trigger items from 512 single-bit inputs, the Phase-II **CTP** will allow 1024 trigger items based on 1024 usable single-bit inputs. The **CTP** combines trigger inputs from the **Global Trigger**, the **MUCTPI** and the Phase-I legacy **L1Calo** trigger systems, in order to form trigger items. The **CTP** combines trigger inputs from the **Global Trigger** – flags indicating which trigger criteria have been satisfied or multiplicities of trigger objects – with multiplicities and flags coming from the **MUCTPI** and Phase-I legacy **L1Calo** trigger systems, in order to form trigger items. The trigger items may be simple requirements, such as an inclusive selection of objects (e.g. muons) above a given  $p_T$  threshold, or complex requirements combining multiple objects. For each item, trigger algorithm flags from the **Global Trigger**, and conditions on the **MUCTPI** and the Phase-I legacy inputs may be combined in coincidence or anti-coincidence. The **CTP** forms the overall trigger decision based on the trigger items that are active, applying additional requirements, such as limiting triggers to a subset of bunch crossings in the **LHC** train, deadtime, and pre-scaling.

The Central Trigger System includes extensive monitoring of trigger rates and of deadtime. This is used in assessing the luminosity and beam conditions, as well as for operational monitoring and optimisation of trigger conditions.

The Central Trigger System also distributes the LHC timing signals and the trigger decision, together with associated data such as a Trigger Type word to **LTI** modules associated to the sub-detectors. The **LTI** will reformat the received **TTC** information, with the possibility of adding asynchronous and synchronous sub-detector data and commands via a programmable interface, and send it to the sub-detector **FELIX** systems or other sub-detector-specific electronics. BUSY signals, i.e. backpressure signals from the subsystem readout, are collected via the **LTIs** and sent back to the **CTP**, where they are combined and used to apply deadtime when necessary. The **LTI** modules provide facilities that can be used by the detector subsystems for various functions, including standalone calibration running and synchronous distribution of detector-specific signals and data. It is assumed that the **TTC-PON** (Passive Optical Network) technology [10.2] is used for the link between the **LTI** and **FELIX** I/O cards.

The Level-0 **MUCTPI** is the component of the Level-0 Central Trigger system which aggregates and merges the trigger information from the barrel muon and endcap muon systems before passing it on to the **Global Trigger** and to the **CTP**. It provides detailed information

---

on the muon candidates to the [Global Trigger](#) and sends muon multiplicity information and trigger flags directly to the [CTP](#). It is also responsible for resolving cases where a muon gives rise to multiple muon candidates, e.g. when it passed through detectors in more than one sector of the muon spectrometer.

Care is being taken to use processing [FPGAs](#) in the [MUCTPI](#) and [CTP](#) with some reserve in resources and speed, beyond the minimum requirements for the algorithms that are currently foreseen. It is justified by the central role played by these systems and cost effective due to the very small number of MUCTPI and CTPCORE units used. This reserve will give considerable flexibility to address new requirements that may emerge in the coming years. An example that we have already considered is a trigger that includes delayed signatures combined with prompt ones. This could be implemented by adding trigger items based on special or duplicated inputs to the [CTP](#) with adjusted time-alignment parameters.

## 10.1 Level-0 CTP

The [CTP](#) is the last stage of the processing chain of the Level-0 trigger system, making the final Level-0 accept decision (LOA). It receives digital trigger inputs from the [Global Trigger](#), the [MUCTPI](#), the legacy Phase-I trigger system, and various forward detectors and sub-detector calibration systems in the agreed form of simple flags and multi-bit multiplicities of trigger objects. While the majority of trigger inputs are received via the [Global Trigger](#), some muon-only inputs such as muon multiplicities and muon-only topological triggers may be produced by the [MUCTPI](#) as in Phase-I. During the commissioning phase, the legacy Phase-I trigger system can provide well understood trigger inputs that can be used while the new [Global Trigger](#) is being commissioned. This includes inputs from the Phase-I [L1Calo](#) system and the muon multiplicities from the [MUCTPI](#) (that may continue to be used once the [Global Trigger](#) system is commissioned).

The trigger inputs arriving at the [CTP](#) from different sources will not generally be aligned in time, i.e. data from the same bunch crossing but from different sources will arrive at the [CTP](#) at different times. The [CTP](#) therefore includes programmable elements to delay the inputs that arrive first, to align them in time with those that arrive last.

The algorithm used by the [CTP](#) to combine the different trigger inputs allows events to be selected on the basis of a programmable trigger menu. An event is selected if it satisfies the criteria of one or more menu items. Each menu item consists of the logical combination of a number of criteria, typically multiplicity requirements for muons, electrons/photons, hadrons/taus and jets, threshold requirements on missing or total transverse energy, and flags indicating events passing the criteria in a [Global Trigger](#) algorithm. Menu items may be enabled or disabled for certain bunches in the LHC, for example be enabled for colliding bunch pairs only, or be disabled during the gaps in the bunch train when calibration pulsers may be fired. The [CTP](#) makes provision for individually prescaling any trigger

menu item, and is responsible for introducing deadtime as required by the detector front-end and readout systems. Deadtime is normally generated internally using algorithms that are described in detail later, and taking into account the BUSY backpressure signals from the sub-detector readout systems.

After the application of deadtime, the **L0A** signal is formed as the logical OR of the enabled trigger items. Along with associated trigger information and timing signals, the **L0A** signal is distributed to the sub-detectors, via optical point-to-point links to sub-detector **LTI** modules, and subsequently via **TTC-PON** links to the **FELIX** systems.

The **CTP** will be implemented as a single **ATCA** shelf [10.3] with the following custom-built blades:

- **CTPMI**: the LHC Machine interface, to receive the beam timing signals from the LHC, most notably the 40 MHz bunch clock and the ORBIT signal. The backplane will be used to distribute the timing signals to all the blades.
- **CTPIN**: an input board to receive non-latency-critical optical and electrical trigger signals, i.e. optical trigger inputs from the legacy Phase-I trigger processors and from the **MUCTPI**, and electrical trigger inputs from various forward detectors and sub-detector calibration systems.
- **CTPCORE**: the board that implements the core functions of the **CTP**, the realtime trigger logic path, pre-scaling, deadtime, readout, monitoring, and the **TTC** interface to the sub-detectors. There will be one instance (**L0CTP**) to be used for the Level-0 trigger system. For the evolved Level-0/Level-1 system, a second **CTPCORE** board with different firmware would be used as the **L1CTP**.

Figure 10.2 shows the basic architecture of the **CTP**, with its main components, its inputs and outputs, and the main flow of information.

### 10.1.1 CTPMI: Machine Interface

In order for ATLAS to be synchronous with the LHC beam, the **CTP** will need to provide two signals to the ATLAS sub-detectors via the aforementioned **TTC** distribution system: the 40 MHz bunch clock and the ORBIT signal. The bunch clock is a 40 MHz clock derived from the LHC RF system and hence synchronous with the arrival of bunches in ATLAS. The ORBIT signal is one pulse per LHC revolution and is synchronous with the LHC bunch train structure. As for Run 1 and Run 2, the LHC will provide the following signals: BC1, BC2, ORBIT1, ORBIT2; i.e. 40 MHz bunch clock and ORBIT signals separate for beam 1 and beam 2 of the LHC. At flat top of the LHC magnet cycle, the two beams will be frequency-locked onto each other, so that BC1 and BC2 become identical.

As no LHC-wide decision has been made for a common receiver for these four signals, the requirements need to be collected in order to understand how these signals are received.



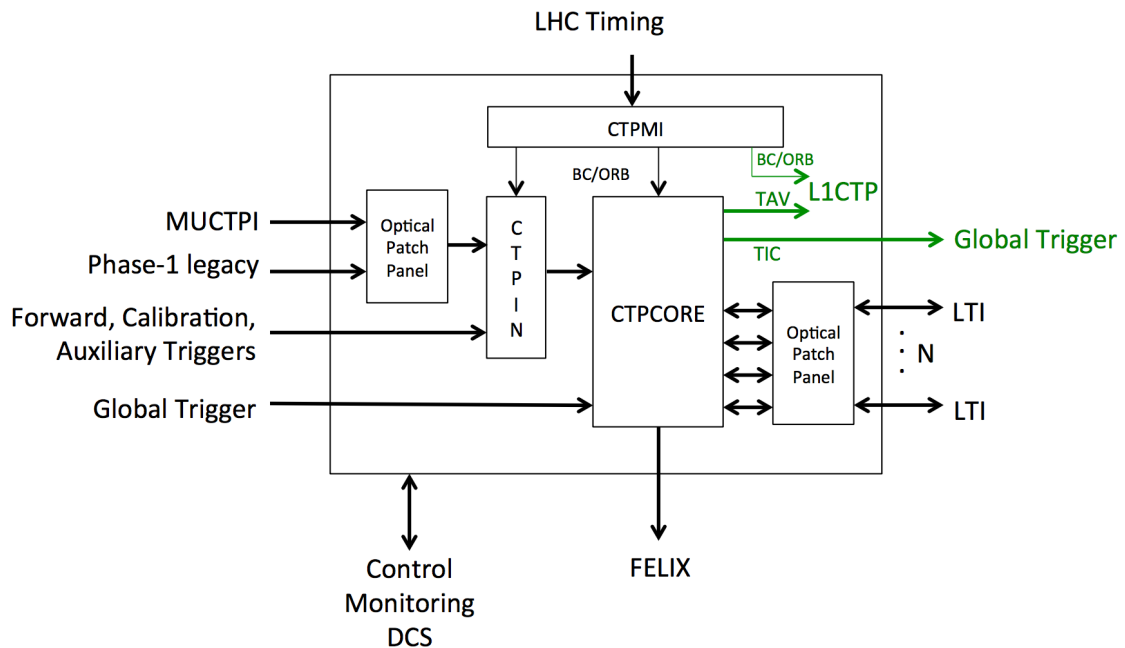


Figure 10.2: Basic architecture of the CTP, showing its main components, inputs, outputs, and the main information flow. Shown in green are additional elements only present in the evolved Level-0/Level-1 architecture.

Sufficient flexibility on the CTPMI will be provided to facilitate later adaptations. The CTPMI will be equipped with an FPGA (e.g. Xilinx Kintex Ultrascale+ KU15P [10.4]) to perform a variety of functions and to provide the required flexibility. Typical functions include synchronisation of the ORBIT signals, fine-tuning of the phase of the clock signals, switching between clocks, providing an internal 40 MHz clock, and jitter cleaning. Depending on whether or not they will be already implemented on a possible LHC-wide common receiver module, these functions could be provided by the CTPMI. The timing signals (bunch clock, ORBIT) are made available on the ATCA backplane for distribution to other modules (CTPCORE, CTPIN).

### 10.1.2 CTPIN: Trigger Inputs

The CTPIN board provides electrical and optical inputs for non-latency critical inputs, such as auxiliary electrical trigger signals from forward detectors and calibration systems, as well as optical trigger inputs from the legacy Phase-I trigger system and the MUCTPI. Up to 24 electrical NIM standard trigger signals can be received via LEMO connectors. Two 12-way ribbon fibre receiver modules [10.5] allow one to connect up to 24 serial optical links, each

with line rates of up to 9.6 Gb/s. The two receiver modules simplify the physical interface to the **MUCTPI** and the legacy Phase-I trigger system and give additional flexibility via sufficient headroom in the number of fibres and bandwidth. The inputs are connected to an **FPGA** (e.g. Xilinx Kintex Ultrascale+ KU15P [10.4]) where the signals are synchronised and aligned in time. The rates are monitored in detail, including per bunch, and up to 512 input bits can be selected to be routed to the output. A 12-way ribbon fibre transmitter module [10.5] allows one to send the 512 selected trigger signals to the CTPCORE module.

It is possible to use more than one CTPIN module, e.g. in case more than 24 electrical trigger signals need to be included.

### 10.1.3 CTPCORE

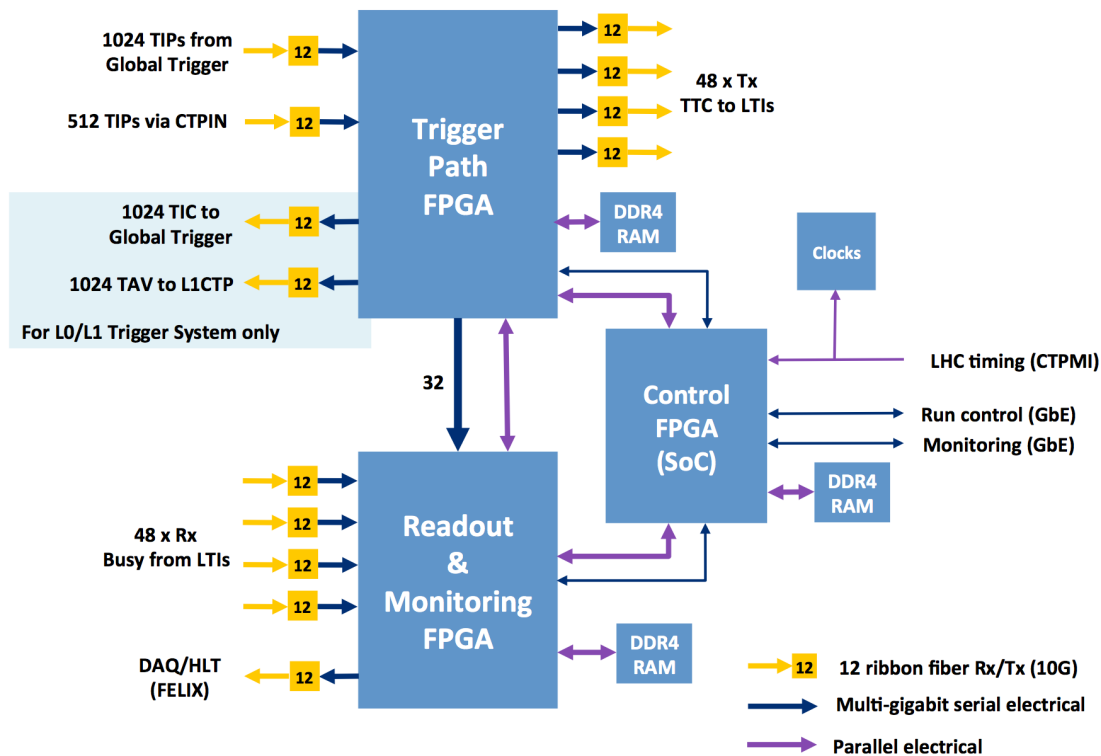


Figure 10.3: Block diagram of the CTPCORE architecture.

Figure 10.3 shows the proposed implementation of the CTPCORE and its interfaces to other components of the trigger system. It will be a single electronics board based on the **ATCA** standard.

The implementation will use **FPGAs** with a large number of on-chip multi-gigabit transceivers/receivers (**MGTs**). An existing candidate **FPGA** from the Xilinx Virtex Ultrascale+ family [10.4] is the VU13P, which has 128 on-chip **MGTs**, a sufficient number of logic cells and sufficient memory to implement the required logic. There will be one such **FPGA** for the real-time trigger path, and a second one for the event readout, monitoring, and dead-time generation. The **FPGAs** will be chosen such that a significant fraction of their resources remain free as margin for the implementation of future functionality. High speed-grade devices will be used to minimise the contribution of the CTP to the trigger latency. An additional **FPGA** with an on-chip processor (System-on-Chip SoC) will be used to control the CTPCORE blade.

**Trigger inputs** The trigger inputs are implemented as 24 optical serial links, grouped into two 12-way ribbon fibre receiver modules [10.5]. These links will be operated at 9.6Gb/s synchronously with the bunch crossing (BC) frequency with 8b/10b encoding, i.e. 4600 bits/BC, thus providing a lot of headroom in the required input bandwidth and connectivity.

The first 12 links are reserved for the **Global Trigger**, from which it is planned to receive 1024 trigger bits. The remaining 12 links are to be used for trigger inputs from one or more CTPIN modules, with up to 512 trigger bits from the **MUCTPI**, the legacy Phase-I trigger system, and from forward detectors and sub-detector calibration trigger signals.

**Trigger formation** The realtime path of the trigger formation in the CTPCORE is shown in Fig. 10.4. Out of the 1536 received trigger inputs, at any given time, the CTPCORE is able to use 1024 in the subsequent **LOA** trigger formation process. The 1536 incoming trigger inputs will be synchronised to the BC clock and passed through a delay pipeline of programmable length in order to align the data to the correct bunch crossing. Programmable and self-generated trigger signals such as random triggers and bunch pattern flags will be added at this stage. A switch matrix allows a selection of 1024 trigger inputs (TIP) to be mapped to the subsequent Trigger Logic block. The Trigger Logic block offers programmable logical functions to combine the 1024 inputs to form 1024 trigger items according to a defined trigger menu. The block will be organised as an array of Look-Up Tables (**LUTs**), in combination with one or more large ternary Content-Addressable Memories (**CAM**). The LUT/CAM scheme builds on the extensive experience gained during Run 1 [10.6] and Run 2, where this scheme was successfully used. It offers a lot of flexibility in the logical combinations of the trigger signals and allows one to change the trigger menu without modifying the firmware, by simply loading different configuration parameters and memory contents into the CTPCORE.

The selection of the **FPGA** for the CTPCORE board is driven by the number of on-chip high-speed serial links as well as by the required logic and memory resources. The required resources are dominated by the LUT/CAM structure used to implement the trigger menu, in particular the CAM resources scale with the square of the number of trigger items. Doubling the number of trigger items with respect to the Run 2 system would require an **FPGA**

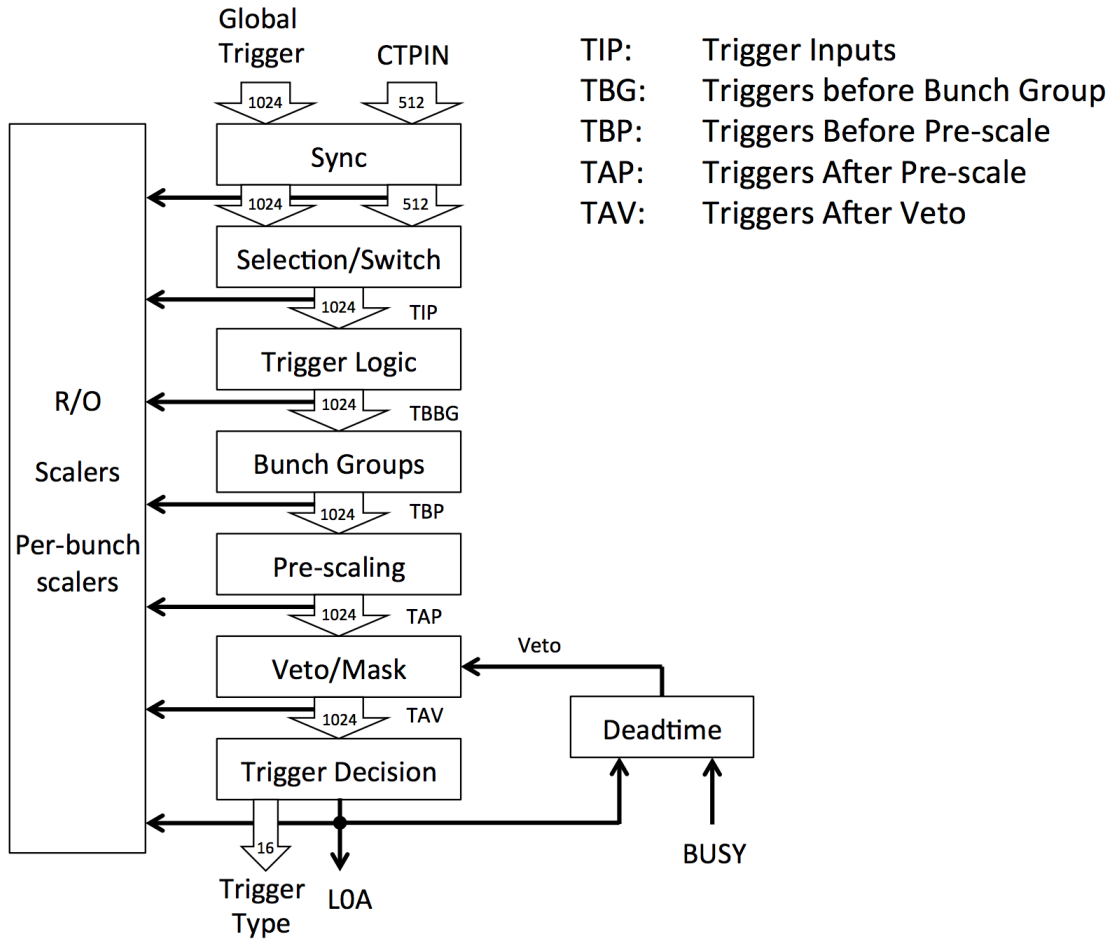


Figure 10.4: Realtime path of the trigger formation in the CTPCORE.

with at least four times the capacity of the Run 2 system to implement the Phase-II CTP-CORE trigger menu. The foreseen [FPGA](#) device (Xilinx Virtex Ultrascale+ VU13P [10.4]) has more than five times the number of logic look-up tables as the device used on the current Run 2 CTPCORE (Xilinx Virtex-7 7VX485T [10.4]) and should therefore have sufficient spare resources to also implement additional Phase-II functionality. Thus, we are confident that we can implement the aforementioned logic for 1024 inputs and 1024 trigger items in a single [FPGA](#) device.

The 1024 trigger items from the Trigger Logic Block are called “Trigger items before Bunch Groups” (TBG). They are each gated with an individually programmable combination of so-called bunch group masks. The bunch group masks implement fully programmable bunch patterns repeated at the LHC revolution frequency. The resulting 1024 so-called “Trigger items Before Pre-scale” (TBP) are pre-scaled with individually programmable pseudo-

random scalers to form 1024 so-called “Trigger items After Pre-scale” (**TAP**). The 1024 TAP signals are gated with the VETO signal, whose formation is described below. A so-called “trigger mask” is also applied to enable and disable individual trigger items. The resulting 1024 so-called “Trigger items After Veto” (**TAV**) are then combined in logic OR to form the **LOA** signal. Along with the **LOA** signal, a 16-bit Trigger Type word is calculated, based on trigger type masks programmed for each trigger item. The resulting Trigger Type word is the logic OR of the trigger type masks of all trigger items active after veto and trigger mask.

The Deadtime block uses the time history of the **LOA** and the logic OR of all sub-detector BUSY signals to generate the VETO signal that is used to gate the TAP signals. The VETO signal is formed as the OR of the combined sub-detector BUSY signals and the result of an array of complex deadtime algorithms which offer protection against trigger bursts for sub-detector front-end buffers. The complex deadtime algorithms are typically leaky-bucket or sliding window algorithms, with programmable parameters and fed by the **LOA** signal of previous bunch crossings. The **FPGA** allows sufficient flexibility to address specific sub-detector requirements and needs based on experience gained during commissioning and early running.

**TTC output** The **LOA** signal, along with the bunch clock and additional information is sent to the sub-detector Local Trigger Interface (**LTI**) modules via high-speed serial optical point-to-point links. Four 12-way ribbon fibre transmitter modules [10.5] are foreseen to interface, through an optical patch-panel, with up to 48 sub-detector **LTI** modules. These links will be operated at 9.6 Gb/s synchronously with the BC frequency, with 8b/10b encoding, which allows the transmission of 192 bits per bunch crossing. Since it is implemented in firmware, the exact data format can be adapted to the final requirements when they are understood in detail. However, the signals and information will include:

- BC
- Bunch Counter Reset
- BCID
- Turn counter value
- **LOA**
- Trigger Type
- **LOA** counter value
- Control commands, resets, etc.

**BUSY collection** The BUSY signals from the up to 48 sub-detector **LTI** modules are received using 48 optical serial links, grouped via an optical patch-panel into four 12-way ribbon fibre receiver modules. While only the BUSY signal needs to be transmitted by the **LTIs**, there is bandwidth available to send additional information, if needed. The BUSY signals will be received in the Readout and Monitoring **FPGA** of the CTPCORE. They can be individually masked and they are combined in logical OR to form the combined sub-

detector BUSY signal, which is subsequently used in the Deadtime generation block. The BUSY signals can be constantly monitored in order to identify the sources of deadtime.

**Event readout** For each L0A event, the CTP will send an event fragment to its FELIX system. A 12-way ribbon fibre transmitter module [10.5] is foreseen for this purpose, to be used as four or more 9.6 Gb/s FELIX GBT-FPGA links operating in “full mode”. Along with standard header and trailer information, bit arrays of the trigger inputs and items at the various stages of trigger processing will be sent as main payload: the trigger input bit pattern TIP, the trigger item bit pattern before prescaling (TBP), after prescaling (TAP), and after veto (TAV), as well as the bit pattern of all trigger inputs contributing to the formation of the L0A (TIC). It will be possible to send parts of this information in a programmable window around the bunch crossing which contains the L0A. For illustration, without any data-compression scheme, one can estimate about 5 kbits of data per bunch crossing being read out, and the additional information, including header and trailers will likely not surpass 1 kbit. Hence, four links at 9.6 Gb/s line rate and 7.5 Gb/s available bandwidth in full mode, will allow, at 1 MHz L0A rate, to read out with a readout window of up to 5 BC.

**Monitoring** The CTP features monitoring counters for all trigger input signals, trigger items before and after pre-scaling and veto (TIP, TBG, TBP, TAP, TAV), L0A, incoming sub-detector BUSY signals, and generated deadtime. They allow continuous counting with intermediate readout accompanied by a 25 ns precision time interval for accurate rate normalisation. Special counter arrays (“per-bunch histograms”) will allow trigger input rates, item rates, and deadtime to be monitored for each individual bunch-crossing identifier.

**Hooks for the evolved Level-0/Level-1 architecture** Two 12-way ribbon fibre transmitter modules [10.5] are foreseen to send Level-0 trigger information to the Level-1 trigger system in the evolved Level-0/Level-1 architecture. See Section 14.4.5 for more details.

**Partitioning** It will be possible to partition the CTP outputs and assign each individual output to one of several partitions. This will allow one to run several data-taking partitions concurrently, which is useful during testing and calibration periods. In this scheme, the trigger path is shared among the three partitions up to and including the pre-scaling. There will be several instances of the Veto/Mask, Trigger Decision, and Deadtime blocks, one per partition. CTP event readout fragments can be sent to FELIX for each partition.

## 10.2 Distribution of Timing, Trigger, and Control Signals (TTC)

Timing, Trigger, and Control (TTC) signals need to be distributed to the sub-detector FELIX systems and other sub-detector-specific electronics. The timing signals include the 40 MHz beam-synchronous clock derived from the LHC, which is used to clock the sub-detector and trigger electronics, as well as the ORBIT signal that is synchronous with the LHC turns. The clock needs to be distributed with high quality and low jitter. Trigger signals include the L0A signal, which has to be sent with a fixed latency with respect to the corresponding

collision, along with associated information, such as a [LOA](#) counter, a bunch-crossing identifier, information on the type of trigger decision, etc. In addition, there will be control signals and data, sent synchronously or asynchronously with the beam, for sub-detector-specific uses. The final contents and format of the [TTC](#) information will be defined taking into account detailed requirements from [TDAQ](#) and the detector systems. More details can be found in [10.7].

The distribution of the [TTC](#) signals is organised as a tree-like network, driven by a [TTC](#) master. During physics data-taking, the [CTP](#) assumes the role of the [TTC](#) master, while during sub-detector tests and calibration runs, an [LTI](#) module associated with the sub-detector can take over this role. In the opposite (“upstream”) direction, [BUSY](#) signals are sent from the sub-detector [FELIX](#) systems or sub-detector-specific back-end electronics. They are collected by the [TTC](#) master and used to throttle the generation of [LOA](#) signals.

The [CTP](#) will have 9.6 Gb/s point-to-point optical serial links to up to 48 [LTI](#) modules, which define the sub-detector partitioning and are associated with the sub-detector back-end or [FELIX](#) systems. For calibration and test runs, the [LTI](#) modules can replace the [CTP](#) and act as [TTC](#) master. The [LTI](#) distributes the [TTC](#) signals to a large number of [FELIX](#) I/O cards. The baseline implementation uses the 9.6 Gb/s [TTC-PON](#) optical link [10.2], which can be split by a ratio of 1:32 to be broadcasted to up to 32 destinations. A solution in which the [LTI](#) has up to 8 [TTC-PON](#) links is described here; hence each [LTI](#) can drive up to  $8 \times 32 = 256$  [FELIX](#) I/O cards.

The aforementioned [TTC-PON](#) 9.6 Gb/s link has a downlink bandwidth of 200 bits per BC and uses time-division multiplexing for the uplink. Each [FELIX](#) I/O card can send up to 56 bits every  $4 \mu\text{s}$  (in the maximum case of 32 [FELIX](#) I/O modules per [TTC-PON](#) link). This solution provides a lot of additional flexibility. For instance, instead of transmitting only a single [BUSY](#) flag from one [FELIX](#) I/O card, the available upstream bandwidth would allow one to send additional information indicating the source of the [BUSY](#).

## 10.2.1 Local Trigger Interface

The Local Trigger Interface ([LTI](#)) is an integral part of the [TTC](#) distribution in [USA15](#). It provides a point-to-point interface to the [CTP](#) and a [TTC-PON](#) interface to the sub-detector [FELIX](#) systems. For subdetector standalone running, it provides the basic [CTP](#) trigger functions.

The [LTI](#) will be implemented as a single [ATCA](#) blade, to be housed either in existing sub-detector [ATCA](#) shelves, or in a dedicated [ATCA](#) mini-chassis installed close to the sub-detectors’ [FELIX](#) systems. Figure 10.5 shows the basic architecture of the [LTI](#) module. It features an [FPGA](#) (e.g. Xilinx Kintex Ultrascale+ KU15P [10.4]) which has connections to all the optical and electrical inputs and outputs and thus gives a lot of flexibility to implement the required functions. The blade is controlled by a System-on-Chip [FPGA](#) [10.4] connected to a gigabit ethernet network. Two 10 Gb/s [SFP+](#) transceivers [10.8] allow the



## 10.2 Distribution of Timing, Trigger, and Control Signals (TTC)

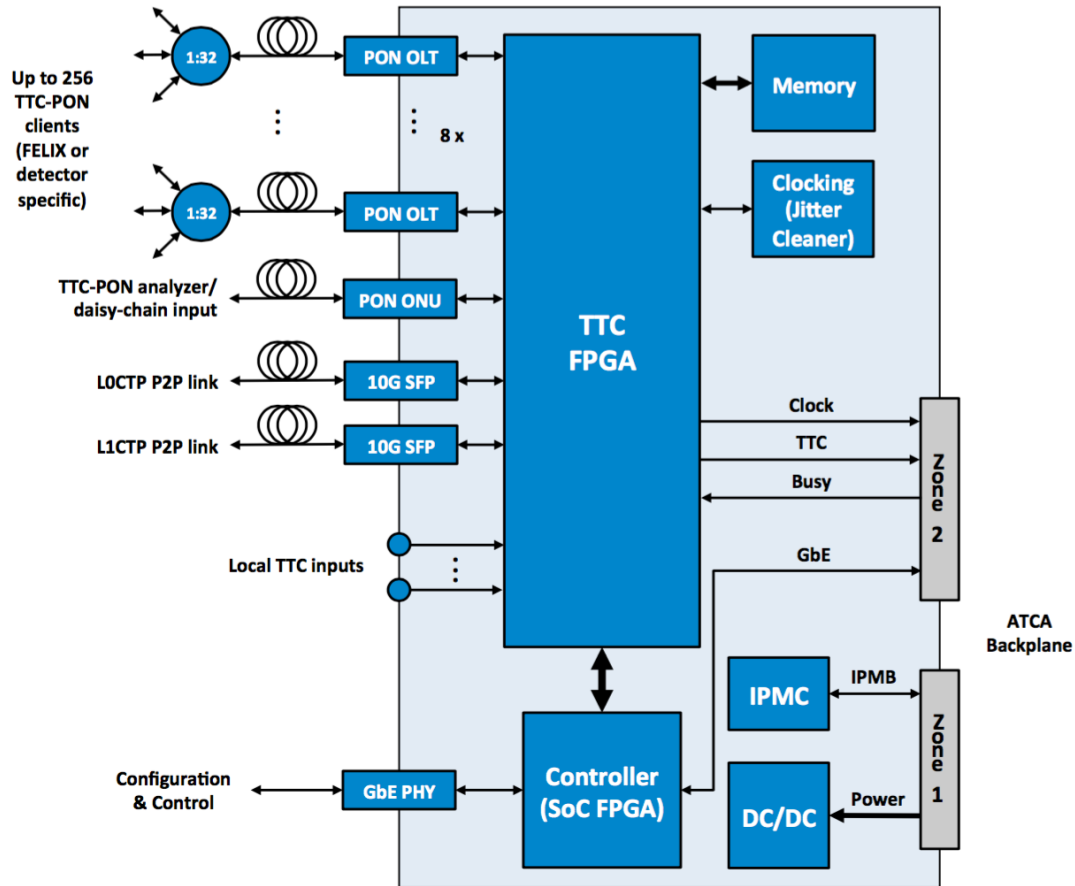


Figure 10.5: Block diagram of the LTI architecture.

optical connection to the L0CTP and possibly the L1CTP (see below), through which the TTC signals are received and the BUSY signal is sent. Up to 8 TTC-PON OLT (Optical Line Terminal) pluggable modules allow one to send TTC information to up to 256 TTC-PON destinations and receive corresponding BUSY information. The module features an additional TTC-PON ONU (Optical Network Unit) pluggable module to be used as a TTC analyser for debugging and monitoring purposes. Additional electrical inputs are foreseen, in order to connect trigger, control, and BUSY signals from external sub-detector electronics. A few electrical outputs will allow one to monitor selected signals with an oscilloscope.

The FPGA is used to implement in firmware the main functions of the LTI: the translation of the incoming TTC signals between the CTP and the TTC-PON format, a programmable interface for sending sub-detector data and commands, and the standalone CTP-like functions. In normal physics running, the LTI receives the main TTC signals optically from the CTP via the 9.6 Gb/s link.



---

It will be possible to add sub-detector specific information to the **TTC** stream and merge it with the **TTC** information from the **CTP**. Typical use-cases include detector calibration information and on-the-fly periodic re-configuration to mitigate SEU effects. A separate interface is provided for each **TTC-PON** output with programmable mechanisms for sending the data in asynchronous or synchronous ways. The information can be inserted in the **TTC** stream in various ways, e.g. synchronously with the LHC turn, during bunch crossings without **L0A**, during longer LHC bunch train gaps, during the time when an externally connected signal is active, etc. The combined **TTC** information is then sent via the **TTC-PON** outputs to up to 32 destinations per output. Each **TTC-PON** also receives the **BUSY** information from all of its up to 32 destinations, via the **TTC-PON** up-link that follows a round-robin time-division multiplexing mechanism. The **BUSY** signals of all **TTC-PONs** are combined in logical OR and sent to the **CTP**. In addition to the **BUSY** state of the particular **FELIX** I/O card, the up-link could carry additional information, such as information on the **BUSY** source to be used for detailed **BUSY** monitoring, or information in the context of detector-specific calibration or configuration procedures.

For sub-detector standalone runs, the **LTI** plays the role of the **TTC** master. For this purpose, it includes **CTP**-like functions: essential trigger functionalities to be able to perform standalone sub-detector data-taking, using external electrical or internally generated signals, with realistic patterns e.g. for high-rate tests. Incoming electrical trigger signals can be synchronised and aligned in time. Additional trigger signals can be generated via a pattern generator and a pseudo-random generator. Trigger logic is provided via a programmable look-up table. The resulting trigger items can be combined with bunch group patterns and can be pre-scaled if necessary. They are gated with a **VETO** signal before the **L0A** signal is formed as the logical OR of all selected trigger items. The **VETO** signal is generated as the logical OR of the **BUSY** signals received via the **TTC-PON** links, an electrical **BUSY** input, and realistically generated deadtime using the same deadtime algorithms as in the **CTP**.

Monitoring counters, including some per-bunch monitoring facilities, are provided for the trigger signals at the input, output, and intermediate stages, as well as for **BUSY** and dead-time signals.

A pattern memory is also provided, which can be used to play back user-defined trigger sequences or to capture incoming signals for debugging purposes.

The **FPGA** will be chosen such that it allows enough flexibility and available resources to be able to embrace new requirements based on future user experience.

### 10.3 Level-0 MUCTPI

The **MUCTPI** processes and combines muon trigger information from the barrel and end-cap trigger system and passes it on to the **Global Trigger** and the **CTP** for combination with other signatures. It resolves cases where a single physical muon gives rise to more than one

muon candidate. It counts such multiple objects only once for the multiplicity calculation (e.g. by only retaining the candidate with the highest transverse momentum,  $p_T$ ) and is able to filter duplicate candidates passed on for further processing.

After the synchronisation, time alignment, and overlap removal stage, the muon information is processed in two parallel paths: the **Global Trigger** and the **CTP** path. In the **Global Trigger** path, the muon candidates are sent to the **Global Trigger** for further processing, with the possibility to limit the number of candidates if necessary. In the **CTP** path, muon multiplicities are calculated for the various  $p_T$  thresholds, taking into account all muon sectors; they are transmitted directly to the **CTP**.

The **MUCTPI** is currently being built for the ATLAS **TDAQ** Phase-I Upgrade [10.9]. It is designed to be forward compatible with the Phase-II upgrade, where it is planned that the Phase-I **MUCTPI** hardware will be re-used with new firmware, possibly using two hardware modules instead of one, as discussed below.

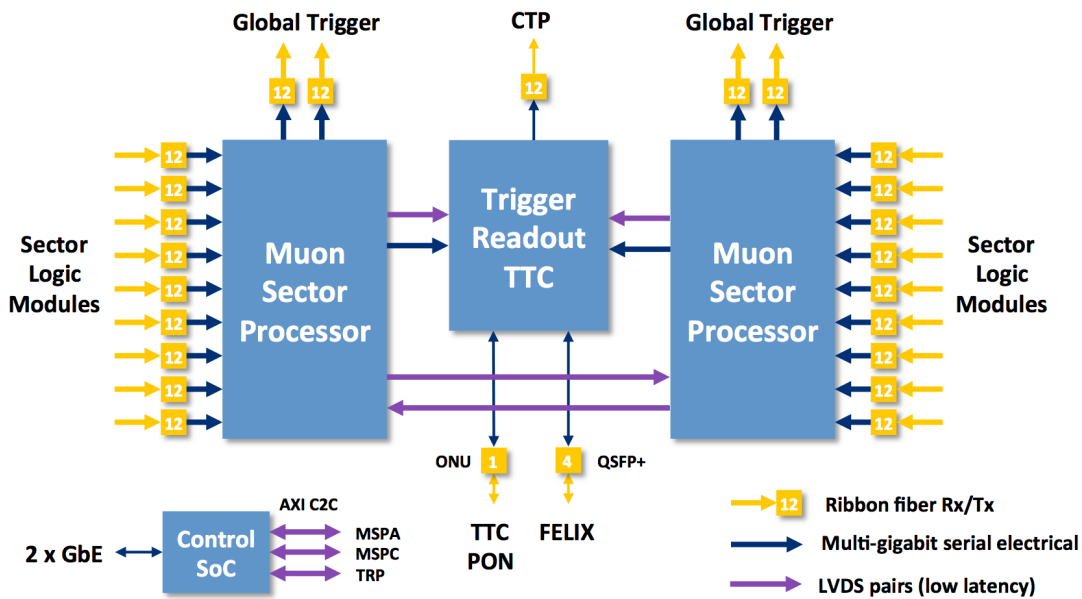


Figure 10.6: Block diagram of the **MUCTPI** board.

Figure 10.6 shows a block diagram of the architecture of the Phase-I **MUCTPI** board. The implementation of the **MUCTPI** is a single electronics board following the **ATCA** standard. The module is based on **FPGAs** with a large number of on-chip high-speed serial links, as well as high-density parallel fibre optics receiver and transmitter modules. The implementation uses a pair of Muon Sector Processor **FPGAs**, each handling half of the inputs from the **LOMuon** Sector Logic modules. They perform the synchronisation, time alignment, overlap removal and further processing tasks on the incoming muon information. A smaller **FPGA**

---

interfaces to the [CTP](#), [TTC](#), and [FELIX](#). The [FPGA](#) for the Muon Sector Processor is chosen such that a significant fraction of its resources remain free as a reserve for the implementation of the Phase-II functionality described here. Each of the two Muon Sector Processor [FPGAs](#) receives and processes the output from up to 104 links from [L0Muon](#) Sector Logic modules, via nine 12-channel parallel fibre optic receiver modules [10.5].

The results from each Muon Sector Processor [FPGA](#) are sent to the [Global Trigger](#) via up to two 12-channel parallel fibre optic transmitter modules [10.5]. The optical links operate at 9.6 Gb/s synchronously with the LHC clock and use 8B10B encoding. Assuming 64 bits per muon candidate, this allows each Muon Sector Processor [FPGA](#) to send up to 72 candidates per BC to the [Global Trigger](#), well in excess of the largest multiplicity relevant for the trigger. The large available bandwidth gives the possibility to send the candidates for a window of several BCs for use in evaluating delayed signals for exotic particles in the [Global Trigger](#). The exact information content and data format of the optical links will be defined based on the available information from the [L0Muon](#) Sector Logic and the requirements of the topological algorithms that will be implemented in the [Global Trigger](#).

A third [FPGA](#) is used to merge the results from the two Muon Sector Processor [FPGAs](#), and to send multiplicities for each  $p_T$  threshold to the [CTP](#) through a single optical link. The merger [FPGA](#) also implements the [MUCTPI](#) readout driver functionality, including the gathering and sending of the full muon trigger information to [FELIX](#) upon [L0A](#) reception.

The [MUCTPI](#) will interface directly to the [TTC](#) system via an [LTI](#) module. It will receive the optical [TTC](#) fibre via a [TTC-PON](#) ONU, and the firmware will be adapted to the new [TTC](#) distribution system foreseen for the Phase-II upgrade.

The Phase-I [MUCTPI](#) hardware aims to be forward-compatible for Phase-II. However, based on the final design of the Phase-II muon trigger system (the design of which has already evolved substantially since the Phase-I [MUCTPI](#) was specified) and based on experience from Run 3 operations, new requirements may emerge which exceed the input bandwidth or muon sector processing [FPGA](#) resources of the Phase-I [MUCTPI](#) hardware module and call for a hardware upgrade of the [MUCTPI](#). For this reason, the baseline is to operate the Phase-II [MUCTPI](#) with two Phase-I [MUCTPI](#) modules, one per detector hemisphere (i.e.  $\eta < 0$  or  $\eta > 0$ ), essentially doubling the input bandwidth and muon sector processing resources. Low-latency LVDS links between the two Muon Sector Processor [FPGAs](#) in each [MUCTPI](#) module are used to share information needed in the overlap-removal stage. There is no geometrical overlap between the muon detectors covered by the two [MUCTPI](#) modules.

## 10.4 Monitoring, Control, Configuration, Offline Software

Each [ATCA](#) board of the Level-0 Central Trigger system will feature extensive environmental monitoring. Parameters being monitored include voltages and currents of all crit-

ical power rails, as well as temperatures of [FPGAs](#), fibre optic transmitter/receiver modules and power converters. A small subset of parameters critical to the health of the board will be collected by the on-board [IPMC](#), which can take actions to protect the board. It also provides those parameters to the shelf manager of the [ATCA](#) which publishes them to [DCS](#). The [DCS](#) path through the shelf manager will also be used to power-cycle the board and to reset the System-on-Chip (SoC). In addition, the SoC has access to the full set of environmental parameters, which will be sent to further processing using Ethernet. If required, they can also be published to [DCS](#). The SoC will also be used for configuration, control, and monitoring of the module through a Gigabit Ethernet interface.

New software will be developed, which includes low-level software for interfacing with the [ATCA](#) blades, tools for diagnostics and tests of the hardware, online software for configuration, control, and monitoring within the ATLAS run control framework, as well as offline software for simulation, reconstruction, and event monitoring.

## 10.5 Latency

Table 10.1 summarises the estimated latency contributions of the critical path through the central trigger system. The latency of the [CTP](#) is for the latency critical path of inputs from

Item	Partial Current Best Estimate ( $\mu\text{s}$ )	Uncertainty/Contingency ( $\mu\text{s}$ )
<a href="#">CTP</a> Processing	0.275	0.100
Link to <a href="#">LTI</a> (5 BC + 30 m fibre)	0.275	0.025
<a href="#">LTI</a> Internal processing	0.100	0.025
Link to <a href="#">FELIX</a> (5 BC + 10 m fibre)	0.175	0.050

Table 10.1: Summary of the expected latency of the [CTP](#) and [LTI](#) system.

the [Global Trigger](#) to the [CTP](#). The de-serialisation at the serial link input is accounted for elsewhere. The quoted latency number of 275 ns for [CTP](#) processing includes the re-synchronisation to the local BC clock, the selection and switching of trigger signals, and the complete trigger path including trigger logic, bunch groups, pre-scaling, gating, and final OR to form the [LOA](#) signal and associated Trigger Type word.

For the link to the [LTI](#), 275 ns is estimated, including the serialisation in the [CTP](#), a 30 m long fibre, and the de-serialisation in the [LTI](#). The internal processing of the [LTI](#) is estimated to be 100 ns, and the link to [FELIX](#) is estimated to be 175 ns, including serialisation in the [LTI](#), a 10 m long fibre to a close-by [FELIX](#) system, and the de-serialisation step in [FELIX](#). The uncertainty is taken into account as contingency and summarised in Table 10.1.

Table 10.2 summarises the latency of the critical path through the [MUCTPI](#). For the Phase-II

Table 10.2: Summary of the expected latency of the *MUCTPI*.

Item	Partial Current Best Estimate ( $\mu\text{s}$ )	Uncertainty/Contingency ( $\mu\text{s}$ )
<i>MUCTPI</i> Processing	0.250	0.100
Link to MUX of <i>Global Trigger</i> (5 BC + 15 m fibre)	0.200	0.025

system, the firmware of the *MUCTPI* has to be completely re-written in order to process potentially twice as many candidates with a spatial resolution better by a factor of 16. The latency of *MUCTPI* processing is estimated to be 10 BC, i.e.  $0.250 \mu\text{s}$ , taking into account the sharing of trigger information between the two Muon Sector Processor *FPGAs* via the low-latency LVDS links. The latency of the link to the MUX of the *Global Trigger* is estimated to be 8 BC, i.e.  $0.200 \mu\text{s}$ , which is the sum of the serialisation in the *MUCTPI*, an assumed fibre length of 15 m, and the de-serialisation in the MUX of the *Global Trigger*.

## 10.6 Project Planning

A tentative schedule for the Phase-II Central Trigger upgrades is shown in chapter 19, for the *CTP*, the *MUCTPI*, and the *TTC* system.

It has been split into a number of activities, some of them common to the three systems. After capturing the requirements and designing the respective system in a conceptual way, the system needs to be specified in detail. For the *CTP* and *TTC*, this leads to the Preliminary Design Review. In parallel, hardware evaluations take place, e.g. concerning the optical links and *FPGAs*, using off-the-shelf evaluation tool-kits. For the *CTP* boards and the *LTI*, it is planned to start with a first prototype. After an evaluation and test phase, a Final Design Review is planned. Depending on the outcome of the review, a pre-production version may be necessary. For the *LTI*, the pre-production version could be produced including a small number of boards aimed to be distributed to the subdetectors for detailed assessment. After a Production Readiness Review, the final modules will be produced.

Concerning the *MUCTPI*, the requirements capture and conceptual design phase will last until one year after the start of Run 3 of the LHC, allowing one to benefit from experience with the Muon New Small Wheel. In a review of the requirements and specifications, a decision is expected for the final Phase-II *MUCTPI* architecture: whether the Phase-I single-blade *MUCTPI* hardware is adequate or whether a two-blade architecture using two Phase-I *MUCTPI* blades is needed. A redesign and construction of a new set of *MUCTPI* modules based on more modern link and *FPGA* technologies is considered as a risk mitigation option. For the baseline architecture with two Phase-I *MUCTPI* blades, one would have to launch the production of the additional Phase-I *MUCTPI* blades.

The evaluation of the prototypes will include interface tests with upstream and downstream systems, such as the [Global Trigger](#), [MUCTPI](#) and [LTI](#) for the [CTP](#), [FELIX](#) for the [LTI](#), and the [LOMuon Sector Logic](#) and the [Global Trigger](#) for the [MUCTPI](#).

For all three systems, firmware, low-level software, test software and high-level software needs to be developed. The firmware implementation starts with board infrastructure firmware and firmware to test and evaluate the hardware. It is followed by the firmware to control the board via the network and firmware for the high-speed links. Some of these firmware blocks are common for the various boards and may be shared. There is also firmware specific for the CTPIN, CTPMI, CTPCORE, [MUCTPI](#), and [LTI](#) boards, implementing the main functions of the respective board. The CTPCORE, for example, will have several firmware blocks that implement the various functions of the CTPCORE: the realtime trigger, the event readout, the monitoring, and the [TTC](#) interface.

Low-level software needs to be developed, including software for basic communication with the board and the board infrastructure. This part will be similar for the various boards and can be shared. It is followed by a low-level library to communicate with the specific registers of each board and a low-level test suite which includes diagnostic tools as well as software to perform full system tests. For the CTPCORE and the [MUCTPI](#) there will be specific configuration software which allows the boards to be configured via a higher-level description of the trigger menu for the CTPCORE and of the overlap handling for the [MUCTPI](#).

The software development concludes with the high-level software, which comprises control and monitoring software, which is integrated in the ATLAS TDAQ run control software framework. In addition, offline software for simulation, reconstruction, data quality, and analysis is required.

All hardware components will be first commissioned in the laboratory, based on trigger patterns generated at the hardware inputs and the comparison of readout and monitoring data with a full system simulation. Once fully commissioned in the lab, the systems will be ready for installation in [USA15](#) during Long Shutdown 3 and final commissioning with the ATLAS subsystems can begin.

## References

- [10.1] R. Spiwojs et al., *The ATLAS Muon-to-Central Trigger Processor Interface (MUCTPI) Upgrade*, Tech. Rep. ATL-DAQ-PROC-2017-013, CERN, Geneva, Jun, 2017.  
<https://cds.cern.ch/record/2270141>.
- [10.2] D. Kolotouros, S. Baron, C. Soos, and F. Vasey, *A TTC upgrade proposal using bidirectional 10G-PON FTTH technology*, *Journal of Instrumentation* **10** (2015) no. 04, C04001. <http://stacks.iop.org/1748-0221/10/i=04/a=C04001>.

- 
- [10.3] *AdvancedTCA Base Specification*, Tech. Rep. PICMG-3.0-R3.0, PCI Industrial Computer Manufacturers Group, 2008.
- [10.4] *UltraScale Architecture and Product Data Sheet: Overview*, Tech. Rep. DS890 v3.0, Xilinx, October, 2017.
- [10.5] *MiniPOD AFBR-814VxyZ, AFBR-824VxyZ, 14 Gbps/Channel Twelve Channel, Parallel Fiber Optics Modules*, Tech. Rep. AV02-4039EN, Avago Technologies, 2013.
- [10.6] ATLAS Collaboration, *Performance of the ATLAS Trigger System in 2010*, [The European Physical Journal C 72 \(2012\) no. 1, 1–61](#).  
<http://dx.doi.org/10.1140/epjc/s10052-011-1849-1>.
- [10.7] F. Lanni et al., *Interfaces with Detector Front-End Systems. Requirement Document for HL-LHC*, Tech. Rep. ATL-DA-ES-0051 v.1, EDMS Id 1563801 v.1, CERN, Geneva, March, 2016. <https://edms.cern.ch/document/1563801>.
- [10.8] *Specifications for Enhanced Small Form Factor Pluggable Module SFP+, Revision 4.1*, Tech. Rep. SFF-8431, SFF Committee, July, 2009.
- [10.9] ATLAS Collaboration, *Technical Design Report for the Phase-I Upgrade of the ATLAS TDAQ System*, Tech. Rep. CERN-LHCC-2013-018. ATLAS-TDR-023, CERN, Geneva, Sep, 2013. <https://cds.cern.ch/record/1602235>.





# 11 Data Acquisition

## 11.1 Introduction

The Data Acquisition (DAQ) acts to facilitate the final stage of the ATLAS online event selection process before permanent storage for offline reconstruction and analysis. The DAQ system transports and stores data accepted by the Level-0 trigger for further processing. The data are then provided on demand to the Event Filter (EF) system, which is then responsible for performing event selection. The DAQ system will change substantially for the Phase-II upgrade in response to evolving functional and performance requirements, driven in large part by increasing trigger rates and data volumes. The upgraded system will benefit from improvements in multiple related technologies. The work will continue the process, started during Phase-I, of replacing obsolete custom hardware and software with implementations built on commodity components.

This chapter describes the overall DAQ system architecture before detailing the design of each functional area, as well as the network which will facilitate all inter-component communication and data transport. Initial focus is given to the DAQ components themselves, moving then to the online software which underpins operations. Finally, a detailed discussion of the Event Filter component is presented, which draws on aspects of both the revised DAQ system design and new online software infrastructure. The EF system itself is described in more detail in Chapter 12.

The baseline architecture is designed to service a single-level hardware trigger running at 1 MHz, with hardware tracking as a co-processor to the EF. The system will support a maximum output rate of 10 kHz of full events for transfer to CERN permanent storage. This does not include rates for partial event building, which are not expected to drive bandwidth or processing requirements. All baseline figures will therefore only refer to the full event building rate.

Given the potential evolution to one featuring a split-level hardware trigger, a discussion of the implications of such a change for the DAQ system is presented at the end of the chapter. Such a change requires both modifications to the design and flexibility to be built into the revised baseline to facilitate future changes.

## 11.2 Overview of System Functionality

The primary functional blocks of the **DAQ** architecture in Phase-II are shown in Fig. 11.1. The detector Readout system in Phase-II will be made up of Front-End Link eXchange (**FELIX**) systems receiving event data from detector front-end (**FE**) links relayed across a commodity multi-gigabit network to software-based *Data Handler* applications running on commodity servers. The Data Handlers will host detector-specific processing such as fragment aggregation and formatting, while also facilitating monitoring operations. **FELIX** will also relay Timing, Trigger and Control (**TTC**) information from the Phase-II **TTC** system to on-detector electronics. Additionally, **FELIX** will provide a link between the Detector Control System (**DCS**), Online SW System and on-detector electronics. Through this channel commands and data will be relayed to detector **FE** components, with data also being distributed from the detector **FE** to **DCS** to facilitate routine operations such as control, configuration, monitoring and calibration. **DCS** will also receive monitoring data from **TDAQ** hardware throughout the system.

The Data Handlers will relay all event data to the Dataflow system for storage in a large buffer system known as the *Storage Handler*. Event building functionality will be implemented as part of the interface between the Data Handler and the Storage Handler. The Storage Handler will then serve data as needed to the **EF**, which will be implemented on a commodity server farm. Accepted events will be sent to the *Event Aggregator* system for packaging and transfer to the CERN computer centre for further processing. Underpinning all **DAQ** system activity will be a common online software framework; responsible for control, configuration and monitoring of the entire ATLAS data taking process. For development and commissioning, all **DAQ** system components will be designed to support standalone data taking either as a complete or partial detector slice as needed by detector operations.

## 11.3 Event Data Format

Detector data will be formatted by the *Data Handlers* after detector specific processing as need by the **Dataflow** and **Event Filter** systems. Logical organisation of the event data will be overseen by the *Event Builder* to optimise throughput of the system and minimise bandwidth transferred to the *Event Filter*. The accepted events will be formatted as required by offline reconstruction. The use of data compression is not currently planned, but this could be reviewed in future if compression is shown to be advantageous.

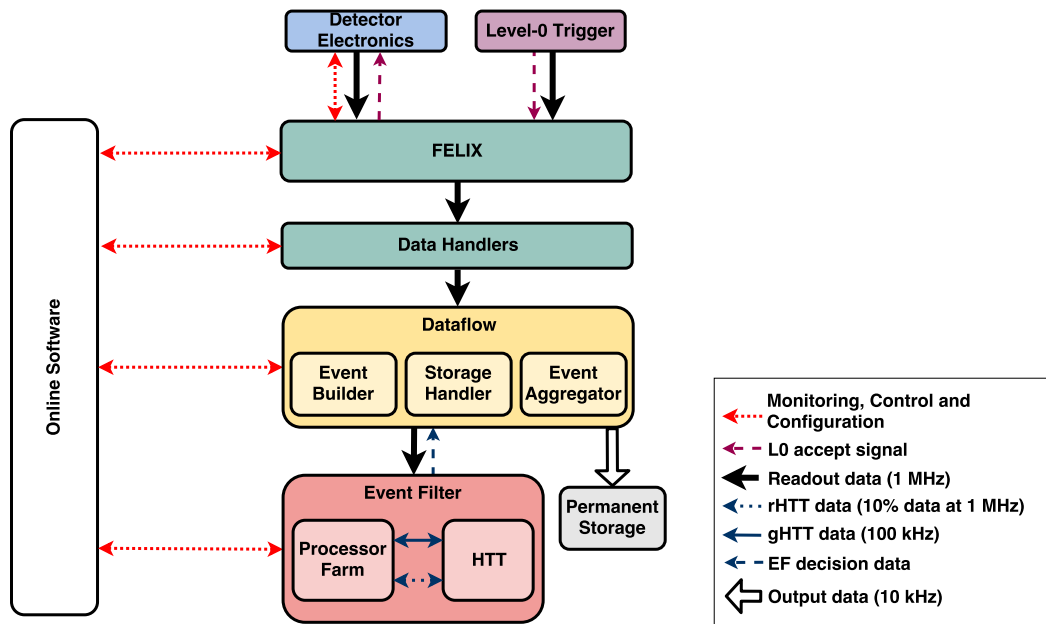


Figure 11.1: The main functional blocks of the *DAQ* architecture for Phase-II, with focus on interfaces with other systems. Dataflow is shown in yellow, Readout in Green. The *EF* components are shown for reference in Red. The arrows linked to Online Software indicate the path for monitoring, control and configuration operations, denoting the integration of all components with the software infrastructure.

## 11.4 Detector Readout

### 11.4.1 FELIX

As illustrated in Fig. 11.2, *FELIX* is the Readout system component which implements the interface to all detector-specific electronics via custom point-to-point serial-links. *FELIX* also acts as the interface to the Data Handlers, monitoring, control & configuration and *DCS* via a commodity multi-gigabit network. *TTC* information will be received by *FELIX* from the Local Trigger Interface (*LTI*) and relayed to detector *FE* electronics systems in a format satisfying system-specific requirements. The main idea behind the *FELIX* concept is the development of a modular system which makes it possible to independently upgrade or modify aspects of the system such as computing and buffering resources, network technology or supported point-to-point (or *PON*) serial-link protocols. The ability to evolve through further upgrades is a key feature of the Readout System when one considers the performance requirements and long development cycle leading to Phase-II, as well as the long lifetime of the ATLAS experiment beyond this period.

The ATLAS choice of optical link technology connecting to on-detector electronics is the Versatile Link [11.1] for Phase-I upgrades and the Versatile Link PLUS [11.2] for Phase-II

## 11.4 Detector Readout

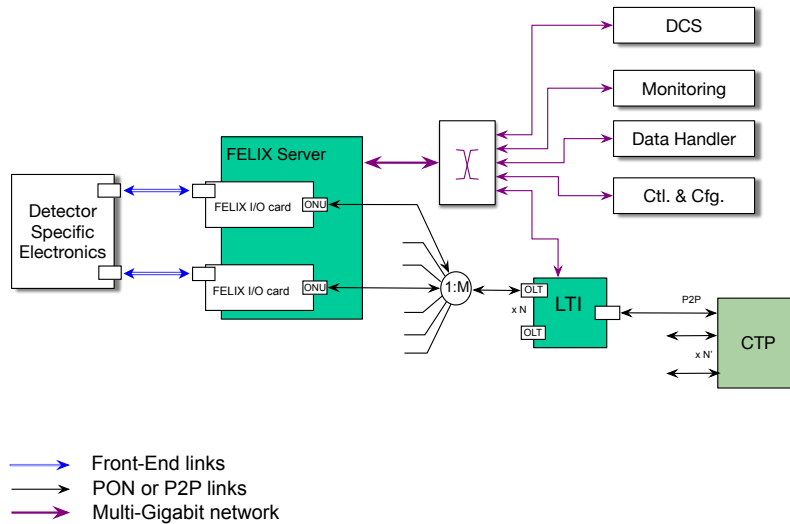


Figure 11.2: Possible Readout system architecture for Phase-II. The location of detector specific processing within the chain will evolve as requirements become better understood and tested.

upgrades. The latter is specified to operate at up to 10 Gb/s for uplinks and at 2.5 Gb/s for downlinks, and therefore is **lpGBT** compatible. **FELIX** will interface to these links with standard transceivers (e.g. miniPODs). Connections within the ATLAS computing cavern between off-detector electronics and **FELIX** will be implemented with standard optical links. **FELIX** will support numerous uplink protocols. These include **lpGBT**, FULL mode [11.3] and GBTx mode [11.4] (as already used in Phase-I) and any other detector specific protocols which will develop over the course of the upgrade, such as the protocol used for readout of the pixel detector. Downlinks will always make use of the **GBT** or **lpGBT** protocol to convey information from the **TDAQ** and **DCS** systems to **FE** electronics. Examples of these data include: **TTC** information (e.g. the bunch crossing clock), configuration and control signals for calibration procedures and **DCS** traffic. **FELIX** must therefore also receive packets from a commodity multi-gigabit network and route them to the relevant serial-link. The connections to **DCS** places a requirement on **FELIX** reliability and uptime beyond that of regular data taking operations.

**FELIX** functions as a router between custom serial links and a commodity multi-gigabit network. It is detector agnostic and encapsulates common functionality. There is no requirement for **FELIX** to decode or process the received data beyond what is required to determine its destination. In the baseline architecture, the detector **FE** electronics do not require **RoI** information. The **RoI** list (along with **L0CTP** output information) will thus be sent to **FELIX** as data fragments from the trigger systems and follow the normal Readout

and Dataflow path. This information will be used by **EF** processors, which will decode the list and retrieve the corresponding data.

**FELIX** will be implemented with a commodity server hosting custom FPGA I/O cards. Each **FELIX** card receives trigger and synchronous commands from one optical fibre (**PON**), which connects the **FELIX** card to one **LTI** module. The same fibre is used by the **FELIX** card to propagate back **BUSY** information. The **FELIX** card will receive complete information for each **LOA**, will pass on this information to the Data Handler(s), which are in charge of inserting it into the data fragments, and will send a subset of the information to the detector front-ends formatted according to the needs of the front-end electronics. **FELIX** will also allow front-end data sources to send data not only to the Data Handler but also to other network end points. For example, hardware trigger processors may sample 'bunch crossings of interest' that are not **LOA** or accumulate data for histograms of input and output data characteristics. These can be transferred to **FELIX** on a dedicated logical link different from **LOA** data and routed to a dedicated monitoring process on the network.

The **LTI** module will also be able to send **FELIX** additional user information, or specific trigger patterns (e.g. for calibration), as agreed with the detectors. The communications between the **LTI** and **FELIX** I/O card will be implemented in firmware at both ends. The essential **TTC** and **BUSY** functionality that exists in the Phase-I system will be complemented by enhanced features, including additional event identification counters and facilities useful for detector calibration, configuration and control. The detailed requirements will be captured in discussion with the detector communities, and the implementation will be optimised and refined in detailed design and prototype-evaluation work.

To meet the requirement of forward compatibility with Phase-II, the Phase-I upgrades of the New Small Wheel, **MUCTPI**, Liquid Argon calorimeter trigger electronics and Level-1 calorimeter and muon trigger systems [11.5] will use **FELIX**-based readout. These upgrades only require a subset of the functionality described, but **FELIX** will be designed to be architecturally compatible across upgrade phases. For Phase-II the design of **FELIX** will be expanded to address Phase-II specific requirements, such as: support for the upgraded **TTC** system, more complex configuration and control scenarios, vertical and/or horizontal aggregation of data across links before routing data fragments to the Data Handlers, expanded data routing capabilities and handling of a higher number of potentially higher speed links. Any Phase-I **FELIX** systems will be replaced with the upgraded Phase-II versions as needed.

As illustrated in Fig. 11.2, the baseline implementation envisages that the **FE** links and **GBT/lpGBT/FULL** mode protocols will be managed on a custom **PCIe** interface board (the **FELIX** I/O card) hosting a custom mezzanine that will implement the **TTC** interface (i.e. the connection to the **LTI**). It is expected that each **FELIX** interface board will be able to support up to forty-eight links, with two such boards hosted in a commodity server. However, the number of links that will be handled by each **PCIe** card largely depends on the required

## 11.4 Detector Readout

---

Table 11.1: Overall Phase-II Readout System Size. The precise number of *FELIX* systems will vary depending on the partitioning model which detector groups choose for their Readout slices. The number of links will increase due to the fact that fibres are usually handled in bundles of 12. The ratio between *FELIX* servers and Data Handler servers will be optimised for I/O and CPU load balancing.

Component	Total Number
Total links from detectors	17093
<i>FELIX</i> I/O card	545
<i>FELIX</i> servers	279
Data Handler PCs	545

link speeds, link utilisation and type of protocol as well as on the evolution of *PCIe* technology or any other equivalent high-speed bus that may be selected. The *PCIe* Gen4 (16 GT/s<sup>1</sup>) standard has already been released, offering double the bandwidth of *PCIe* Gen3 (in common use at the time of writing). The first devices compatible with *PCIe* Gen4 started appearing in the market in 2017. Based on this model of implementation, and assuming the choice of a bus at least as performant as *PCIe* Gen4 x16, the size of the Phase-II *FELIX* system is summarised in Table 11.1. A full breakdown of the expected number of links per subdetector is presented in Table 11.2. In places where the use of 10 Gb/s *lpGBT* with high-link utilisation (> 50%) is expected, 24 links from the front-end have been assumed for each board. In the case of lower link speeds (< 5 Gb/s or low utilisation), and less *FPGA* resource intensive link protocols, such as for the *ITk* Pixel detector, 48 links per board are envisaged.

The Phase-I *FELIX* has been demonstrated to be able to drive 24 *GBT* links and 12 FULL mode links. In the case of *GBT* the limiting factor is the number of available *FPGA* lookup tables, while in FULL mode the limitation comes from the maximum bandwidth of the *PCIe* Gen3 bus in use. More information can be found in the presentation prepared for the *FELIX* Phase-I progress review [11.6]. For Phase-II, *FELIX* systems will need to support 24 links, with the exception for the *ITk* Pixel *FELIX* systems which will need to support 48. This goal is achievable, since the *ITk* Pixel uplinks will resemble FULL mode links, which require far fewer *FPGA* resources to implement than *GBT* links. By 2019 *PCIe* Gen4 development systems are expected to be available. These will be used to start prototyping and identifying any limitations to be overcome.

### 11.4.2 Data Handler

The Data Handlers will receive data from *FELIX* via a commodity multi-gigabit network. The system will be implemented as a series of applications running on commodity serv-

---

<sup>1</sup> GigaTransfers per second, a common metric for bus performance.

Table 11.2: Summary of Phase-II Detector Readout Link and Bandwidth Requirements. Downlink refers to data travelling toward the front-end electronics, and uplink to data travelling from the front-end toward the rest of the DAQ system. Detectors with existing FELIX installations from Phase-I will be updated with new hardware as required.

Detector	Number of FELIX boards	Number of Links	Bandwidth (Gb/s)
ITk Pixel downlink	224	1285	2.5
ITk Pixel uplink		10596	5
ITk Strips downlink	76	1552	2.5
ITk Strips uplink		1824	10
LAr LASP downlink	36	100	2.5
LAr LASP uplink		770	10
LAr LDPB downlink	8	31	2.5
LAr LDPB uplink		155	10
L0Calo downlink	8	16	2.5
L0Calo uplink		120	10
NSW downlink	96	864	5
NSW uplink		1440	5
NSW Trigger Processor	4	64	5 or 10
Global Trigger downlinks	4	48	2.5
Global Trigger uplinks		96	10
Tile	8	160	10
TGC	8	192	10
MDT	64	1536	10
RPC	4	64	10
CTP	1	12	10
MUCTPI	1	4	10
LUCID uplink	1	24	5
ZDC uplink	1	24	5
AFP uplink	1	12	5

ers which facilitate detector-specific processing, e.g. formatting and/or monitoring, within a common DAQ infrastructure. After completion of detector-specific actions all data will then be passed to the Dataflow system for further processing by the Event Filter (see Section 11.5). To meet the requirements of detector-specific trigger-aware monitoring, automated recoveries and book keeping, the Data Handler will also receive Level-0 trigger information via FELIX.

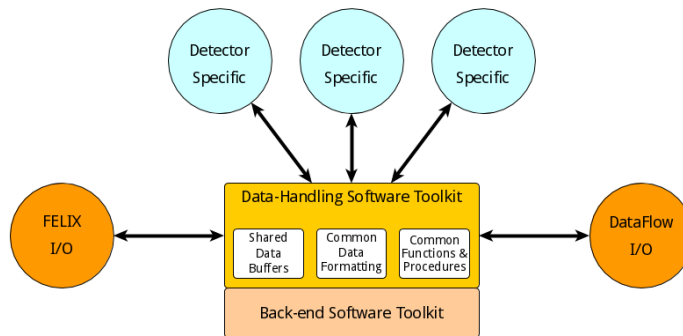


Figure 11.3: *Main components of the Data Handler infrastructure.*

The detector-specific data processing in the Data Handler, shown in Figure 11.3, as well as the configuration/calibration, control, and monitoring functions shown in Figure 11.2, will be implemented using common software tools. Raw data processing, which up to this point has been implemented in firmware on-board the Readout Driver (ROD) hardware, will now be implemented within customisable Data Handler software applications. On average, it is expected that a data handler will be capable of sustaining a bi-directional I/O capacity of order 100 Gb/s. The ratio between the number of FELIX boxes and Data Handlers for each detector is not yet well-defined. The number of Data Handlers, and the number of links each handle, can be easily adjusted given the flexibility of the network slice design. The final ratio will depend on I/O parameters such as bandwidth, i.e. how heavily subscribed each FELIX input link is and the data size for LOA fragments, and transaction rate, i.e. how many readout links per FELIX box will receive data for each LOA and whether all front-end sources report for every accept. The ratio will also depend on the total CPU load (within the Data Handlers) requested for DAQ data formatting and for detector specific needs, such as calibration and monitoring operations. For modelling this the assumption can be made that a few thousand operations per incoming fragment are possible.

### 11.4.3 R&D leading to the final design and implementation of the Readout System

The rapid technology evolution in the area of high-speed interconnects makes it premature to freeze the design of the Readout System at this point. There is nevertheless a baseline



architecture, in-line with the implementation of **FELIX** for Phase-I (hardware shown in Figure 11.4), relying on **PCIe**, multicore commodity servers and commodity high-speed network technologies. The **FELIX** firmware designed for the Phase-I is in an advanced stage of development [11.7]. The design provides a good estimate of **FPGA** resource utilisation, from which it is possible to make extrapolations for **FELIX** in Phase-II. The system size presented in this chapter is based on these extrapolations. By basing prototype implementations on this architecture, it will be possible for **TDAQ** and detector communities to explore how the required Phase-II functionality and performance could be achieved. As an example, the Phase-I **FELIX** is being used to demonstrate the capability to perform complex calibration procedures for the **ITk** in the context of the **ITk** Pixel demonstrator programme. Results are expected in the course of 2018 and will make it possible to tune both the future design and implementation based on first-hand experience and collaboration between **TDAQ** and detector communities. A similar programme of R&D will occur for **DCS**, in order to inform final decisions as to the location of **DCS**-specific processing operations (from the network and software layer to the level of the **FELIX** firmware itself).

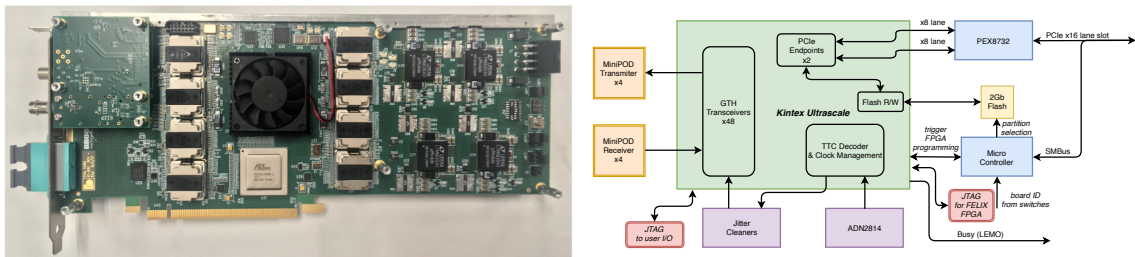


Figure 11.4: Phase-I **FELIX** hardware platform, known as the **FLX-711**. Final prototype shown on the left, block diagram on the right.

Similarly, by evolving the Phase-I **FELIX** firmware and Readout software it will be possible to study the optimisation of **FPGA** resources in order to deal with a higher number of **FE** Links. It will also be possible to develop and evaluate more complex data routing algorithms based on the **Level-0 IDentifier (L0ID)/BCID**, as well as the implementation of 'vertical or horizontal' data aggregation before forwarding the data to the Data Handlers. In this context vertical aggregation refers to the grouping of multiple event fragments from the same logical detector link before forwarding the data to a Data Handler; whereas horizontal aggregation refers to the aggregation of data fragments from multiple logical detector links with the same **L0ID/BCID** before forwarding them to a Data Handler. **FELIX** data-routing capabilities are an important ingredient not only for the design of the Data Handler, but also for the **DAQ** system as a whole. For instance, being able to route data based on the **L0ID/BCID** opens the possibility of partitioning the Dataflow/**EF** system into identical subdetector-wide slices. Each slice would be made of a subset of Data Handlers, each connected to all Readout paths from a subdetector, each seeing all of the subdetector's output for any event for a fraction of the **L0A** rate.

In addition to these areas of study, R&D will continue on the matter of high-speed server buses, either evolutions of [PCIe](#) or any alternatives that may appear in the coming years. Work will also focus on high-performance I/O software. Studies will be performed in the areas of data compression, which may be useful in minimising network bandwidth as well as Dataflow storage requirements. Last but not least, alternative form factors for the [FELIX](#) and Data Handler may be considered, assuming they retain the modularity of the [FELIX](#) concept in terms of independent upgradability of computing resources, network technologies and serial-link handling.

### 11.5 Dataflow

The Dataflow system buffers, transports, aggregates and compresses event data. It is responsible for the transport of data from the output of the detector Readout to CERN permanent storage. The main functional elements of the Dataflow system are the Event Builder, the Storage Handler and the Event Aggregator. The Event Builder is the logical interface with the Readout System. It receives event data from Data Handler systems and assembles event records, associating each incoming event fragment with its parent. The Event Builder also manages the storage volume of the Storage Handler system. The Storage Handler is a high-throughput large-volume storage system which buffers event data before and during processing by the Event Filter. For events accepted by the Event Filter, the Event Aggregator collects, formats, compresses, and transfers the output to CERN permanent storage. Fig. 11.1 shows these main functional blocks as part of the Phase-II Dataflow system.

The Dataflow system receives events from the Readout System's Data Handlers at 1 MHz and buffers them in the Storage Handler with total writing (i.e. input) traffic of 5.2 TB/s. The size of this traffic is driven by the overall event size. This is shown in comparison to Phase-I for the input stage of the system in Table 11.3. The Storage Handler provides the Event Filter with access to Level-0 trigger information, and approximately 10% of data, at 1 MHz and full event access at 400 kHz for a total reading (i.e. output) traffic of 2.6 TB/s. The Event Aggregator receives the full event record at the output rate of 10 kHz for a total throughput of 60 GB/s out to CERN permanent storage. The traffic requirements for the different components of the Dataflow system are summarised in Table 11.4.<sup>2</sup>

The next sections describe the different components of the system in more detail and provide information independent from the implementation. The last section presents and discusses the implementation-dependent aspects.

---

<sup>2</sup> Not including contributions from the *High Granularity Timing Detector* or small contributions from other subsystems.

Table 11.3: Average Event Size (before *HLT* selection) as extrapolated from Run 2 data to the Phase-I conditions (including the new detector systems introduced by Phase-I upgrade), and from estimated by the detector systems in Run 4 after the Phase-II upgrades. Forward detectors are not listed since the associated event size is negligible.

Detector System	Extrapolated Data Size [MB]	
	Phase-I	Phase-II
Pixel	0.3	2.4
Strip	0.2	0.5
TRT	0.5	-
LAr	0.7	0.7
Tile	0.1	0.2
Muon	0.5	0.8
TDAQ	0.6	0.6
<b>Total</b>	<b>2.9</b>	<b>5.2</b>

Table 11.4: Phase-II Dataflow traffic requirements.

Component Connection	Traffic	
Detector Front-ends to <a href="#">FELIX</a>	5.2 TB/s	
<a href="#">FELIX</a> to Data Handlers	5.2 TB/s	
Data Handlers to Event Builder/Storage Handler	5.2TB/s	
Storage Handler to Event Filter	2.6 TB/s	
Event Filter to <a href="#">HTTIF</a>	Event Filter to <a href="#">rHTT</a>	175 GB/s
	Event Filter to <a href="#">gHTT</a>	560 GB/s
Event Filter to Event Aggregator and Permanent Storage	60 GB/s	

### 11.5.1 Event Builder

The Event Builder is the logical interface between the Dataflow and Readout systems and acts as the primary intelligence managing event data across the Dataflow system. It is also the mechanism by which backpressure from the Dataflow and downstream systems are communicated to the Readout system.

The Event Builder will be implemented as a software interface to the Storage Handler. The Data Handlers will write data to the logical storage volume via this interface. The interface will be responsible for associating event data with their parent [LOA](#) as they are received. Across the Event Builder system events must be tracked at the [LOA](#) rate of 1 MHz. In the case of an operational issue the Event Builder interface will throttle Data Handler traffic to reduce the rate of written data, by asserting back pressure to the Readout system.

### 11.5.2 Storage Handler

The Storage Handler buffers data received from the Readout system through the Event Builder to decouple the Readout and Event Filter. The large storage volume needed to achieve the required 7.8 TB/s aggregate I/O traffic allows for increased event processing time, the potential of processing deferment, and an increase in operational robustness in the [EF](#).

The Readout system accesses the Storage Handler through the Event Builder software layer. Event Filter processors are then notified of the presence of new data within the storage volume and commence processing as permitted by their current load. Accepted events are then made available to the Event Aggregator for transfer to permanent storage.

### 11.5.3 Event Aggregator

The Event Aggregator system receives events from the Event Filter, performs compression and prepares output files for transfer to permanent storage. The Event Aggregator receives the 60 GB/s output from the Event Filter farm. In terms of implementation the system will make use of a common storage hardware platform shared with the Storage Handler.

The Event Aggregator provides a buffer area capable of storing up to 48 hours of accepted event data as required by CERN-IT. This decouples online data taking operations from non-ATLAS systems and enables the Dataflow system to cope with disruptions or malfunctions of the data transfer to the permanent storage systems outside of ATLAS experimental control.

The Event Aggregator will be responsible for ensuring that events associated with different output streams are sent to Tier-0 as part of the requested stream. The decision as to which stream an event should be associated with will be taken by the Event Filter.

The Event Aggregator will support partial event building based on the result of Event Filter processing.

Finally, the Event Aggregator will be responsible for any necessary communication with the Tier-0 processing centre if anything beyond data transfer is required.

### 11.5.4 Storage System Sizing & Technology

Looking at currently available technology, and extrapolating based on industry predictions, it is possible to produce an estimated scale for the Phase-II storage system. Current commercially available data centre solutions provide storage space within performance-tiered clusters using a variety of different media types with different throughput and storage capabilities including: [HDD](#), [SSD](#), [Non-Volatile Memory Express \(NVMe\)](#) and random-access

memory (RAM). Adopting the same policy, with the use of different media types according to role, could allow the system to meet its high-throughput requirements while using the most appropriate and cost-effective storage type for the component in question. For example, on the Data Handler side, very high-throughput media such as RAM could provide the necessary performance, while more reliable media such as HDDs could provide the stability needed for the Event Aggregator side. For the Storage Handler, the throughput requires that Solid State Drives (SSD) be the dominant technology. The system is sized according to this assumption.

Over the past ten years HDD capacity grew by a factor of 10 and the trend is expected to continue through to 2025. The same improvement in capacity can be reasonably expected for SSD technologies. Today the highest capacity SSD is of order 4 TB. Today's SSD technologies provide 2 GB/s of aggregated throughput; this throughput scales with capacity up to the extent that the connectivity can support. With PCIe Gen4 connected SSDs, we assume 5 GB/s achievable throughput as a baseline for future PCIe connected SSDs which will have a capacity of 20 TB.

The Storage Handler system needs to provide 5.2 TB/s of writing and 2.6 TB/s of reading traffic. The full 7.8 TB/s I/O throughput can be provided with 1800 SSDs providing 36 PB of total storage. This represents more than an hour of event buffering at the full Level-0 rate before and during Event Filter processing.

To meet the requirements of CERN-IT the Event Aggregator must buffer 48 hours of full event data at 10 kHz with an output event size of 6 MB, giving a storage volume requirement of 10 PB. This capacity would require 500 SSDs which are naturally included as part of the Storage System needed for the Storage Handler. The required output throughput of 60 GB/s is negligible compared to the total 7.8 TB/s of the Storage Handler, and a common storage system is design for the input and output buffer to more efficiently use resources and provide greater system flexibility.

As a whole, the storage volume supporting the Dataflow system could consist of 1800 SSDs supporting 5.2 TB/s of input traffic, 2.6 TB/s of output traffic and 36 PB of storage capacity, among which up to 10 PB will be used by the Event Aggregator.

### 11.5.5 Factors Affecting Implementation

Due to the high-throughput requirements, the design of the Dataflow system is closely coupled with that of the network system which supports communications (see Fig. 11.8). The ideal case, where the network can be considered an infinite resource linking every system to every other system (i.e. all-to-all) regardless of functionality or physical location is not feasible from a cost or logistical perspective given the required bandwidths. The Dataflow system design and implementation must take into account the constraints emerging from the supportable network topology and bandwidth provided between clients.

From the Dataflow point of view the network design provides higher-bandwidth and lower-latency connections between subsets of Data Handlers and subsets of Storage Handler units by connecting them all via the same switch hardware within a subset. Every Data Handler has a set of storage units for which it will be much more efficient to write to. Connectivity between a specific Data Handler and the Storage Handler units not directly connected to its switch is provided by interconnecting switches through an intermediate router. The local network areas are referred to as Dataflow *slices* as shown in Fig. 11.5. Intra-slice communication will be considered efficient and fast while inter-slice communication will be considered limited.

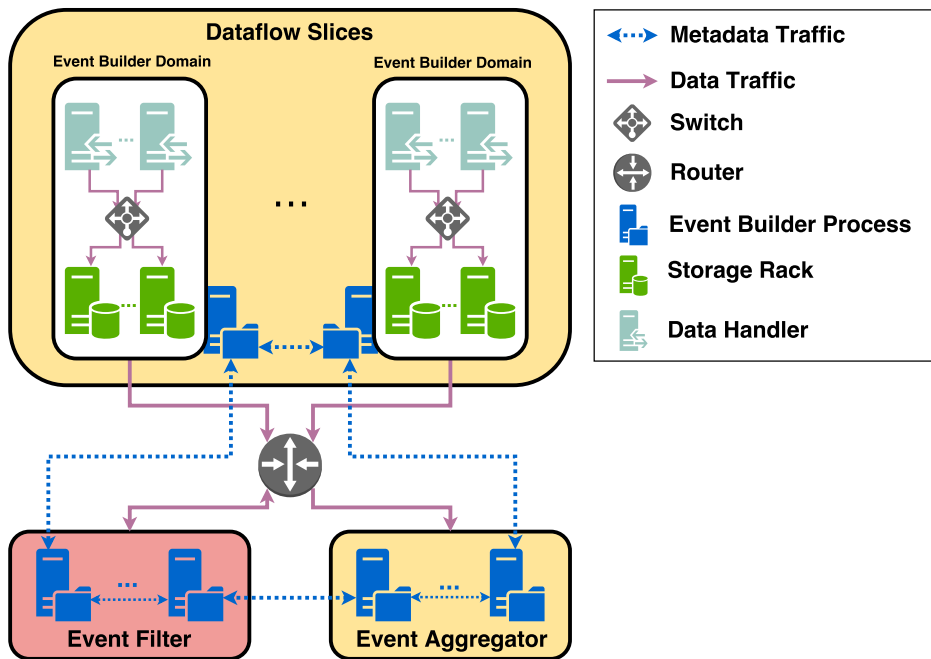


Figure 11.5: Logical flow of the Dataflow system showing the slices of network traffic from Data Handlers to preferred Storage Handler units.

### 11.5.6 System Workflow

Fig. 11.6 shows a high-level description of the communication between the Dataflow components. It describes a comprehensive understanding of the life cycle of an event within the Dataflow system and remains true regardless of the final implementation.

The Data Handlers write data to the storage system using the Event Builder interface. The location of the data is irrelevant to the Data Handlers and the Event Builder guarantees that written data is optimally placed within the Dataflow slice and will eventually be accessible to the Event Filter. This communication step is used by the Event Builder to prevent Data Handlers from writing new events, and therefore asserting back pressure to the Readout

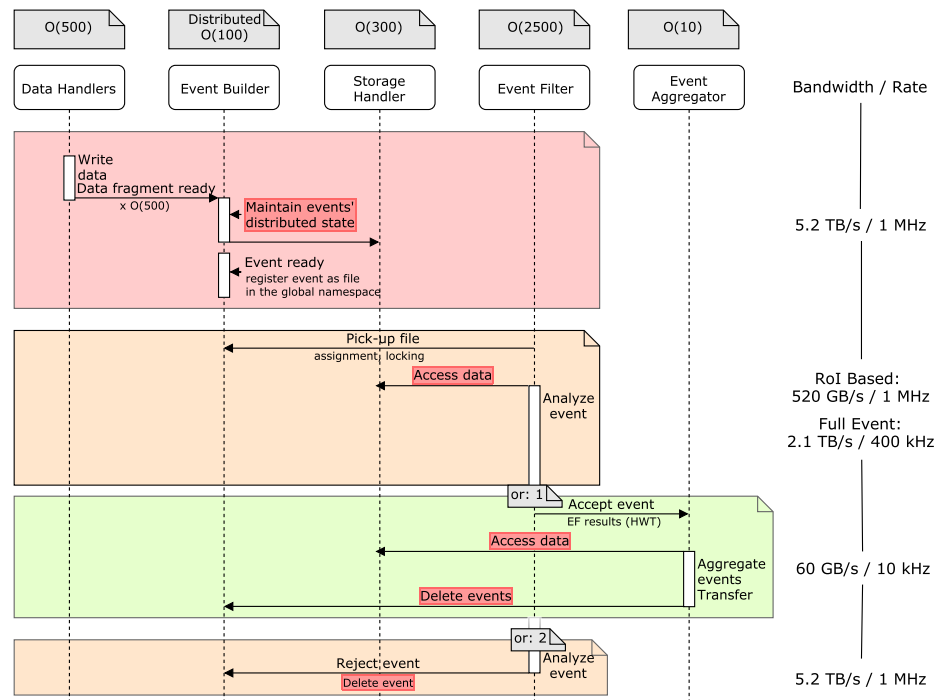


Figure 11.6: Logical communications between different components of the Dataflow system. The red boxes show the communications involving inter-slice traffic, the dashed Vertical lines indicate the logical component responsible for each action. The arrows represent all communication between components, including the initial message (indicated by arrow direction) as well as any response and follow up high bandwidth traffic if present. The Event Builder serves as an interface for both the Data Handler to write data to the storage volume and the Event Filter Processing Units reading data from the storage volume.

system. The main task of the Event Builder is to gather the data from the different Data Handlers into a single event entity. The event entity can be a physically contiguous file or a logical aggregation of independent blocks distributed across the storage cluster. By tracking the event entities the Event Builder is responsible for the maintenance of the Storage Handler volume. This event building step necessarily involves inter-slice communication. In the logical event building case, maintaining the state of the distributed event fragments accounts for far less network traffic than the event data itself needed for physical event building. For optimal use of storage media, as well as network resources, sets of built events could be aggregated into larger files for access by the Event Filter. The exact number of events in each file would then be dependent on the technologies in use.

Once an event is built it is published by the Event Builder to the Event Filter and can be requested by or pushed to a given Event Filter node. The Event Filter sees the event as part of a single file whether or not it is physically built. From the storage system and network point of view, if the file is contiguous or distributed in blocks implies very different access



patterns. In either case there is many-to-one communication at some point in the chain that requires special consideration in the design of the Dataflow and network systems.

The two different forms of event data (distributed and contiguous) may coexist in the system. The Event Builder could, for example, initiate a background physical event building process for long-lived events when computing, storage and network resources are available. The Event Builder may also use a timeout policy to move event data to different media types throughout their lifetime. The Event Builder guarantees that these background activities are not visible to the Event Filter or the Event Aggregator.

If the event is accepted by the Event Filter it writes the result of the Event Filter processing to a designated location within the event file, and the Event Aggregator is informed of the file and can access it for aggregation and external transfer. The result written to the Event Aggregator can include information governing output stream allocation and partial event building. When the transfer is complete the file is deleted to release storage resources. If the event is rejected, the file is similarly deleted from the storage system.

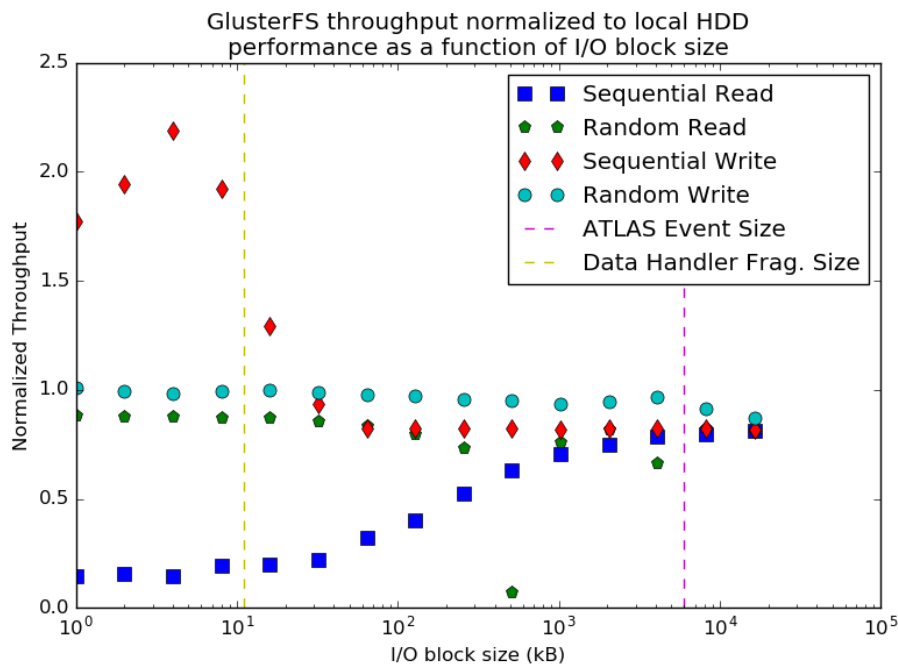


Figure 11.7: The measured throughput of GlusterFS *DFS* normalised to the performance of a single local process. Measurements are shown for sequential reading and writing as well as random reading and writing as a function of IO block size. The distributed read and write processes are coordinated across multiple hosts allowing the performance to potentially be greater than the local process when throughput is limited by host CPU, for example sequential write of small block sizes.



**A DFS-based Implementation** The rapid evolution of both hardware storage technologies and software storage solutions makes the definite choice of an implementation premature, but from what is known so far, a proposed solution has been developed based on linking the different Dataflow system components through a distributed file-system (DFS). In a DFS-based implementation, all system components (i.e. Data Handlers, Storage Handler, Event Filter, Event Aggregator) host part of the file-system on their local storage media. The location of all data on the system is tracked by metadata managed by the DFS. All Dataflow operations can then be enacted by modifying file metadata to schedule transfer to another storage volume, and the synchronisation of metadata between hosts. In this model, the event data files for processing exist in a global namespace, which is visible to all Dataflow system components.

The first step in the processing of event data is for the Data Handlers to open a file, located in the Storage Handler volume, and start writing into it. Local metadata servers then implement event building functionality by ensuring that the locations of event data are correctly tracked across the system. In this way the Event Builder component of the system is implicitly implemented without the need for more dedicated technology. The existence of the file in the global namespace will make it visible and accessible to the Event Filter and the Event Aggregator. Processing or accepting an event can then be achieved by moving the file to a specific directory monitored by the Event Filter or Event Aggregator, therefore making it a metadata-only operation. Event rejection can itself be implemented as a simple file deletion which is also a metadata-only operation. In the proposed solution the Storage Handler system provides its own storage units and devices but may also take opportunistic advantage of other systems' storage media to aggregate storage space where possible and pertinent.

From a networking perspective, DFSes like CephFS [11.8], GlusterFS [11.9], or OctopusFS [11.10] provide a network topology-aware placement of file blocks. With this it is possible to ensure each Data Handler will only write event fragments to parts of the file within its Dataflow slice. There is therefore no need for large scale data transfer between the storage system slices, only the transfer of the metadata needed to keep track of the contents of the file in each slice. The DFS metadata contains the storage block-to-file level information as well as the usual file system metadata (permissions, timestamps, etc.). At present, no DFS supports the opening of the same file from several hundred writers, but CephFS and OctopusFS do offer distributed metadata management that is required for sliced event building.

Storage tier-aware DFSes like OctopusFS or ScaleIO [11.11] take care of automatically managing the different capabilities of mixed media storage to balance throughput and reliability. Most DFSes (HDFS [11.12][11.13], GPFS [11.14], GlusterFS, OctopusFS, etc.) have built-in mechanisms to manage data replication to provide device, unit and rack fault tolerance. These mechanisms can be disabled to save storage space and storage and network bandwidth, and potentially overloaded to implement an automated Event Filter processing timeout mechanism.

In Fig. 11.7 the performance impact of a DFS is compared with the throughput that can be achieved by a single local process as a function of application-to-storage communication block size. Shown for GlusterFS, the writing performance is at or above the local throughput for I/O block sizes up to the ATLAS event size. The distributed read and write processes are coordinated across multiple hosts allowing the performance to potentially be greater than the local process when throughput is limited by host CPU, for example sequential write of small block sizes. Random read access shows a degradation of less than 10%. The CPU overhead associated with reading data from a DFS relative to the local storage media has been measured in CephFS, ScaleIO and GlusterFS to be negligible at less than 1-5% of the system CPU. Any performance impact associated with a DFS implementation is assumed to be within the uncertainty of the functional requirements of the Dataflow system.

The implementation that has been described in this section is based on the known functionality of third-party DFSes used as points of reference, with third-party solutions in general seeming the optimal choice for any final implementation. Such solutions may require additional customisation, the level of which depends on the evolution of the market. The final form of the system will favour commodity solutions where possible, with the desired features having been highlighted. In the event that the development of a custom solution provides a more efficient and cost-efficient approach it will be adopted. Market survey, experimentation and specific studies will be necessary to maintain an up-to-date knowledge of the available third-party solutions' capabilities and limitations. DFSes that minimise inter-slice communication to metadata-only traffic, trading off extra latency to achieve the required throughput, will be considered the top candidates. Automatic mechanisms internal to DFSes (e.g. load balancing, data replication, tiered-cache management) will provide a smarter and more upgradeable system supported by a potentially large developer community.

## 11.6 Network

The Readout and Dataflow components discussed so far rely on a high-throughput network for all communication and data transport. In this section the architecture of such a network will be discussed, focusing not just on data transport, but also on reliability and redundancy.

### 11.6.1 Network Domain Description

The ATLAS DAQ system consists of two well-separated network domains. The first domain, known as the Readout network (or FELIX network), is a high-throughput system connecting the FELIX servers and the Data Handlers with additional requirements for DCS, control and monitoring infrastructure interconnection. The high-level architecture for this

component is envisaged to be an extension of the Phase-I Readout network [11.15] comprising all the ATLAS subdetectors.

The second network domain, known as the Dataflow network, provides connectivity between all the Dataflow components and the EF farm. Its architecture has to be completely revised with respect to the Phase-I system as requirements will change significantly.

## 11.6.2 Network Technology Choice

There are currently several network technologies with promising future performance projections in terms of bandwidth and maximum latency that can be considered for the Phase-II networks. These are Ethernet, InfiniBand and OmniPath. While Ethernet is the market-leading technology for a wide range of applications (including data centre interconnection), InfiniBand and OmniPath are technologies developed specifically for High Performance Computing applications. It is not currently possible to make firm conclusions about the most suitable technology for Phase-II, as many factors are currently unknown and difficult to predict. For example: price per port and per Gigabit, lossless operation of Ethernet, effectiveness of remote direct memory access (RDMA) for high-speed Ethernet, medium-range link range, etc.

The current baseline is a system completely based on Ethernet. However, the use of two different network technologies in the two network domains presented is not excluded. While to first order this may appear inefficient, there might be some requirements only attainable with a specific technology (e.g. ultra-low latency communication in the Readout network) for each use case.

In terms of technology assumptions which can be made at this point, the first one is that by the time of Phase-II it will be possible to purchase 100/400 GbE devices, 400 Gb/s 'Next Data Rate' (NDR) Infiniband or its equivalent in OmniPath. This assumption is supported by the latest official roadmaps [11.16][11.17]. The second assumption is that Ethernet NICs, InfiniBand host channel adapters (HCAs) and OmniPath host fabric adapters (HFAs) will have enough bandwidth towards the system CPU and memory to run at the required speed [11.18].

In order to select the best suitable technology the R&D strategy has been divided into three steps: technology tracking, technology evaluation and network slice construction. The technology tracking will be performed with Mellanox (for InfiniBand), with Intel (for Omnipath) and with the typical Ethernet vendors that regularly visit LHC experiments. Following the Phase-I R&D model, early stages of technology evaluation should be performed in High Performance Computing (HPC) clusters owned by third parties. This was proven to be enough to learn the basics of a given technology. Finally, when the candidate technology is selected a testbed setup should be built to perform an in-depth evaluation with a real DAQ system slice.

### 11.6.3 Network Architecture

As mentioned above, the architecture for the Readout network is similar to the Phase-I [FELIX](#) network. The design is based on organising the system in slices and connecting all the nodes in a given slice to the same high-throughput device. Each slice contains a certain number of [FELIX](#) servers and the set of Data Handlers they need to communicate with. All the network slices are connected together with a pair of routers on a higher hierarchical level allowing all-to-all communication for [DCS](#), control and monitoring purposes.

The redundancy of the described setup could be improved by connecting the [FELIX](#) server links to a pair of switches working in active-active mode. This solves the problem of handling individual link failures (the most common type of failure) making use of the full capacity installed in normal operation. On the other hand, the service would be degraded, but not stopped, in case of a switch failure.

For the Dataflow network, the proposed architecture is based on aggregating Data Handler servers with pizza-box switches and connecting the Storage Handler units to the same network switch. A set of Data Handler nodes will then always write to the same set of Storage Handler units, maximising the throughput and reducing the latency. The switches need high-throughput uplinks to the network core routers, so as to allow all-to-all communication between the Storage Handler and the rest of the system.

For Event Filtering connectivity, the servers are stored in racks and connected to Top-of-Rack (ToR) switches. Every ToR switch is connected to the core network with enough uplink capacity to ensure the required throughput. Finally, events aggregated by the Event Aggregator need to be sent to permanent storage. For this, high-throughput Long-Range links are needed between the core routers at Point 1 and the CERN data centre.

A baseline network implementation is depicted in Figure [11.8](#). The shape of the core router cluster is an implementation detail and different solutions can be envisaged: from a pair of chassis routers to a distributed leaf-spine switch topology [[11.19](#)]. The advantage of using a router cluster is that it can provide a basic level of network redundancy simply by connecting pairs of uplinks to different devices.

The described architecture is suitable for Ethernet networks. However, if an [HPC](#) technology were to be used (where all the devices have the same link speed), it could be possible that link aggregation could be achieved by installing many links of the same speed. In this case, the number of ports available on the network devices for connecting endpoint hosts would be reduced and even more devices would be needed. In this situation, the architecture of the network would certainly need to be revised and tailored to a given technology.

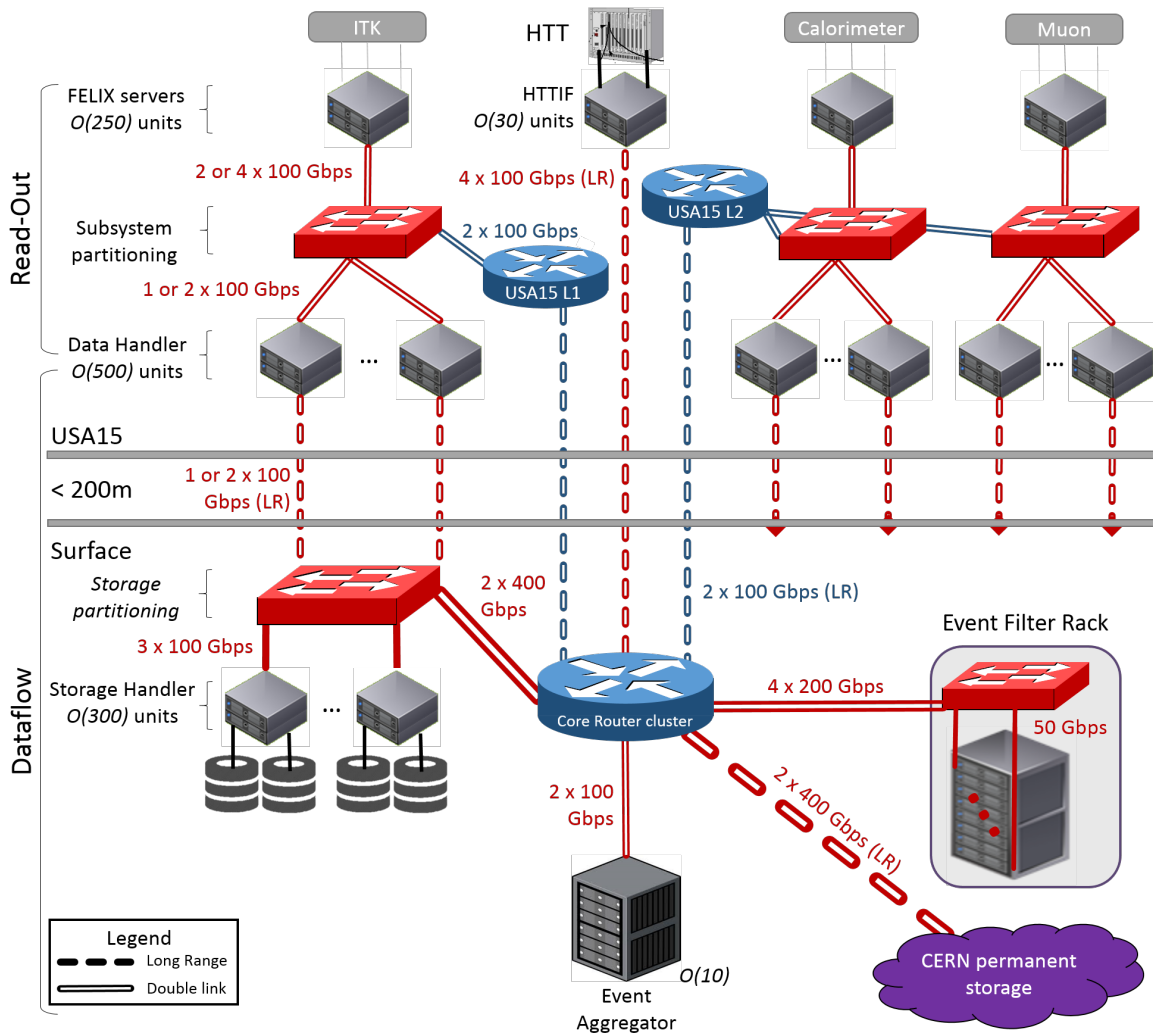


Figure 11.8: Baseline network implementation using Ethernet technology. Network redundancy scenarios are not depicted.

## 11.6.4 Control Network

The amount of control traffic during a data taking session is not expected to grow significantly in the new system. Therefore, the need for a dedicated network is not considered necessary for control and monitoring traffic. Instead, new virtual networks can be defined on top of the available high-throughput devices and make use of traffic engineering and Quality-of-Service techniques to guarantee the non-interference of different flows. Moreover, the control traffic can make use of all the installed network capacity during software infrastructure initialisation and configuration steps. On the other hand, a flat non-redundant 1 Gbps Ethernet network may still be needed for server management traffic via [Intelligent Platform Management Interface \(IPMI\)](#).

### 11.6.5 Network Installation

In the case where the final network design is Ethernet based, the available Medium-Range interconnection technology will have an impact in the network installation. There are currently two industry standards used in big data centres (not yet standardised by IEEE) for 100 GbE: PSM4 and CWDM4. PSM4 stands for "Parallel Single Mode 4 lane" and can reach 500 m. CWDM4 stands for "Coarse Wavelength Division Multiplex" and can reach 2 km. For 400 GbE similar solutions are expected to cover similar distances for reasonable prices. However, depending on the price ratio between Medium-Range 100 and Medium-Range 400 GbE the installation of an aggregation layer could be performed in the ATLAS service cavern (USA15) or in the surface data centre (installation of many 100 GbE links versus installation of fewer but more expensive 400 GbE links).

For practical purposes, patch-panel based solutions are preferred over active-fibre based ones for USA15 to surface links [11.20]. The reason is that with active cables we need to account for enough spares to deal with all the transceiver failures that might arise during operational periods. Laying new active cables is only possible during scheduled maintenance periods such as the yearly winter shutdown. On the other hand, with a patch panel solution we only need to account for spare fibres, which are passive components with low probability of failure, and account for few spare transceivers that can easily be replaced when necessary.

## 11.7 Online Software

So far in this chapter the focus has mainly been on architecture and connectivity of individual components of the [DAQ](#) system. Another core part of the design is the software infrastructure which will manage and operate the components of the system, and integrate them with the wider ATLAS data taking environment. The Phase-II upgrades provide an opportunity for significant re-designs and improvements of various parts of the system to make use of new technologies and standards. In this next section a detailed description of the proposed upgraded online software infrastructure will be presented, alongside an overview of the studies currently in progress in order to better understand and explore all aspects the design. Most of the software elements within the online setup will be maintained and upgraded adiabatically in the next several years by the ATLAS [DAQ](#) Operation group. However, some of the core software infrastructure may be significantly redesigned to benefit from the availability of new technologies

### 11.7.1 System Overview

The online software infrastructure used to facilitate ATLAS data taking is comprised of a small set of common libraries on which the [DAQ](#) and offline software projects depend. The



software stack is built on top of a large number of external libraries, which are provided in a dedicated release by the [LHC Computing Grid \(LCG\)](#) project. Most of the online software is written in C++, although some of the [DAQ](#) software uses Java, and Python is the standard for more complex scripting tasks. With the recent move to git for version control and CMake [11.21] as a build system, most of the software infrastructure is now using standard open-source tools for these tasks, as well as for continuous integration and testing.

The [DAQ](#) software contains components responsible for configuration, control and monitoring of the full ATLAS system during operations. Few fundamental changes are expected in terms of the responsibilities of this part of the system. The main exception is the increased use of [ATCA](#) over [VME](#) for Phase-II. The current software provides support for Single-Board Computers inside [VME](#) crates, which control much of the front-end hardware and Readout. The transition to [ATCA](#) for Phase-I seems to go hand-in-hand with solutions where a processor is integrated directly into an [FPGA](#) hosted on custom hardware. The processors are often [ARM](#) processors instead of Intel, and might run a different Linux distribution than the rest of ATLAS. Effort will be needed to provide coherent support for configuration and control of these systems across ATLAS.

The data acquisition part of the online software is expected to need to be completely rewritten compared to Run 2 and Run 3. This begins with the Data Handler component, which is the interface to [FELIX](#), and includes the Event Builder, Storage Handler and Event Aggregator.

Another theme of the software upgrade will be the move to containerised solutions. There is an ongoing effort to reduce the dependencies between online and offline software as part of the Phase-I upgrade. This will make it possible to run Event Filter processes in their own environment inside a container as described in the next section. The same approach can be used for monitoring based on the offline Athena framework [11.22]. Monitoring in general will be a major component of the upgrade effort. As discussed below a number of changes will be required to the current monitoring infrastructure to support the proposed new system architecture.

Furthermore, control and monitoring software currently runs on dedicated machines with some limited support for failures of critical services. Moving as many of these services as possible to containers and running them on a common infrastructure will open new possibilities for fault tolerance, reliability and scalability.

Underpinning the operation of all upgraded systems will be the role of system administration, stretching from process monitoring to infrastructure management and facilitation. Such matters will be covered throughout this section, culminating in a more dedicated discussion once all the areas of interest have been introduced.

### 11.7.2 EF Farm Management

The EF computing farm hosts both the processing units (PUs) and all the supporting services needed to implement the last step of the event selection component of the TDAQ system. The farm is made up of several thousand commodity servers, each executing one or more processes. A robust and reliable mechanism for the management of all processes running in the EF farm is a requirement to guarantee stable and efficient execution of the EF service. In modern software architectures, the management of large clusters of computing nodes is delegated to so-called 'Cluster Orchestrator' services. In a system like the TDAQ system, a Cluster Orchestrator will fulfill a series of well-defined basic requirements:

- It will support different types of application lifecycle (i.e., always-running, run-to-completion and cron-like services);
- It will allow both elastic and static allocation of processes to computing nodes;
- It will be able to dynamically handle cluster resources (i.e., enabling/disabling computation units at runtime, efficient exploitation of the available CPU power and memory);
- It will scale to thousands of hosts;
- It will make it possible to control (i.e., starting, stopping) and monitor the status of all active processes;
- It will make it possible to completely describe the requirements for all processes needing to be started, including the definition of command-line parameters and environment variables to be passed to the executable.

A survey of the offers currently available on the open-source market (based on the requirements above) highlighted Kubernetes [11.23] to be a perfect candidate as an orchestrator for the EF computing farm.

Kubernetes was announced by Google to the open-source community in 2014 [11.24] and is based on 15 years of experience at Google in managing and orchestrating large clusters. Since its first release, the Kubernetes open-source community has experienced steady - growth, reaching more than 1500 commits per month and more than 150 contributors per month in February 2017 [11.25]. Today, Kubernetes is a mature product contributed to by several technology partners like RedHat, CoreOS and Intel.

Kubernetes can be described as 'a system for automating deployment, scaling and management of containerised applications'. Among several supported features, Kubernetes provides a set of services facilitating easy and effective management of applications in a cluster:

- Scheduling of applications based on required resources and other constraints;
- Automatic re-scheduling of applications when the application itself fails or the node where the application is running dies;
- Built-in support for service discovery and load-balancing;



- Management of several storage back-ends, allowing transparent mounting of both local and network storage volumes;
- Easy (via command line tools or UIs) and automated (based on CPU usage) application scaling.

Kubernetes requires applications to be packed into software containers. Containers exploit virtualisation at the level of the operating system and are lightweight and simpler to build than Virtual Machines, which instead exploit hardware virtualisation. Packing an application into a container makes it possible to create immutable images disentangling the application itself from the host operating system. In such a way, containers do not only provide strong resource isolation but also make the development, integration and deployment cycle easier, thus simplifying software portability and distribution.

Kubernetes supports Docker [11.26] containers. Docker is currently the market-leading container platform.

**EF Processing Units in Software Containers** As a proof of concept, a small Kubernetes cluster (4 nodes) was set up using the CERN IT Virtual Infrastructure [11.27], with the goal of running EFPU instances in software containers. The PUs themselves were emulated with the offline version of today's HLT software (i.e., AthenaHLT), using a realistic trigger menu. A Docker container image was created starting from a base SLC6 [11.28] image and adding a few additional packages. The EF software was retrieved directly from the CVMFS [11.29] installation repository. Kubernetes was able to transparently mount the CVMFS volume, making it possible to keep the size of the Docker image to a few hundred megabytes. In order to better simulate data processing activity, two additional mount points were added to the container: an *input* directory with data files containing real events, and an *output* directory receiving the results of event selection algorithms. The AthenaHLT image was distributed to and executed by the Kubernetes cluster. Events were correctly retrieved from the input data files and processed by the PUs, with selection results stored in the output directory.

**Performance and Scaling** In the most recent release (1.6) available at the time of writing, Kubernetes is able to handle computing clusters with up to 5000 nodes populated with up to 150,000 containers [11.30]. To satisfy this requirement Kubernetes satisfies two performance goals for the reported cluster size: 99% of the calls to its backend (e.g., the calls to inspect the state of a container) return in less than 1 s and containers (with pre-pulled images) are able to start within 5 s with a probability of 99%<sup>3</sup>. It is worth noting that the

---

<sup>3</sup> Performance of a big cluster is sensitive to the size of the Kubernetes 'master' node (i.e., the cluster control plane). The Kubernetes documentation indicates that the reported performance figures were achieved running tests on the Google Compute Engine using a *n1-standard-32* [11.31] virtual machine for the master node.

performance goals were achieved with the cluster being *fully populated*; giving also an estimation of the time needed to restart a container in case of failures. Larger clusters are also supported, but with degraded performance.

The time needed to completely fill the cluster represents another crucial performance and scaling figure, particularly important for system operations. Since Kubernetes does not provide any official result in that respect, some specific and dedicated experiments were performed on a cluster made up of about 1000 virtual cores. The cluster was organised in the following way: one Kubernetes master node (32 CPU cores and 60 GB of RAM) and 240 slave nodes (4 CPU cores and 8 GB of RAM each). All the nodes were equipped with CPUs of the Intel Broadwell family running at 2.2 GHz. The latest Kubernetes version (1.5) available on the CERN virtual infrastructure was used. The tests aimed at measuring the time needed to scale an application to a certain number of replicas (from one to five instances per host). In order to minimise the impact of the started applications on the measurement (i.e., they may consume CPU cycles competing with the Kubernetes system), a *pause* container was used and its image was pre-pulled into the cluster. Such a container sleeps for an undefined period of time after being started, with very minimal resource usage.

Fig. 11.9 shows the time needed to scale the *pause* container up to five replicas per node as a function of the cluster size, for a total of 1200 started containers in a 240 host cluster. The number of replicas was chosen to match the number of applications executed on each HLT host during Run 2. The size of the cluster could be easily changed by enabling or disabling the corresponding hosts in the Kubernetes scheduler. The measurements are reported for different values of the Kubernetes *Query Per Second (QPS)* configuration parameter set. The *QPS* set is used to set a limit on the maximum number of requests the different Kubernetes components can handle. In such a way it is possible to avoid overloading the system, resulting in a possible denial of service. The default *QPS* values are quite conservative and defined to safely allow Kubernetes to run on a wide range of hardware platforms. Tests were executed increasing *QPS* values to up to four times their defaults. Kubernetes performance proved to be strongly dependent on the *QPS* configuration. As an example, the time to fully scale the *pause* container to 1200 replicas in a 240 hosts cluster decreased from about 74 seconds down to 27 seconds with the most aggressive *QPS* configuration.

Table 11.5: Cumulative time to start the *pause* containers at the 25<sup>th</sup>, 50<sup>th</sup>, 75<sup>th</sup>, 95<sup>th</sup> and 100<sup>th</sup> percentile (240 nodes and 5 replicas per node).

	Percentile				
	25 <sup>th</sup>	50 <sup>th</sup>	75 <sup>th</sup>	95 <sup>th</sup>	100 <sup>th</sup>
QPS x1	27 s	42 s	56 s	69 s	74 s
QPS x2	16 s	24 s	31 s	37 s	45 s
QPS x3	12 s	17 s	23 s	26 s	35 s
QPS x4	10 s	14 s	18 s	21 s	27 s

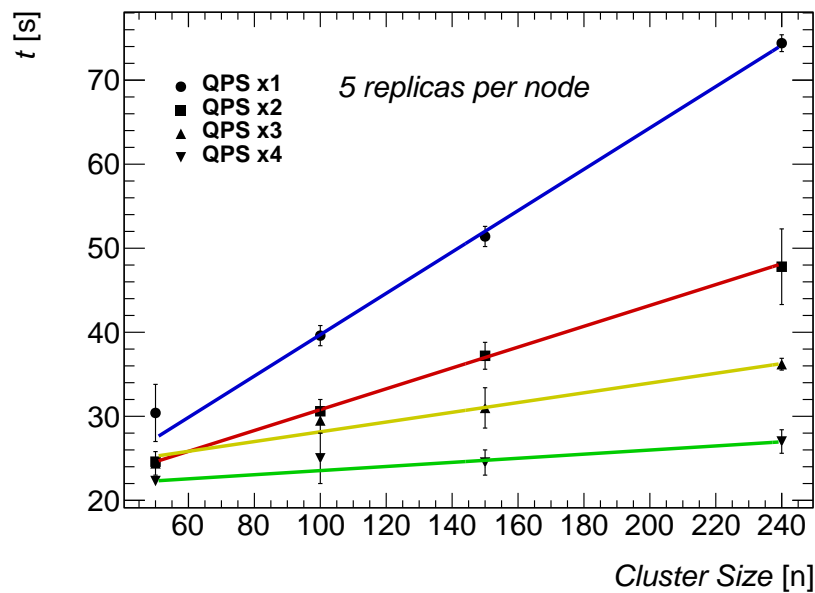


Figure 11.9: Time to scale the pause container up to five replicas per node as a function of the cluster size. The plot reports the measured times for four different values of the QPS parameter set.

The impact of QPS settings can be also seen in Fig. 11.10, showing the number of started containers as a function of time for different QPS values. The QPS does not only strongly impact the rate at which Kubernetes manages to deploy applications, but it also has an impact on the time needed to start the first container (from 11 seconds down to 5 seconds for QPS values four times higher than the default configuration). At the same time it is evident how the time needed to have all the containers up and running is dominated by a few outlier instances beyond the 95<sup>th</sup> percentile (see Table 11.5).

In general, the executed tests demonstrated that:

- The capability for Kubernetes to deploy containers in a cluster scales linearly with the size of the cluster itself. Assuming no higher order effects with larger clusters (Kubernetes officially supports 5000 hosts clusters), an EFPU service instance can be fully deployed on each node of a 2500 host cluster in about 30 seconds;
- Kubernetes performance is highly dominated by its QPS configuration. QPS values four times higher than their defaults make it possible to reach a sustained deployment rate of almost 70 containers per second (to be compared to about 20 containers per second with the out-of-the-box configuration). The Kubernetes development road map aims to reach a rate of 100 containers per second on a 5000 host cluster in upcoming releases.

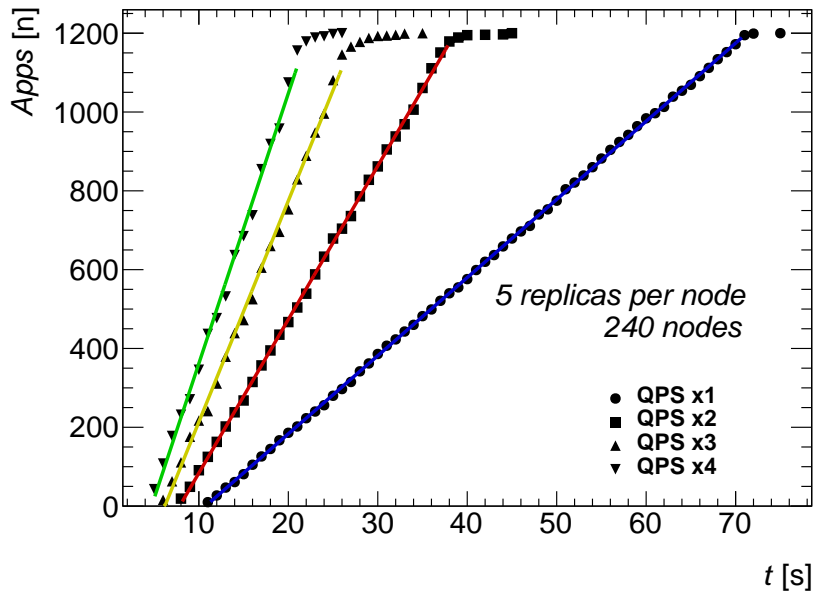


Figure 11.10: Number of started pause containers as a function of time for a cluster of 240 hosts and five replicas per host. Time is counted from the moment the command to deploy the containers is sent to Kubernetes. Measured times are reported for four different values of the QPS parameter set.

Overall Kubernetes performance has proven to be sufficient for its usage as an orchestrator of the EF computing farm. Even so, it will be worthwhile to keep monitoring upcoming Kubernetes releases in order to track and verify evolving performance figures.

### 11.7.3 Operational Monitoring of the Online System

The status and health of every host in the online computing system (from the current [HLT](#) farm to the control nodes) must be constantly monitored to ensure the correct and reliable operation of the whole online system. The monitoring system is not critical for data taking, but it is the first line of defence: it has to promptly provide alerts in case of failure and warn of impending issues whenever possible. The reason for this is to minimise the risk of downtime, which is particularly important for mission critical systems (e.g. [DCS](#) systems) which are also a single point of failure.

The currently implemented monitoring system is based on Icinga 2 [11.32], for active checks and alerting, and Ganglia [11.33] for data storage and to provide performance data for debugging and to complement Icinga 2 information. At the host level, the system is complemented by [IPMI](#) [11.34] and [Simple Network Management Protocol \(SNMP\)](#) [11.35] that are used to retrieve additional data directly from the host, for hardware health and system information respectively.

This implementation has knowledge of how the physical hosts are performing, and is able to monitor the status of various services. The monitoring system relies on ping requests, custom-built service checks, etc. to determine the state of the machine. These checks can generally ensure that the needed services and applications are running, but not necessarily how well they are running.

The monitoring system has proven its reliability and effectiveness in production, providing alerting and advanced performance monitoring. Recent improvements include the automation and improvement of checks and notifications, and the generation of Icinga 2 configuration files in order to simplify system maintenance and guarantee complete coverage of the online system.

The next evolution in operational monitoring would require a change in perspective. The implementation of business logic rules would allow the monitoring system to move away from checking specific aspects of a node (for example, network speed and disk status) to a more holistic, global overview of all the services that are being provided.

This would imply focusing on the services that the servers are collectively responsible for providing and to react to a service failure (rather than a machine failure). Of course, the information provided by hardware checks should not be neglected as it would still be necessary to maintain the overall physical health of the nodes.

With this perspective, the prioritisation and schedule of the alerts and notifications would have to be reviewed and adjusted to focus on what has to be immediately addressed and what can be addressed during the next scheduled repair opportunity (such as a technical stop). By applying these aspects of business logic, it would be possible to more effectively triage and react to the various issues that arise during day-to-day operations while still maintaining acceptable levels of service quality.

Looking ahead, a variety of new monitoring solutions are now available, with further developments expected in the coming years. With the Phase-II upgrade, there is an opportunity to re-evaluate the current implementation and investigate the possibility of either improving the current design or upgrading the monitoring system with new tools.

Following a survey of the open-source market, Prometheus was identified as a possible improvement to the current implementation. Prometheus is a system and service monitoring platform capable of collecting a number of metrics from any number of configured targets. It is able to evaluate rule expressions and can trigger alerts if some arbitrary condition is met. A key factor to consider is the scalability of the implementation – an aspect that Prometheus has been shown to handle very well. The largest publicly-documented Prometheus deployment monitors ~20 million nodes, each with ~13 exporters across 13 geographic regions using ~80 Prometheus servers [11.36]. However, caution should be taken when considering these new technologies, as many of them are intended to be used in cloud-based infrastructure, and is therefore not necessarily applicable to TDAQ servers.

Prometheus can also be integrated with the [TDAQ](#) software infrastructure (through the use of exporters and client libraries) in order to provide more in-depth health monitoring (known as 'white box' monitoring). Combining these checks with physical host and service monitoring (for example, the number of database queries per second) will allow a more complete overview to be built. By adopting this approach, it would be possible to monitor the actual performance of the various applications and components in order to detect issues sooner.

### 11.7.4 Physics Monitoring

**Motivation** The Online Physics Monitoring facility is one of the key tools for maximising the efficiency of the experiment. The system can be used to verify the quality of data being taken, as well as the trigger efficiency at all levels using several complimentary approaches. The primary mechanisms used are based on event sampling and the generation of monitoring histograms. In Phase-II the majority of the components of the Run 2 and Run 3 [DAQ](#) system, as well as the current High Level Trigger, will be replaced with new designs. As such, the monitoring infrastructure must adapt to service the new designs. In this section the requirements for the new system will be discussed.

**Event Sampling** The quality of the data being taken can be affected at any time by a malfunction of detector equipment or Readout electronics. Such an issue will equally affect all physics events, so this can be monitored based on statistical samples of the full event stream. The sampling of partial or full events can be done at different levels of the Dataflow chain allowing optimisation of the resources used for different types of monitoring.

A sampling-based approach could be used in Phase-II to assess, for example, the selection performance of the Level-0 trigger. Such an assessment cannot be done using a full event stream due to the high trigger rates. Conversely, Event Filter selection performance can be monitored with the analysis of a full event stream due to the distributed nature of the [EF](#) farm. [EF](#) monitoring is based on the analysis of dedicated trigger efficiency histograms, which are built by every individual [EF](#) processing unit and periodically gathered to obtain full [EF](#) statistics.

The changes in the Dataflow system architecture which are planned for the Phase-II upgrade will affect the implementation of Physics Monitoring, in particular the Event Sampling sub-system, as most of the [DAQ](#) components which supported event sampling for Run 2 and Run 3 will be removed or replaced for Run 4. This implies that a new sampling implementation will have to be provided for the new Dataflow system components. At the same time, requirements for event sampling will be reviewed in order to improve performance and usability of the Monitoring system based on experience obtained during Run 2 and Run 3. In particular, the new Event Sampling facility will be required to:

- Sample events from a specific physics stream;

- Sample events according to an arbitrary set of properties from event headers;
- Support unbiased sampling of events before the EF decision;
- Support sampling of full events as well as fragments of events which correspond to a particular slice of the detector or Readout electronics;
- Facilitate sharing of sampled events to independent monitoring tasks;

The new Event Sampling facility will need to interact with the new Dataflow system components, namely Data Handlers and Storage Handler, in order to fulfill the sampling requirements mentioned above. The design and the implementation of these components will impact the spectrum of technologies considered for the Event Sampling facility. Depending on the specifics of Dataflow component implementation it may be possible to use third party software solutions for event sampling, or it may be necessary to proceed with building and maintenance of a custom in-house software service for that purpose.

**Event Filter Histogram Management** The implementation of the Online Histogramming service, which provides histogram publication and gathering, will need to be adapted to the new Event Filter component. The monitoring of histograms produced by the EF will be vital to track the operational state of event selection code running in the new architecture. The transition to the new system will begin with a requirements review and a redesign of the existing facility. In particular, the new Histogramming service will be required to support the following operations:

- Publication of an arbitrary set of histograms by every EF processing unit
- Merging of identical histograms produced by EF processing units
- Display of a pre-selected subset of produced histograms
- Application of pre-defined Data Quality assessment algorithms to a pre-selected subset of the produced histograms
- Archiving of a pre-selected subset of produced histograms, both periodically as well as at the end of a data taking session

While some of these requirements are already addressed by the original implementation of the online monitoring, those related to communication with the new Event Filter component would need to be thoroughly analysed and a new implementation of an appropriate interface will need to be provided.

**Physics Monitoring High Level Design** In order to fulfill the requirements the Physics Monitoring (PM) infrastructure will need to communicate with the other sub-systems and applications as shown in Figure 11.11. The input information for the Physics Monitoring will be provided by the Dataflow sub-system, which will supply samples of raw physics events to be analysed by dedicated Monitoring Tasks, which will fill histograms using these data. The PM infrastructure will be responsible for shipping raw events from the Dataflow



sub-system to these monitoring applications as well as for storing the histograms they produce. The PM infrastructure will also expose the stored histograms to the other applications, which can use them for data quality assessment, archive them for off-line analysis or display to the experts and the shift crew.

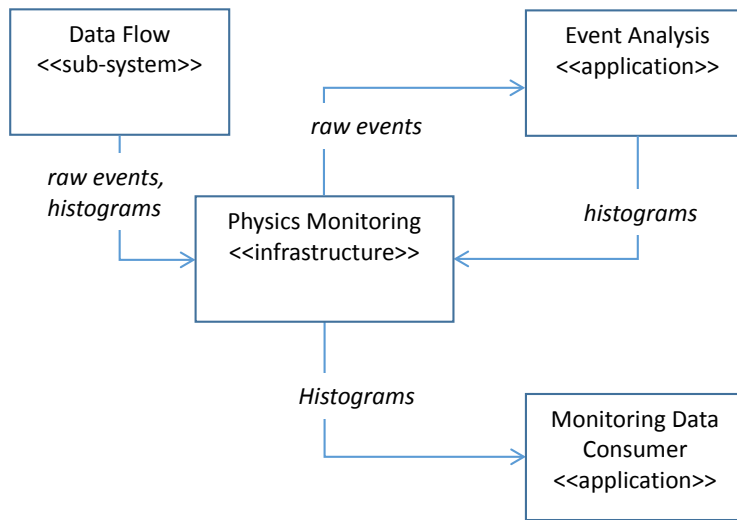


Figure 11.11: *Physics Monitoring interactions with external sub-systems and applications*

In order to simplify development and maintenance of the PM infrastructure it will be broken down into 4 independent sub-systems which deal with different types of monitoring data:

- Event Sampling sub-system: responsible for transportation of physics events from the Dataflow sub-system to event data analysis applications
- Information Sharing: works as a medium for sharing histograms produced by data analysis applications with the other sub-systems and applications
- Information Gathering sub-system: sums up information objects, e.g. histograms, produced by homogeneous applications, e.g. monitoring tasks or EF processing units, in order to get complete statistics for a given system component, e.g the EF farm
- Data Quality Assessment sub-system analyses the monitoring histograms in order to spot any possible malfunction of the detector or the TDAQ system itself

Figure 11.12 shows how these sub-systems will interact with one another in order to provide the required functionality.



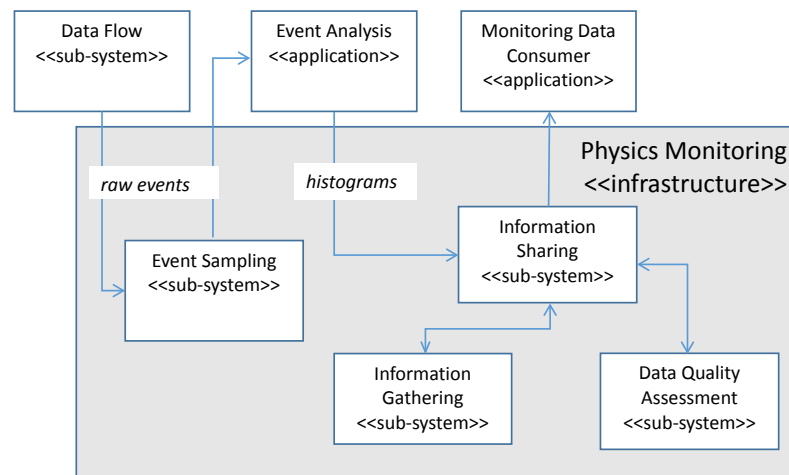


Figure 11.12: *Physics Monitoring sub-systems*

**Event Sampling sub-system** The Event Sampling sub-system provides a way of subscribing to a specific type of raw physics event while it passes through the Dataflow sub-system. Figure 11.13 shows how the Event Sampling sub-system interacts with the other components to accomplish its task. When the Event Sampling system receives a subscribe request it stores selection criteria which are passed as parameters of the request. It will set up a hook function to an appropriate place in the Dataflow system. The Dataflow sub-system calls this function each time a new event passes through so the Event Sampling sub-system can check if the given event matches the provided criteria. If it does match a copy of the event is sent to the subscriber while in the opposite case the event is ignored by the Sampling sub-system. Ideally Event Sampling should also take into account the processing speed of the subscriber in order to send only as many events as the subscriber can afford. If the subscriber does not want to receive any more events it can send an unsubscribe request to cancel the subscription.

**Information Sharing and Gathering sub-system** All histograms produced online by an Event Analysis application or by an EF Processing Unit are registered to the Information Sharing (IS) sub-system, which will implement a subscribe/callback pattern to share them with other interesting applications. When another application or sub-system, e.g. Information Gathering (IG), wants to get some histograms it sends a subscribe request to the Information Sharing system, providing it with a list of interesting histograms. From that moment IS takes responsibility for sending all requested histograms to the reader whenever they are updated by the providers. These interactions are shown in Figure 11.14.

The Data Quality Assessment (DQA) subsystem will also subscribe to merged histograms, which provide complete statistics for a given system component, and will receive them

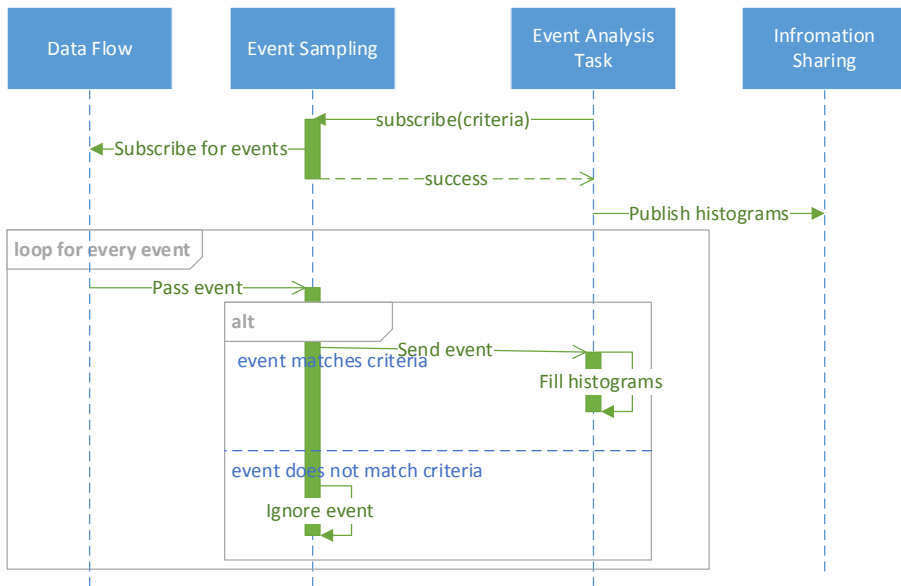


Figure 11.13: *Event Sampling sub-system interactions*

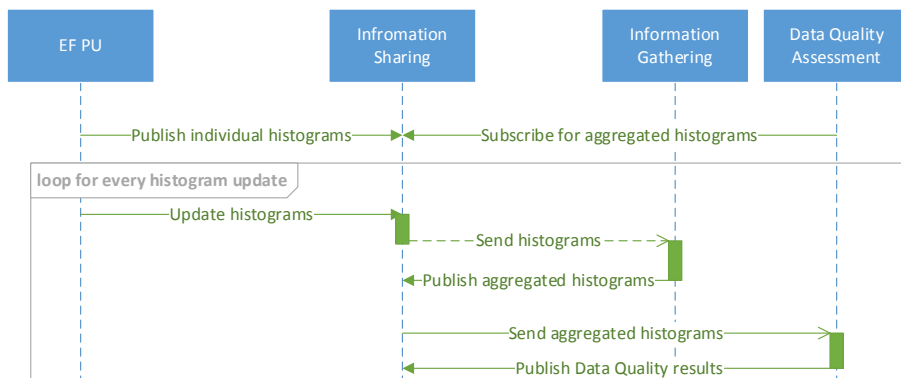


Figure 11.14: *Information Sharing and Gathering sub-systems interactions*

from the IS system whenever they are updated by the IG sub-system. DQA will apply some pre-defined automatic analysis algorithms to those histograms in order to spot anomalies which may point to problems with the current data taking session. If such an anomaly is detected DQA will raise an alarm to notify the shift crew about the issue, which will allow potential problems to be fixed early on.

### 11.7.5 System Administration

As with any large computing installation, system administration is key to successful operation of the ATLAS [TDAQ](#) infrastructure. System administration is an area that evolved substantially in the past few years, with the advent of large commercial data centres, virtualisation, and cloud computing. While the modernisation of individual host management and monitoring tools are areas of continuous development, the Phase-II upgrade could be an opportunity to revise the boundaries between system administration, [DAQ](#), and [DCS](#).

The borders between [DAQ](#), system administration and [DCS](#) that used to be well defined have been eroded by recent developments such as: technology evolution that has made networking pervasive in all aspects of computing; the introduction of embedded computers with fully fledged operating systems in parts of the experiment (which traditionally used lower level electronics devices) and the invention of containers to constrain and define the scope of applications running on servers. Such developments provide an opportunity to rethink the interplay and integration between these three domains.

Although containerisation brings many performance benefits, it also poses potential security risks which would need to be addressed. Because the operating system kernel is shared among containers on the same host, any vulnerabilities in the host would also be present within the containers. To address this, some level of vulnerability scanning and quarantine enforcement will be necessary to ensure the continued security of the container infrastructure [11.37].

Traditionally, system development and operations were two distinct fields. However, recent trends mean that what is known as 'DevOps' is now the paradigm of choice for the integration of complex systems. This requires the involvement of operations and development engineers working together during the entire system lifecycle; from the design, through the development and deployment process, to subsequent production support, and finally to deprecation. With these trends in mind, the approach taken during the last major shutdown (LS1) was to establish a more flexible infrastructure and replace tools which had become outdated and cumbersome over time. The development and continuous improvement of the current toolset is ongoing and will continue to be maintained during future upgrades.

As mentioned also in Section [11.7.2](#), in modern software architectures cluster orchestration services are used to manage large clusters of computing nodes. In this respect, container orchestration exists across the boundary between the application layer (Trigger/[DAQ](#) software) and the system/OS layer. Going forward, the usage of containerised systems may require a review of currently available data storage topologies in order to provide a distributed storage implementation, and improve the scalability of the current storage infrastructure. It would also be necessary to integrate the container infrastructure with the configuration management system that will be in use at the time of deployment. By leveraging some of these more recent technologies, it will be possible to improve computational efficiency

as well as the maintainability of the computing infrastructure, thereby offering a number of significant advantages over the current implementation.

### 11.8 Standalone DAQ Mode & Commissioning

While the integration of the full **DAQ** system for ATLAS data taking is the ultimate goal of all development activity, it is important to maintain support for use cases outside of primary data taking. Critical among these are the use of the **DAQ** architecture for subdetector development and commissioning (away from Point 1) as well as testing and calibration once the final system is in service.

To facilitate such operations to date, the system has been designed in a modular way, with flexibility in terms of hardware platforms. Such a design makes it possible to set up small scale data taking slices featuring subdetector components under testing alongside primary **DAQ** system components, meaning the environment at Point 1 can be accurately emulated. For subdetectors at Point 1 it is also possible to take control of specific slices of the ATLAS detector, alongside related **DAQ** system components, in order to run testing and calibration to prepare and maintain the system ahead of continued data taking.

In order to maintain support for such activities in the future, any upgrade of the **DAQ** system and its components should ensure that it is possible to run logical slices of the system both at Point 1 and at smaller scales elsewhere. Such a design will also make it easier to test and commission **DAQ** system components ahead of their final integration into the ATLAS system for Run 4. The **DAQ** component commissioning process itself will then take place at multiple sites, each with slices of the system according to their development requirements. The modularity of the system components should then make it possible to test combinations of components (e.g. **FELIX** and Data Handler, or Data Handler and Event Builder) in isolation before integrating into larger slices. A fully integrated 'whole system' slice will also be created at CERN as part of the final step before installation in Point 1.

### References

- [11.1] *Versatile Link*, (online). <https://espace.cern.ch/project-versatile-link/public/default.aspx>.
- [11.2] *Versatile Link PLUS*, (online). <https://espace.cern.ch/project-Versatile-Link-Plus/SitePages/Home.aspx>.
- [11.3] *Specifications for the FELIX FULL mode link*, (online). <https://atlas-project-felix.web.cern.ch/atlas-project-felix/user/docs/FullMode.pdf>.

- 
- [11.4] P. Moreira, R. Ballabriga, S. Baron, S. Bonacini, O. Cobanoglu, F. Faccio, T. Fedorov, R. Francisco, P. Gui, P. Hartin, K. Kloukinas, X. Llopart, A. Marchioro, C. Paillard, N. Pinilla, K. Wyllie, and B. Yu, *The GBT Project*, .  
<https://cds.cern.ch/record/1235836>.
- [11.5] P. Farthouat et al., *Report of the preliminary design review of the FELIX for Phase 1*, Tech. Rep. ATU-TC-ER-0048 v.1, EDMS Id 1518157 v.1, CERN, Geneva, Aug, 2015.  
<https://edms.cern.ch/document/1518157>.
- [11.6] *FELIX Firmware Status Report for Phase-I Design Review.*, (online).  
[https://indico.cern.ch/event/676493/contributions/2786317/attachments/1557222/2449510/PR\\_13-11-2017\\_firmware.pdf](https://indico.cern.ch/event/676493/contributions/2786317/attachments/1557222/2449510/PR_13-11-2017_firmware.pdf).
- [11.7] ATLAS FELIX Group Collaboration, *FELIX Phase-I Technical Specification and Implementation*, Nov, 2017. [https://svnweb.cern.ch/cern/wsvn/FELIX/trunk/Documents/Phase1\\_FDR/FelixPhase1FDR.pdf](https://svnweb.cern.ch/cern/wsvn/FELIX/trunk/Documents/Phase1_FDR/FelixPhase1FDR.pdf).
- [11.8] *CEPH Filesystem*, (online). <http://docs.ceph.com/docs/master/cephfs/>.
- [11.9] *GlusterFS Documentation*, (online). <http://docs.gluster.org/en/latest/>.
- [11.10] E. Kakoulli and H. Herodotou, *OctopusFS: A Distributed File System with Tiered Storage Management*, in *Proceedings of the 2017 ACM International Conference on Management of Data, SIGMOD '17*, pp. 65–78. ACM, New York, NY, USA, 2017.  
<http://doi.acm.org/10.1145/3035918.3064023>.
- [11.11] *Software Defined Block Storage – ScaleIO*, (online).  
<https://www.dellemc.com/en-us/storage/scaleio/index.htm>.
- [11.12] *Apache Hadoop*, (online). <http://hadoop.apache.org/>.
- [11.13] K. Shvachko, H. Kuang, S. Radia, and R. Chansler, *The Hadoop Distributed File System*, in *2010 IEEE 26th Symposium on Mass Storage Systems and Technologies (MSST)*, pp. 1–10. May, 2010.
- [11.14] F. Schmuck and R. Haskin, *GPFS: A Shared-disk File System for Large Computing Clusters*, in *Proceedings of the 1st USENIX Conference on File and Storage Technologies, FAST'02*, pp. 16–16. USENIX Association, Berkeley, CA, USA, 2002.  
<http://dl.acm.org/citation.cfm?id=1973333.1973349>.
- [11.15] ATLAS Collaboration, *Technical Design Report for the Phase-I Upgrade of the ATLAS TDAQ System*, Tech. Rep. CERN-LHCC-2013-018. ATLAS-TDR-023, CERN, Geneva, Sep, 2013. <https://cds.cern.ch/record/1602235>.
- [11.16] *The 2016 Ethernet Roadmap*, (online).  
<http://ethernetalliance.org/roadmap/>.
- [11.17] *Infiniband Roadmap*, (online). [http://www.infinibandta.org/content/pages.php?pg=technology\\_overview](http://www.infinibandta.org/content/pages.php?pg=technology_overview).

## 11.8 Standalone DAQ Mode & Commissioning

---

- [11.18] *PCI Express Specifications and ECRs for Review*, (online).  
<https://pcisig.com/specifications/review-zone>.
- [11.19] *Brocade Data Center Fabric Architectures*, <http://www.brocade.com/en/backend-content/pdf-page.html?/content/dam/common/documents/content-types/solution-design-guide/brocade-dc-fabric-architectures-sdg.pdf>.
- [11.20] *Active Optical Cables VS Optical Transceivers*, (online).  
<http://www.mellanox.com/blog/2016/12/aocs-active-optical-cables-why-are-2-transceivers-and-fibers-priced-less-than-one-connectorized-transceiver/>.
- [11.21] *CMake Project Page*, (online). <https://cmake.org>.
- [11.22] *The Athena Framework*, (online).  
<https://atlassoftwaredocs.web.cern.ch/athena/athena-intro/>.
- [11.23] *Kubernetes: Production-Grade Container Orchestration*, (online).  
<https://kubernetes.io/>.
- [11.24] *An update on container support on Google Cloud Platform*, (online).  
<https://cloudplatform.googleblog.com/2014/06/an-update-on-container-support-on-google-cloud-platform.html>.
- [11.25] *The Kubernetes Open Source Project on Open Hub*, (online).  
<https://www.openhub.net/p/kubernetes>.
- [11.26] *Docker: the world's leading software container platform*, (online).  
<https://www.docker.com/>.
- [11.27] *CERN Cloud Infrastructure*, (online). <https://openstack.cern.ch/>.
- [11.28] *SLC6: Scientific Linux CERN 6*, (online).  
<http://linux.web.cern.ch/linux/scientific6/>.
- [11.29] *CernVM File System*, (online).  
<https://cernvm.cern.ch/portal/filesystem>.
- [11.30] *Scalability updates in Kubernetes 1.6: 5000 node and 150000 pod clusters*, (online).  
<http://blog.kubernetes.io/2017/03/scalability-updates-in-kubernetes-1.6.html>.
- [11.31] *Google Cloud Platform: Machine Types*, (online).  
<https://cloud.google.com/compute/docs/machine-types>.
- [11.32] *Icinga 2*, (online). <https://www.icinga.com/products/icinga-2/>.
- [11.33] *Ganglia Monitoring System*, (online). <http://ganglia.sourceforge.net/>.

- [11.34] *IPMI: Intelligent Platform Management Interface Specification, Second Generation V2.0, Revision 1.1*, (online). <https://www.intel.com/content/dam/www/public/us/en/documents/product-briefs/ipmi-second-gen-interface-spec-v2-rev1-1.pdf>.
- [11.35] *A Simple Network Management Protocol (SNMP): RFC 1098*, (online). <https://tools.ietf.org/html/rfc1098>.
- [11.36] *Deploying Prometheus at DigitalOcean*, (online). <https://promcon.io/2016-berlin/talks/deploying-prometheus-at-digitalocean/>.
- [11.37] N. Bila, P. Dettori, A. Kanso, Y. Watanabe, and A. Youssef, *Leveraging the Serverless Architecture for Securing Linux Containers*, in *2017 IEEE 37th International Conference on Distributed Computing Systems Workshops (ICDCSW)*, pp. 401–404. June, 2017.





## 12 Event Filter

The [Event Filter \(EF\)](#) is responsible for executing sophisticated event selection algorithms at high rate, in a massively parallelised form, within the hardware and software infrastructure discussed in [Chapter 11](#). The selections taking place in the [EF](#) reflect the physics goals of the collaboration, and are the means by which it is possible to deliver a rich experimental dataset within the agreed rate and bandwidth budget. Much of the rejection power of the [EF](#) will be based on hardware-based tracking, discussed in more detail in [Chapter 13](#). As such, this chapter will focus on the software component of the [EF](#), as well as presenting how all information sources are combined to reach the final [EF](#) decision. [Section 11.7.2](#) describes the farm management aspects of the [EF](#).

The challenge for the [EF](#) in Phase-II is to support the increased event input rate, while mitigating the rise in algorithm execution times resulting from the increased level of pile-up at higher luminosity. The physics motivation and expected physics performance of the [EF](#) trigger algorithms are discussed in [Chapters 2](#) and [6](#), respectively. In particular, [Table 6.4](#) is used as the basis for the estimation of the required [CPU](#) compute power.

Upgrades are required both to the farm hardware, to provide the needed compute power, and to the selection software to provide the required selection power and algorithm speed. In the present (Run 1-Run 3) trigger architecture, the track-based selections that provide much of the rejection power can only be performed in the [EF](#). For Run 3 we expect a fully commissioned [Fast Tracker \(FTK\)](#) system, which will augment the current software tracking with hardware-based full-event tracking. This will allow the [EF](#)-based event selection to move even closer to the offline one. In Phase-II, hardware-based track selections will also be possible inside [RoIs](#). In order to maintain rejection, suppress pileup and keep efficiency high, the [EF](#) will need to make use of more [CPU](#)-intensive offline-type selection techniques, including a greater use of global (full-event) tracking.

The most affordable computing platform (commodity server PCs) is developing towards systems hosting multi-core [CPUs](#) with an increasing core-count and heterogeneous hardware architectures. The exact architecture is however difficult to predict. As an example of the quickly evolving market, in November 2017 Intel made the following two announcements: the 8<sup>th</sup> generation Intel Core processors will come with a “custom-to-Intel third-party discrete graphics chip from AMD’s Radeon Technologies Group” [[12.1](#)] all in the same processor package; and the next generation Intel Xeon Phi is cancelled in favour of a “new platform” [[12.2](#)] but without giving more details on this new platform. Considering other recent events, it is likely that we will see [CPUs](#) with [GPGPUs](#) and/or [FPGAs](#) all integrated

into the same package allowing high-speed memory transfers between them. Therefore, the upgraded EF software should allow for both parallel algorithm execution as well as exploitation of internal parallelism of those which are the most costly in terms of CPU. The optimal degree of parallelism (processing multiple events at the same time, parallel regional reconstruction within the same event, or parallelism within individual algorithms) will be found by balancing the potential benefits in throughput and the effort needed to modify and maintain the code.

The EF in Phase-II will be required to interface to the following subsystems (see Fig. 12.1):

- The hardware track trigger system, to request tracks from regions via **rHTT** (tracks with  $p_T > 2$  GeV) or the full **ITk** via **gHTT** ( $p_T > 1$  GeV) both covering the region  $|\eta| < 4$ .
- The Dataflow system, which will distribute event processing tasks, provide input data to and collect the output from the EF for each event.
- Online software will provide monitoring services and control of EF applications.

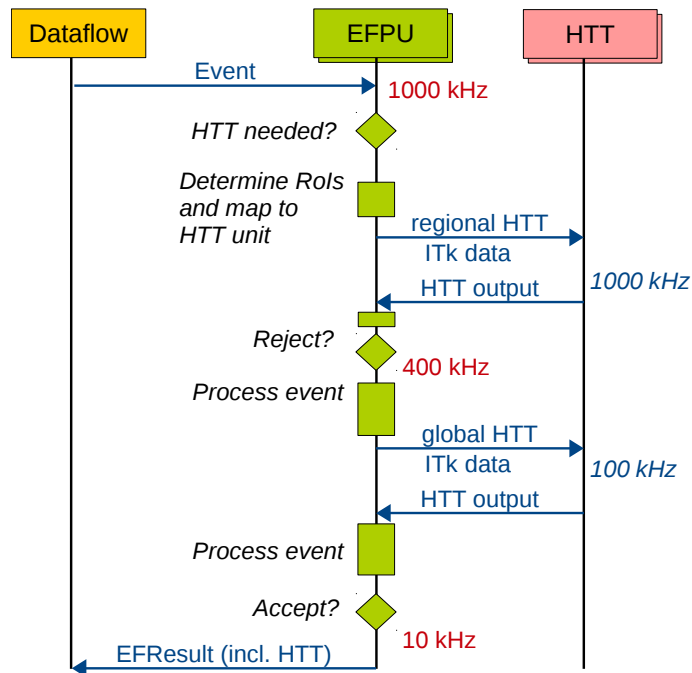


Figure 12.1: Diagram showing the interaction between the EF processing unit (EFP), Dataflow system and the hardware tracking (HTT). After receiving the event data from the Dataflow system, the EFP decides, based on the Level-0 decision, if regional hardware tracking is needed for this event. If yes, requests are sent to the relevant rHTT units along with the necessary ITk data. After receiving the HTT result the EFP decides if the event should be rejected or continues with event processing including a possible second request for full event hardware tracking via the gHTT based on the Level-0 decision or additional EF-based selections.

The **EF** processing is expected to occur promptly, i.e. with a minimum delay after the L0 accept decision. The large buffer in the **Storage Handler** will allow for processing timeout settings which are less strict than those used in the Run 2 and Run 3 systems. It is also conceivable that processing could continue beyond the end of a data-taking run. However, delayed processing that makes full use of the **LHC** inter-fill periods is currently not seen as desirable from an operational point of view. The option of an **EF** calibration loop (similar to the prompt calibration loop deployed at Tier-0) that would derive certain calibrations before the **EF** processing starts has been studied. While there is no clear use-case for calorimeter or muons, a full ITk alignment could be derived within one hour of data-taking. However, assuming the new ITk detector is at least as stable as our current detector it is not clear that this would result in significant improvements for the **EF** selection. In any case, should a use-case emerge, the **Storage Handler** could serve as a buffer of the L0 accepted data until calibrations have been derived.

## 12.1 Event Filter Hardware

Current estimates of the evolution of event processing times indicate the need for  $4.5^{+2.7}_{-0.7}$  million HEP-SPEC06<sup>1</sup> (MHS06) to handle a Level-0 rate of 1 MHz with an initial reduction to 400 kHz achieved entirely by using information from Level-0 and **rHTT** as shown in Table 6.4. The **CPU** estimate is based on an extrapolation of current (Run 2) **CPU** usage, taking into account scaling with pile-up, **CPU** time reduction due to the use of hardware tracking, and other software improvements. More details on the estimation, the uncertainties and the pile-up dependence are given in Section 12.4.

All the commodity computing power must be accommodated within a fixed rack-space located in the surface computing infrastructure. This rack-space is shared with other components of the data-acquisition system, such as storage and networking. An extrapolation based on the evolution of compute-power in the ATLAS **TDAQ** farm over the past ten years (see Fig. 12.2a) results in an estimated computing capacity of 1.5 kMHS06 per dual-socket server or approximately 3000 motherboards on the time-scale of Phase-II. Assuming the current format of four motherboards in a 2U server chassis, these can be accommodated in 38 racks of the type that will be available in level one of the SDX building. The scale of the system is summarised in Table 12.1.

Table 12.1: *Summary of **EF** farm size estimates, based on projections of compute capacity requirements and compute power of servers from current data, as described in the text.*

Compute capacity required	4.5 MHS06
Equivalent dual-socket servers	3000
Racks	38

<sup>1</sup> <http://w3.hepik.org/benchmarking.html>

## 12.2 Selection software

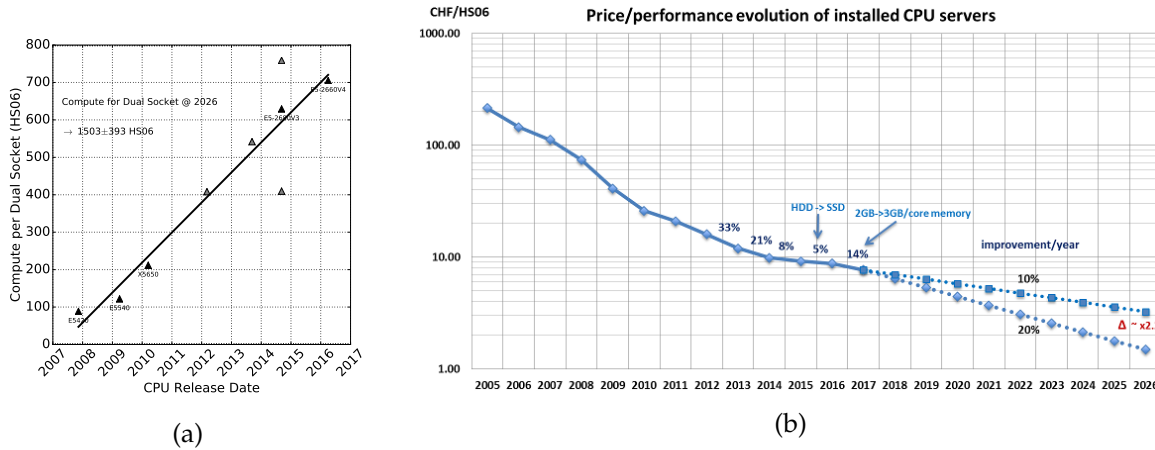


Figure 12.2: (a) Evolution of the compute power per dual socket motherboard installed in the ATLAS TDAQ farm over the last ten years. (b) Observed and expected price/performance evolution provided by CERN IT with the reasonable (1.4 CHF/HS06) and pessimistic (3.3 CHF/HS06) scenario for prices in 2026 [12.3].

Predicting how the cost of commodity computing technology will evolve in the next decade is very difficult. The cost on the other side is expected to be driven mainly by financial and market developments rather than technology. Furthermore there is a possibility that a disruptive technology has a large unforeseen impact. We assume a cost of 2 CHF/HS06 by 2026, consistent with the *reasonable* (1.4 CHF/HS06) and *pessimistic* (3.3 CHF/HS06) scenarios provided by CERN IT (see Fig. 12.2b). Based on this, we expect a cost of 3000 CHF per motherboard.

It is envisaged that the EF farm could be a rather heterogeneous system, possibly containing different classes of hardware such as GPGPUs, FPGAs and commodity servers. Advances in these areas will be closely monitored (see Section 12.3), but the current baseline assumption is that CPUs will provide the required compute density on the time-scale of Phase-II. However, since the technology decisions shaping the EF infrastructure will be taken as late as possible, it is important to establish interfaces that will allow for operation of a heterogeneous system.

## 12.2 Selection software

The selection software upgrades fall into two parts: further evolution of the framework and development of selection algorithms to meet the Phase-II requirements.

The present software framework is undergoing a major upgrade to provide the new functionality needed for the start of Run 3. The new framework, AthenaMT [12.4][12.5], is being implemented as a common trigger and offline computing framework. Key features are

---

built-in: multi-threading support, provision for seamless integration of offline algorithms and infrastructure to support external accelerators. The AthenaMT framework will be able to run existing [EF](#) and offline algorithms with limited changes, but additional changes will be needed to fully exploit the potential of the new framework. Most of the changes in the selection software are expected to be done in time for Run 3. Further changes for Run 4 will be needed to adapt the system to the evolved [DAQ](#) architecture, i.e. with regards the [Storage Handler](#) and hardware tracking interfaces. In addition, new trigger hardware (e.g. the [Global Trigger](#)) will have to be included into the trigger configuration database. In case [GPU](#) or [FPGA](#) accelerators come to be used, the necessary framework services to offload the computing load to these devices will have to be developed unless the manufacturers provide compilers that automatically make use of these accelerators.

Significant effort will be required to upgrade the selection software to provide the required rejection in the [EF](#) within [CPU](#) resource constraints. Experience from Run 1 and Run 2 has shown that the rates of certain triggers, especially multi-jet triggers and triggers based on  $E_T^{\text{miss}}$ , rise rapidly with increasing pile-up. It is expected that an initial reduction from the Level-0 rate of 1 MHz to 400 kHz can be achieved by the use of hardware-reconstructed tracks and other information from the [Global Trigger](#). After this initial 'fast' rejection, the remaining required [EF](#) rejection power will be achieved by importing techniques currently used offline to provide selections that are robust against the effects of pile-up. The latter often rely on tracking information to correct the energy of a calorimeter cluster. More details on the tracking use cases can be found in Section [6.12](#).

The tracking trigger is subdivided into *fast tracking* and *precision tracking* stages. The fast tracking consists of trigger-specific pattern recognition, whereas the precision stage relies heavily on offline tracking algorithms. The tracking algorithms are typically configured to run within an [RoI](#) identified by L0. To reduce [CPU](#) usage even further, the offline track-finding is seeded by tracks and space-points identified by the fast tracking stage. The biggest change is expected to the inner detector tracking software, as it will have to be adapted to the new [ITk](#) detector and the hardware-reconstructed tracks provided by [rHTT](#) and [gHTT](#). The former has already been done for the offline tracking software and is being used routinely for upgrade studies of the [ITk](#) detector. Since the trigger tracking software is largely based on the offline tracking software, any additional trigger-specific changes with regards to [ITk](#) are expected to be of limited scope. A lot of experience in the use of hardware-reconstructed tracks is expected to be gained once the [FTK](#) system is fully operational by the end of Run 2.

ATLAS will rely on simulation of the upgraded trigger for development of [EF](#) software and performance studies prior to the upgrade and processing of fully-simulated data for analysis of high-luminosity data. The same [EF](#) selection software can be run on simulated data in the ATLAS production system, as is the case already with the current [HLT](#) selection software. This is straightforward thanks to the common framework and large overlap of services, algorithms and event data classes with reconstruction software. Simulation of the

detectors and hardware triggers will provide the input data to the EF. The ATLAS simulation software already includes hardware-based tracking for FTK. It will be updated to include full simulation of the HTT system, which is currently developed standalone. A fast simulation of the HTT will also be included, which will reduce processing time significantly.

The current muon trigger reconstruction [12.6] is composed of a fast muon and inner detector track reconstruction, followed by a combination step and final precision reconstruction in both the inner detector and muon systems. Both the trigger-specific and offline reconstruction software will have to be adapted to include and leverage the information provided by the new MDT Level-0 trigger as well as hardware-reconstructed tracks, both of which should result in a significant CPU usage reduction. It is expected that at least a factor two improvement in the muon reconstruction time is necessary on the timescale of Run 4.

The first stage in the calorimeter reconstruction involves unpacking the data from the calorimeters. The unpacking can be done in two different ways: either by unpacking only the data from within the RoIs identified at L0 or by unpacking the data from the full calorimeter. The RoI-based approach is used for well-separated objects (e.g. electron, photon, muon, tau), whereas the full calorimeter reconstruction is used for jets and global event quantities (e.g.  $E_T^{\text{miss}}$ ). In both cases the raw unpacked data is then converted into a collection of cells. Two different clustering algorithms are used to reconstruct the clusters of energy deposited in the calorimeter, the sliding-window and the topo-clustering algorithms [12.7]. While the latter is significantly slower it provides performance closer to the offline reconstruction. The calorimeter software will remain largely unchanged, but with the aim to further harmonise the trigger and offline reconstruction. The focus will have to remain on improving the performance of the software, i.e. since some of the fast rejection algorithms will move to the calorimeter hardware trigger. Possible new developments include the use of hardware-based topo-clusters from the Global Trigger.

### 12.3 Study of GPGPU usage in the EF

A demonstrator has been developed to evaluate the potential benefit of offloading specific trigger reconstruction tasks to GPGPUs. Selected trigger reconstruction algorithms from the current (C++) software framework (Athena) have been re-written as Compute Unified Device Architecture (CUDA) modules to be executed on a NVIDIA GPGPU. Within the CPU code, the selected algorithms are replaced by code that converts the required input data to a format suitable for the GPGPU and makes a request, including the data payload, to a CPU-based server application, known as the Accelerator Process Extension (APE) server. The server forwards the work to a GPGPU. On completion, the server returns the output to the CPU client that then performs a conversion of the output data back to the Athena data-format. The APE server handles requests from multiple clients and can be configured to

send work to one or more **GPGPU** (or other resource). The increase in throughput (events processed per second) that can be achieved when **GPGPU**s are added to a **CPU**-only system depends on the fraction (by execution time) of the **CPU** workload that is exported to the **GPGPU**. For example, throughput could be doubled if it were possible to export 50% of the **CPU** work-load to **GPGPU**. The cost-effectiveness of this solution depends on the relative costs of the **CPU** and **GPGPU** hardware and the relative algorithm execution times on the two platforms, which determines the relative number of **CPU** to **GPGPU** needed to perform the task.

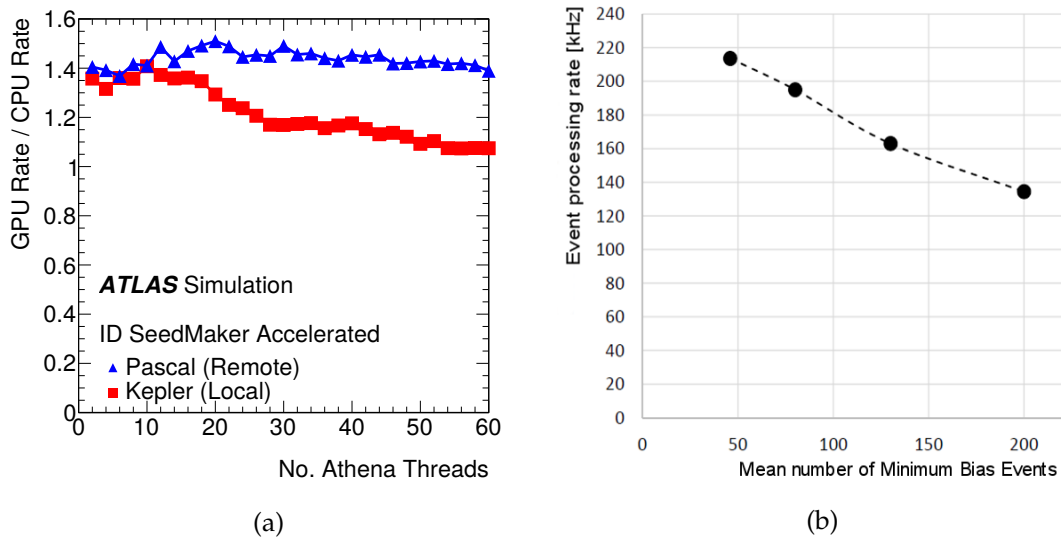


Figure 12.3: (a) Throughput increase as a function of the number of Athena worker threads when executing only ID reconstruction algorithms. (b) Event processing rate for **GPGPU**-based pattern matching as a function of the number of pile-up events.

Processing modules for **GPGPU** have been implemented for **CPU**-intensive parts of the inner detector tracking (data preparation and track-seeding) [12.8], calorimeter topological clustering [12.9], cluster-splitting, jet reconstruction and a Hough transform based algorithm for muon tracking [12.10]. Throughput measurements have been made for a system comprising two 14-core Intel Xeon E5-2695 **CPU** (2.3 GHz clock-speed) and a system with the same **CPU** plus a **GPGPU** serving as accelerator. For the **GPGPU** two scenarios were tested:

- i) up to four Nvidia GK210GL Kepler architecture **GPGPU** (in two K80 PCI cards) on the same **PCIe** back-plane as the **CPU**,
- ii) a GTX1080 **PCIe** card with a (newer) Pascal architecture **GPGPU** (GP104) in a separate unit with a local **CPU** acting as a server and connected to the client node via a 10 GbE network.

For the inner detector track-seeding algorithm a speed-up factor of 28 (5.8) is obtained on the Pascal (Kepler) **GPGPU** relative to execution on **CPU**. Overheads for data conversion



### 12.3 Study of GPGPU usage in the EF

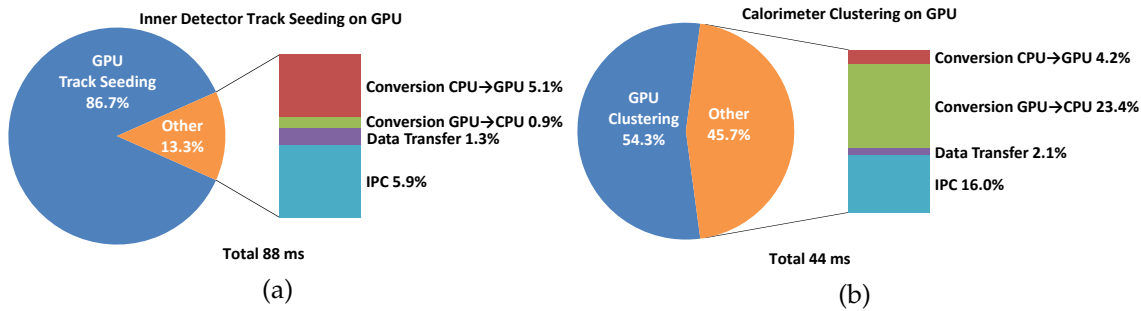


Figure 12.4: Breakdown of the time per event (measured on the Kepler system) for (a) inner detector track-seeding and (b) calorimeter clustering offloaded to a GPGPU for the kernels running on the GPGPU and the overhead associated with offloading the work (other). The overhead comprises the time to convert data-structures between CPU and GPGPU data-formats, the data transfer time between CPU and GPGPU and the Interprocess communication (IPC) time that accounts for the transfer of data between the Athena processes and the process handling communication with the GPGPU.

and inter-process communication reduce this to an effective speed-up of 15 (5) for the Pascal (Kepler) GPGPU cards respectively. Fig. 12.3a shows the throughput increase factor (defined as the ratio of the event rate with GPGPU to the CPU-only event rate) as a function of the number of Athena clients for full-event ID track reconstruction accelerated by a Pascal-architecture and a Kepler-architecture GPGPU. The input was a simulated Phase-I tt dataset converted to raw detector output format (bytestream). An average of 46 minimum bias events per simulated collision were superimposed. In the CPU-only case, the track-seeding algorithm takes 28% of the total event processing time, which limits the maximum possible throughput increase to a factor of about 1.4. The Kepler GPGPU shows evidence of saturation for 20 or more clients, while no saturation is seen for the Pascal GPGPU with 60 clients due to the shorter algorithm execution time on this GPGPU. The speed-up factor obtained for the calorimeter algorithm is 3.6 on the Kepler GPGPU, but in this case the shorter algorithm execution time on the GPGPU is completely offset by the data-conversion and inter-process communication overheads, meaning no throughput increase was obtained. The latter can be observed in Figs. 12.4a and 12.4b showing the breakdown of the time per event for inner detector track-seeding and calorimeter clustering, respectively. This illustrates the importance of implementing a suitable event data format in the offline and trigger code to avoid expensive data-format conversions. More details can be found in [12.11].

Based on the measurements made with the demonstrator, it is estimated that using the tested hardware it would cost approximately the same to increase the farm throughput by adding GPGPU or CPU. For example, it is estimated that two hypothetical racks, one with 120 CPU and another one with 80 CPU plus 40 GPGPU, would have approximately the same cost, throughput and cooling requirements (assuming that one GPGPU serves two CPU and gives a 50% increase of the CPU throughput and that a unit with 4 GPGPU costs about half of a unit with 8 CPU). However, the cost-effectiveness of adding GPGPU



---

to the EF depends on the evolution of CPU and GPGPU in terms of price, performance and packaging. Implementing trigger algorithms to run on GPGPU requires redesigning and parallelising the code to efficiently use the GPGPU hardware. For the demonstrator, this task took approximately 0.5 staff years per algorithm and relied on expert knowledge and experience to implement performant code.

The Inner Detector (ID) algorithm discussed above employs a combinatorial approach to inner detector pattern recognition. A second GPGPU demonstrator has been developed to assess an alternative pattern-matching approach similar to the method employed by rHTT and gHTT. A high-end gaming card (Nvidia Titan X) with a GP102 Pascal GPGPU chip has been used for these studies. The pattern data were stored in the GPGPU registers and event data were read from the GPGPU device memory for each new event. For this setup, the number of patterns per GPGPU is limited by the GPGPU register size to about 0.57 million. The patterns are processed by 56 blocks of threads (the GP102 has 28 multiprocessors that can each run two blocks simultaneously); each block has 1024 threads and each thread handles 10 patterns. Measurements have been made with L1Track patterns with eight layers, including one Pixel layer, using a simulated Phase-II dataset consisting of single muon events with the addition of a number of minimum bias events following a Poisson distribution about a mean that was varied for successive measurements. Fig. 12.3b shows the event processing rate as a function of the mean number of superimposed minimum bias events. Using GPGPU-based pattern matching, an event processing rate of about 140 kHz can be achieved for a pile-up of 200 minimum bias events. Based on a total bank size of 6 billion patterns (as for gHTT), a total of 10 thousand GPGPUs would be required for the full system.

In summary, GPGPUs have been studied as a potential commodity hardware accelerator to which suitable compute-intensive processing tasks can be offloaded from the main CPU. A study for the Phase-I upgrade based on current technology has concluded that GPGPUs and CPUs could be used to increase throughput of the EF at roughly the same cost and with similar power and space requirements. The relative cost-effectiveness of these technologies for the Phase-II EF depends on the relative evolution of CPU and GPGPU in terms of price, performance and packaging. The decision on the commodity compute hardware for the EF farm will be taken nearer the time of purchase following an evaluation of hardware accelerators. The above work should be taken as an indication that the software can be successfully adapted to other architectures in order to perform a full cost/benefit evaluation.

## 12.4 Model for CPU estimation

To estimate the EF CPU requirements for Phase-II for a given pile-up value  $\mu$  the following model is used:

## 12.4 Model for CPU estimation

---

1. The **HLT CPU** usage of a data-taking run at  $\mu = 43$  is measured and broken down into its individual components: ID Data-preparation (11%), ID Tracking (41%), Calorimeter (11%), Muon (26%) reconstruction and object reconstruction for electron, muon, jet,  $E_T^{\text{miss}}$ , etc. (11%).
2. The **Inner Detector (ID)** tracking is further split into the trigger-specific *fast tracking*, *precision tracking* and TRT tracking. The latter is subtracted for all further projections as this detector no longer exists in Phase-II.
3. For each object type (electron, muon, jet, etc.) a rate scale factor is derived based on the expected trigger rates after the regional tracking cuts shown in Table 6.4.
4. The **CPU** times for the individual components are scaled with the relevant rate and pile-up (see below) scale factors and the total **CPU** requirement in units of HS06 is calculated.

The following additional assumptions are made for the **CPU** projections:

- (A) The **CPU** time for software tracking for ITk scales with  $0.5 \times f_\mu$  compared to today's inner detector tracking where  $f_\mu$  is the pile-up ratio (e.g. ITk tracking at  $\mu = 200$  is 5 times slower than today's tracking at  $\mu = 20$ ). This advantageous scaling results from the improved layout and design of the ITk detector for high-pileup conditions [12.12].
- (B) The inner detector data preparation and muon spectrometer tracking scales with  $f_\mu$ , i.e. linear with pile-up. This assumption is based on past experience taking into account code improvements as the pile-up increases.
- (C) The calorimeter, jet and  $E_T^{\text{miss}}$  reconstruction times are constant with respect to pile-up when the appropriate pile-up noise corrections are applied. This has been verified during the current data-taking run.
- (D) The **CPU** time required by today's *fast tracking* will be reduced by 80% thanks to the use of tracks from **HTT**. The remaining time is required for re-fitting the tracks obtained from **HTT**. This is based on measurements with the Run 2 **FTK** system.
- (E) Tracking for hadronic triggers will be provided by **gHTT**. Consequently the **CPU** times for vertexing, beamspot measurement and other tracking-based pre-selections (e.g. the first stage of the two-stage tau tracking [12.6]) are reduced to zero.
- (F) The unpacking and conversion time for hardware-reconstructed tracks received from **HTT** is 0.05 ms per track, which is extrapolated from today's 0.1 ms per track in the **FTK** system. We assume 2000 tracks at 100 kHz from **gHTT** and 50 tracks at 1 MHz from **rHTT**.
- (G) The precision muon reconstruction is a known bottle-neck in today's **HLT** reconstruction software. We assume that a factor two speedup can be achieved on the timescale of Phase-II as described in Section 12.2.
- (H) The jet and  $E_T^{\text{miss}}$  reconstruction times are increased by 50% to take into account the additional use of (hardware) tracks in their calculation.
- (I) Topo-clusters for the full event (in addition to the **RoI**-based topo-clusters used in the electron and tau reconstruction) are created at a rate of 100 kHz based on the rate of hadronic triggers.

Given the above model, the CPU estimates have some rather large uncertainties assigned to them. An attempt to quantify these has been made and is summarised in Table 12.2. Since the correlations of these uncertainties are unknown, the total uncertainty is given for the cases of fully (un)correlated uncertainties. For simplicity we quote in the following the average between these two extreme values as the uncertainty on the CPU estimate.

Table 12.2: *Uncertainties on the CPU estimates in percent. The total uncertainty is given for the cases of fully (un)correlated uncertainties. Since the correlations are unknown we choose the average between these two extreme values as our model uncertainty.*

Trigger and object rates after rHTT rejection are larger than expected	+20%
The CPU time for software tracking for ITk (A) scales with $1.0 \times f_\mu$	+22%
The speedup in the <i>fast tracking</i> (D) due to the use of hardware tracks is only 50% instead of 80%	+11%
Hardware-reconstructed tracks are not re-used in the software tracking for electron and muon signatures (apart from the initial rejection)	+8%
$\pm 50\%$ uncertainty on the unpacking times for HTT tracks (F)	$\pm 3\%$
$\pm 50\%$ uncertainty on the speedup of the precision muon reconstruction (G)	$\pm 12\%$
Topo-clusters are created for 300 kHz of events	+9%
Hardware topo-clusters from the Global Trigger can be re-used in software	-4%
Total (correlated)	+83% -20%
Total (uncorrelated)	+36% -13%
Total ((corr+uncorr)/2)	+59% -17%

Table 12.3 summarises the CPU requirements in million HEP-SPEC06 (MHS06) 200 pile-up interactions, split into the different reconstruction domains.

Table 12.3: *EF CPU requirements in million HEP-SPEC06 (MHS06) for 200 pile-up interactions for the different reconstructions domains.*

CPU [MHS06]	$\mu = 200$
HTT Unpacking	0.30
ID Dataprep.	0.91
ID Tracking	0.99
Muon tracking	1.37
Calo	0.32
Egamma/Tau	0.08
Jet/MET	0.50
Total	$4.47^{+2.7}_{-0.7}$

## 12.4 Model for CPU estimation

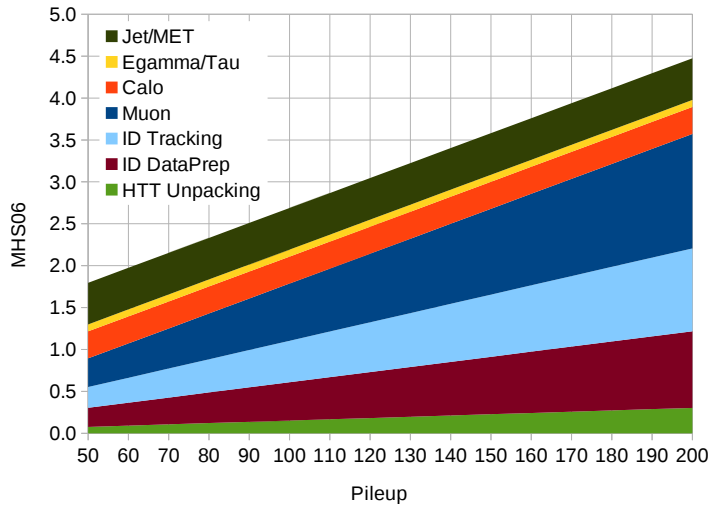


Figure 12.5: *EF CPU extrapolation vs pile-up in million HS06 for the different components of the EF reconstruction software. Only the effects of pileup are considered but trigger rates remain the same.*

Based on the above model an attempt can be made to estimate the CPU requirements without rHTT. In case no rHTT is available, the input rate to the EF increases from 400 kHz to 1 MHz and all regional tracking needs to be done in software. In order to estimate the CPU required in this scenario the same model as explained above is used with the following changes. For each object type (electron, muon, tau, etc.) the L0 rate shown in Table 6.4 is used to determine the rate of regional tracking (e.g. this results in a factor five increase for single electrons). In addition, assumption (D) above no longer holds and the full time for software tracking needs to be accounted for (resulting in a  $\approx 70\%$  increase of the ID data preparation and tracking CPU requirements shown in Table 12.3). On the other side, the time for converting rHTT tracks (F) can be neglected. Putting all this together would result in an additional CPU requirement of 13.4 MHS06 for the EF farm.

Without gHTT all full-event tracking will have to be done in software. From offline tracking studies with the ITk layout it is estimated that 270 HS06 per event are required to perform full-event tracking at  $\mu = 200$  and with a  $p_T$ -cut of 900 MeV [12.12]. For 100 kHz of events this would result in an additional CPU requirement of 27 MHS06.

For both cases, the calculation is based on the currently available ITk tracking software used in the preparation of its TDR [12.12]. Further improvements in the tracking software are to be expected over the next decade, which could reduce the cost of software tracking. On the other side, the large uncertainties in the market evolution for commodity computing make it difficult to do a direct cost comparison between a software-based and a custom hardware tracking solution with the latter being less susceptible to market driven cost changes.

---

## References

- [12.1] *New Intel Core Processor Combines High-Performance CPU with Custom Discrete Graphics from AMD to Enable Sleeker, Thinner Devices*, (online).  
<https://newsroom.intel.com/editorials/new-intel-core-processor-combine-high-performance-cpu-discrete>.
- [12.2] *Unleashing High-Performance Computing Today and Tomorrow*, (online).  
<https://itpeernetwork.intel.com/unleashing-high-performance-computing>.
- [12.3] H. Meinhard and B. Panzer-Steindl, *Technology Evolution, WLCG Workshop 2016*,  
<https://indico.cern.ch/event/555063/contributions/2285842/>.
- [12.4] J. Baines, T. Bold, P. Calafiura, S. Kama, C. Leggett, D. Malon, G. Stewart, and B. Wynne, *ATLAS Future Framework Requirements Group Report*, Tech. Rep. ATL-COM-SOFT-2014-048, CERN, Geneva, Dec, 2014.  
<https://cds.cern.ch/record/1974156>.
- [12.5] J. Baines, T. Bold, D. Oliveira Damazio, S. Martin-Haugh, J. Masik, T. Martin, F. Pastore, H. J. Stelzer, W. Panduro Vazquez, R. White, W. Wiedenmann, F. Winklmeier, B. Wynne, C. Bernius, M. Hodgkinson, C. Leggett, E. Moyse, I. Shapoval, S. Snyder, and G. Stewart, *Design of the High Level Trigger in AthenaMT*, Tech. Rep. ATL-COM-DAQ-2017-047, CERN, Geneva, Jun, 2017.  
<https://cds.cern.ch/record/2268736>.
- [12.6] ATLAS Collaboration, *Performance of the ATLAS Trigger System in 2015*, *Eur. Phys. J. C* **77** (2017) 317, [arXiv:1611.09661](https://arxiv.org/abs/1611.09661) [hep-ex].
- [12.7] ATLAS Collaboration, *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, [arXiv:1603.02934](https://arxiv.org/abs/1603.02934) [hep-ex].
- [12.8] D. Emelianov and J. Howard, *GPU-Based Tracking Algorithms for the ATLAS High-Level Trigger*, *Journal of Physics: Conference Series* **396** (2012) no. 1, 012018.  
<http://stacks.iop.org/1742-6596/396/i=1/a=012018>.
- [12.9] ATLAS Collaboration, P. Loch, S. Menke, P.-A. Delsart, C. J. Young, Z. Marshall, and B. Axen, *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run I*, <https://cds.cern.ch/record/1967028>.
- [12.10] R. Duda and P. Hart, *Use of the Hough transformation to detect lines and curves in pictures*, *Communications of the ACM* (1972) no. Volume 15 Issue 1, 11–15.
- [12.11] ATLAS Collaboration, P. Conde Muiño, *MultiThreaded Algorithms for GPGPUs in the ATLAS High Level Trigger*, Tech. Rep. ATL-DAQ-PROC-2016-045, CERN, Geneva, 2016. <https://cds.cern.ch/record/2239823>.

## 12.4 Model for CPU estimation

---

[12.12] ATLAS Collaboration, *Technical Design Report for the ATLAS Inner Tracker Pixel Detector*, Tech. Rep. CERN-LHCC-2017-021, CERN, Geneva, (*in preparation*).

## 13 Hardware-based Tracking for the Trigger (HTT)

Charged particle reconstruction in the high-pileup conditions expected at the HL-LHC presents a special challenge for the trigger. Use of information from the upgraded tracking detector ITk as early as possible in the trigger selection is a key ingredient in the ATLAS Trigger strategy for the Phase-II TDAQ upgrade. This chapter describes the ATLAS baseline design of a Hardware-based Tracking system for the Trigger (HTT).

Different technologies may be considered for a tracking system for the EF: 1) a hardware-based system (HTT) based on custom-designed Associative Memory (AM) ASICs for pattern recognition and FPGAs for track reconstruction and fitting, and 2) commodity CPU-based servers with or without accelerators (e.g., GPGPUs). The baseline option, in which the HTT is used as a tracking co-processor in the EF, meets the high trigger rate and throughput requirements at the HL-LHC. Several reasons motivate this decision, which include considerable experience in the AM technology, the potential for short latency, a lower power budget and less demanding space requirements compared to other technologies, its cost effectiveness and the independence of its cost from the commodity computing market, availability of in-house expertise, and the capability to evolve the HTT system for use in the hardware-based Level-1 trigger, should ATLAS need to change to a dual L0/L1 trigger system, as described in Chapter 14. Studies of GPGPU usage in the EF, are detailed in Section 12.3, and a discussion of the estimated EF CPU resources that would be required without the HTT is presented in Section 12.4.

As illustrated in Fig. 12.1, depending on the trigger signature, two types of requests from the EF will be transmitted to the HTT: finding tracks in regions of interest identified by the previous stage of selection (regional tracking, rHTT), and reconstructing tracks in the entire ITk coverage (global tracking, gHTT). Both regional and full-scan track reconstruction over the full ITk acceptance ( $|\eta| < 4$ ) are provided by the same hardware system.

The rHTT searches for all tracks with  $p_T > 2$  GeV in limited regions around Level-0 trigger objects that can profit from track information (for example, single high- $p_T$  leptons and four-jet triggers, as summarised in Table 6.5). The rHTT will operate at the full L0A rate of 1 MHz and will process on average 10% of the ITk detector data in these events (see Section 6.12 for a detailed discussion of this design parameter). For these tracks, single-stage reconstruction is performed using eight ITk detector layers<sup>1</sup>; this choice is the result of optimising the overall rHTT performance. As demonstrated by the studies presented in



Chapter 6, regional tracking will accomplish a significant reduction in the Level-0 output rate from 1 MHz to 400 kHz; the remaining events will be further processed by the EF.

The gHTT searches for all tracks with  $p_T > 1$  GeV at a nominal rate of 100 kHz (see Section 6.12 for a detailed bottom-up estimate of this rate). For these tracks, in addition to a first stage of track reconstruction (in common with the rHTT tracks), a second stage of processing is performed over all ITk detector layers to provide high-quality tracks with better purity and track parameter resolutions than those found by regional tracking. These tracks enable EF algorithms such as lepton isolation, primary vertex reconstruction,  $b$ -tagging, and missing transverse energy reconstruction as described in Chapter 6 and summarised in Table 6.6.

The implementation of the HTT design is an evolution of the FTK [13.1] system being commissioned in Run 2, which in turn builds on a similar system developed for the Silicon Vertex Tracker [13.2][13.3] at the Collider Detector Facility (CDF) at the Fermilab Tevatron [13.4]. For the Phase-II upgrade an ATCA-based hardware implementation is foreseen, where tracking processor carrier boards each house two mezzanine cards. A high-level overview of the HTT system is given in Fig. 13.1; the number of boards in the system is allocated to have a six-to-one ratio of first-stage to second-stage track reconstruction. AM ASICs are used for pattern recognition, while FPGAs are used for hit clustering, track fitting, track extrapolation, and duplicate removal. A comparison between the HTT and the FTK is presented in Section 13.3.

Track reconstruction in the HTT proceeds as follows. The inputs to the HTT are ITk hits, which are transmitted via commodity network along with the tracking request(s) from the EFPUs. There are two stages of processing. The first stage of processing is the same whether the request is for regional or global tracking. The hits from the eight ITk layers are clustered into consecutive ITk strip or pixel channels (so-called “superstrips”). Next, these superstrips are compared to a large bank of pre-computed template patterns (derived from simulated single-muon tracks); this comparison is performed in the AM ASICs. For each hit pattern that matches a template, the track parameters and quality are computed from the corresponding full-resolution hits in a FPGA. Tracks which share more than a given number of ITk hits (duplicate tracks) are removed. rHTT tracks are complete after this stage and are transferred back to the EGPU. For global HTT processing, each eight-layer track candidate goes through a second stage of processing in which its track fit is extrapolated to the remaining ITk layers and associated to any matching hits. A full track fit is then performed to achieve the best possible track parameter resolution. Duplicate track removal is also performed on gHTT tracks after the second-stage track fit. gHTT tracks are then transferred back to the EGPU for use in the final EF decision.

---

<sup>1</sup> The eight ITk layers used in the first-stage processing are chosen amongst the strip and pixel detector layers as follows: up to seven strip layers, if available, complemented by the appropriate number of pixel layers (the one(s) furthest away from the luminous region) to reach a total of eight. Hence, the choice of layers and the mixture of strip vs. pixel layers varies as a function of  $\eta$ .



An overview of the **HTT** hardware implementation is presented in Section 13.1, followed by a description of the interface to the **EF** in Section 13.2 and a comparison to **FTK** in Section 13.3. The functional description of the system is given in Section 13.4. Performance studies are shown in Section 13.5 and, finally, Section 13.6 describes each individual hardware component.

### 13.1 Overview of the HTT Architecture

The **HTT** is organised as an array of independent tracking units called **HTT** units as shown in Figure 13.1. The main hardware building blocks of the **HTT** are **ATCA** boards called Tracking Processors (**TPs**). There are two types of **TP** boards: Associative Memory Tracking Processor (**AMTPs**) and Second-Stage Tracking Processor (**SSTPs**). Each **TPs** connects to the **EF** processor farm through a **HTT InterFace (HTTIF)**.

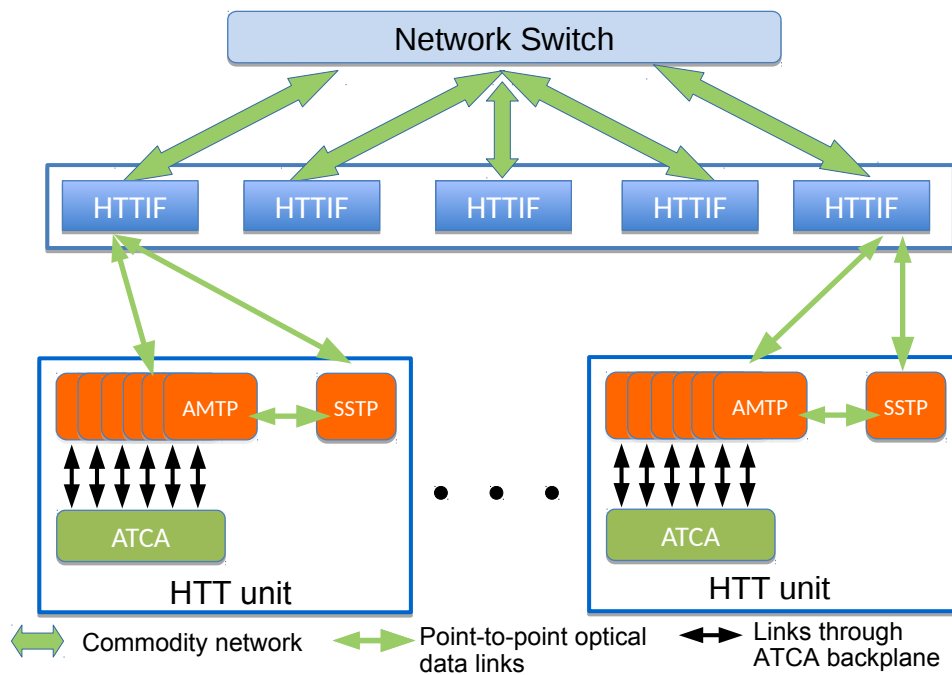


Figure 13.1: Overview diagram of the **HTT** system showing interconnections within **HTT** units and with the **HTTIF**.

Each **HTT** unit comprises a set of six **AMTP** boards and one **SSTP** board and will perform both regional and global tracking in a specific  $\eta - \phi$  region of the track parameter phase space. Each **TP** holds two mezzanine cards. The **AMTP** boards each hold two **Pattern Re-**

## 13.1 Overview of the HTT Architecture

cognition Mezzanines (PRMs) and the SSTP boards each hold two Track-Fitting Mezzanines (TFMs), as shown in Figure 13.2.

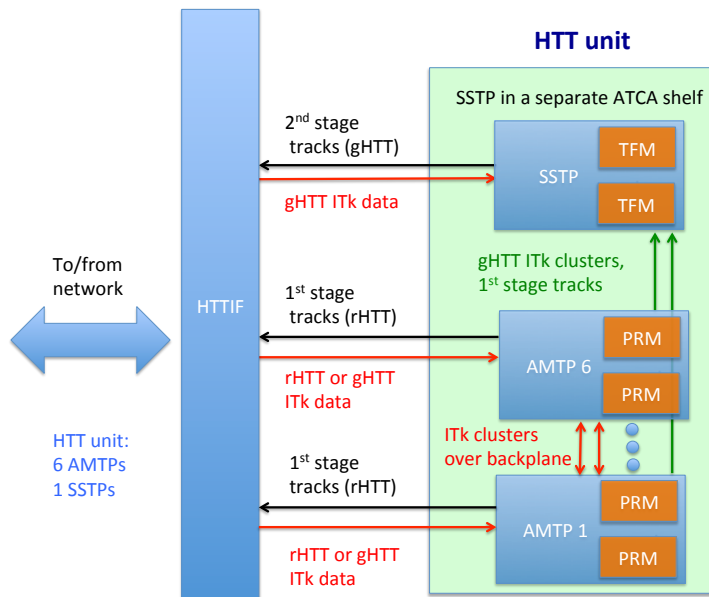


Figure 13.2: Diagram of a **HTT** unit showing interconnections within **HTT** modules and with the **HTTIF**. Each **HTT** unit comprises six **AMTPs** and one **SSTP**. Each **AMTP** and **SSTP** house two mezzanine cards called **PRMs** and **TFMs**, respectively.

The track-finding algorithm is divided into two stages; the first stage is used for both regional and global tracking, while the second is only used for global tracking.

Clustering is the first processing step of the **HTT**. Both the **AMTPs** and **SSTPs** cards share **ITk** clusters; all clusters for a given event are aggregated and then sent for processing to mezzanine cards. The heart of the first-stage processing is **AM ASICs** pattern matching. The **AM** pattern matching occurs in the **PRM** and incorporates 8 layers of the **ITk**. Candidate tracks from **AM** pattern matching are used to retrieve full resolution clusters and perform first-stage track fitting in the **FPGA**.

The second step is done only for global tracking where the tracks found in the **AM** step are extrapolated to the remaining **ITk** layers to find all hits belonging to them, and a full track fit is performed to achieve the best possible track parameter resolution, as required for the use cases of global tracking. This operation is called second-stage processing and is performed in the **TFMs** that each house two **FPGAs**; two **TFMs** are mounted on each **SSTP**.

---

Several important considerations were taken into account in the system design. First, the partitioning of the **HTT** system is optimised by extrapolating the **FTK FPGA** resource usage to the **gHTT** using a ratio of 6 **AMTPs** to 1 **SSTP**. This results in a 50% logic utilisation in the **TFM FPGAs**, which satisfies the target requirement. Second, the **HTT** system is designed for regional tracking in 10% of the detector at 1 MHz using the outer eight **ITk** layers, which is sufficient to reduce the Level-0 output rate to 400 kHz. The use of inner layers also for regional tracking at constant system cost has been evaluated. For example, a reduction of regional processing from an average of 10% of the detector to an average of 7% of the detector (with a corresponding reduction in the number **AMTP** cards to be replaced with additional **SSTP** cards and **HTTIF** servers) would allow second-stage processing for regional tracking for an average of 7% of the detector data per event. Finally, the distribution of data from **HTTIF** to the seven **TP** cards is balanced to send a similar amount of **ITk** data per **TP** card. This is done to balance the amount of cluster-finding processing on each **TP** board.

The concept of a **HTT**  $\eta - \phi$  region is distinct from the concept of a **RoI**. A **HTT**  $\eta - \phi$  region comprises all the **ITk** detector elements intercepted by tracks that have track parameters  $\eta$  and  $\phi$  within the specified  $\eta - \phi$  range,  $p_T$  above a certain threshold (2 GeV for **rHTT** and 1 GeV for **gHTT**),  $|z_0| < 15$  cm and  $|d_0| < 2$  mm. Therefore, two **HTT**  $\eta - \phi$  regions adjacent in  $\phi$  share a number of **ITk** detector elements due to the curvature of tracks, while two **HTT**  $\eta - \phi$  regions adjacent in  $\eta$  will also share a number of **ITk** detector elements due to the range of track  $z_0$  that they cover. This overlap is discussed further in subsequent sections and is taken into account fully in the dataflow estimates.

## 13.2 Interface with Event Filter

Tracking requests with the corresponding **ITk** data will be sent from **EFPUs** to the **HTT** over the network. The size, shape and number of **ITk** regions to be processed for regional tracking can be freely chosen within the **HTT** processing limits (up to 10% of **ITk** data at 1 MHz). A static mapping between **ITk** modules and the **HTT** units will be used. An **ITk** module will be mapped by the **EF** to one or more **HTT** units to allow for overlaps. Each **HTT** unit will be assigned to find tracks in a given **HTT**  $\eta - \phi$  region that corresponds to all tracks with the  $\eta$  and  $\phi$  parameters within that range. In order to do this, the **ITk** module mapped to a **HTT** unit will include those needed to have full acceptance for the bending of the tracks in the magnetic field and for the size of the luminous region (beam spot).

The steps that will return **HTT** tracks are as follows:

1. The **EFPU** algorithm decides whether the request is for regional or global tracking. For global tracking, data from all **ITk** modules mapped to a given **HTT** unit will be sent to the unit for tracking. In the case of regional tracking, the **EFPU** will determine which **ITk** modules correspond to the **RoI** where tracks are requested. All information available at this stage of **EFPU** processing can be used to select the **ITk** modules. The tracking requests will be sent to all relevant **HTT** units.

2. The **EFPU** will build a message for each **HTTIF** that is involved in the processing of the given event. The message will contain data for the corresponding modules and for each module the indication of the **HTT** destination(s) (**HTT** card or input link). Multiple destinations are indicated to duplicate data to different **HTT** units connected to the same **HTTIF** within the **HTTIF** itself. The single message is important for keeping together all data to be processed for one request by the 4 **HTT** units connected to the same **HTTIF**. The message will also contain the indication of regional or global tracking to determine the  $p_T$  threshold and whether to perform second-stage processing.
3. The **HTTIFs** will process messages, distributing module data to optical links feeding the **TP** cards
4. The **HTT** units will process incoming data providing first-stage or second-stage tracks back to the **HTTIF**.
5. Through the data processing, an identifier is kept that will allow returning the tracks found to the requesting **EFPU**.

Once an **EFPU** receives the tracks found by all relevant **HTT** units, a check for duplicate track removal will be performed amongst tracks found in **AMTPs** covering adjacent  $\eta$  or  $\phi$  regions. The chances of duplicate tracks appearing in different units is small, given that each **AMTP** has a unique set of patterns stored in its **AM ASICs**. For both regional and global tracking the duplicate removal is done at the level of the **AMTP** as described. For global tracking a second pass of duplicate removal is performed in the **SSTP** after second-stage track finding.

Figure 12.1 shows that for some events the **EFPU** will send, in a first step, an **HTT** request to the regional tracking and in a second step to the global tracking.

### 13.3 Comparison to FTK

The **HTT** design builds on the **FTK** experience. The basic firmware algorithms from **FTK** are also present in the **HTT** system as described in this chapter. From the hardware point of view, the **HTT** improves in modularity, reducing the type of cards to a single main board and two mezzanine cards. This reduces the number of interfaces between systems and with a corresponding reduction in the amount of firmware needed. The **HTT** will allow expedited commissioning for three main reasons. The first is intrinsically related to the higher modularity that will allow a minimal processing slice to be built with just two main cards. The second is that the **HTT** will be organised as a co-processor farm, as indicated in Fig. 13.1, in which each unit is independent of the others. An **HTT** unit is built of just 7 modules. For comparison the **FTK**, which works as a single system, combines 322 modules working together. The third is that the **HTT** will receive data from the network rather than directly from the detector. This will allow the commissioning to be initiated before the start of Run 4 data-taking. Most of the commissioning will be performed with simulated data

---

only using pre-production hardware as described in 16.4.4. The proposed system design is more integrated and structurally simpler than the current FTK, which will make the synchronisation of the components easier. Furthermore, the general network access for HTT would in principle allow its usage also as an offline co-processor for Monte Carlo events.

## 13.4 Functional description of HTT

### 13.4.1 Data preparation

The data input to HTT are hits from the ITk. The pixel data is made into clusters using clustering algorithms. The clusters are converted to so-called superstrips, which are groups of consecutive silicon strip or pixel channels (see Fig. 13.4). Each superstrip is given a unique SuperStrip Identifiers (SSIDs). Pixel detector elements have coordinates in two dimensions and are formed into superstrips by dividing by both a superstrip width and a superstrip length. The superstrip width can be any integer value. The choice of superstrip dimensions is closely linked to the choice of the number of layers to use at the AM step and the choice of the actual layers, for a given fixed size of the pattern banks, and the choices presented in this Chapter are the outcome of optimisation studies. The list of strip and pixel superstrip widths used in the performance studies is shown in Table 13.1. These widths were found to yield a relatively low number of pattern matches from combinatorial background, while offering high efficiency for single muons and a total number of patterns within the expected hardware capacity. The choice of superstrip width is important as it has a strong effect on the number of patterns required and the number of false matches; hence, different superstrip widths are used in different parts of the detector.

### 13.4.2 Pattern matching

The HTT pattern recognition method is based on the FTK method currently used for the ATLAS Inner Detector, applied to the ITk layout and the HL-LHC levels of pile-up. In this method, a large number of single-muon tracks in simulated training events is used to form template patterns with coarse-granularity SSIDs from a number of layers of the ITk. A collection of these template patterns is called a pattern bank.

In the hardware implementation (described in detail in Section 13.6), these banks are stored in AM ASICs in the PRM. A pattern describes a sequence of eight SSIDs in different layers of the detector.

## 13.4 Functional description of HTT

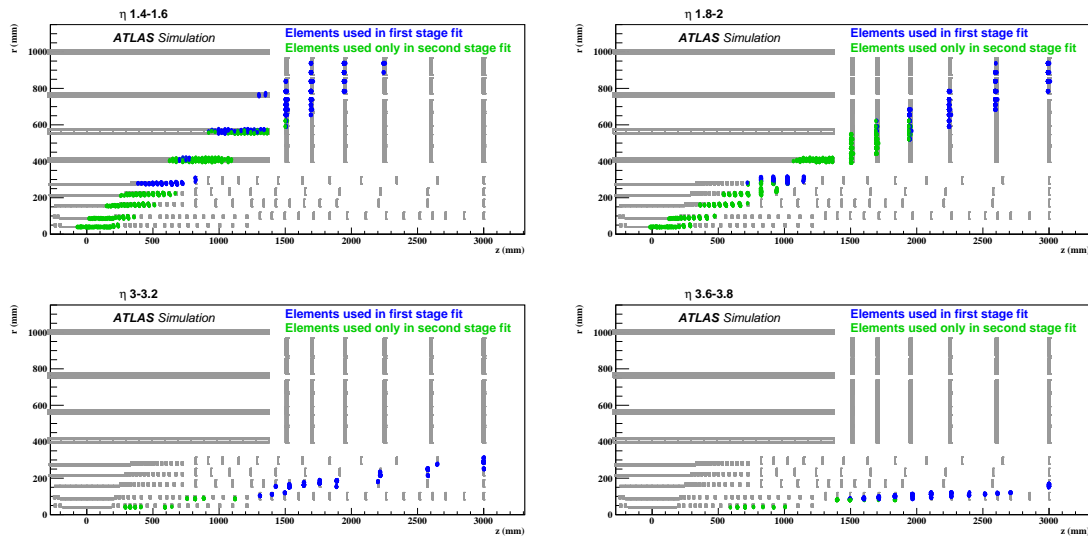


Figure 13.3: Layers used for 1<sup>st</sup> and 2<sup>nd</sup> stage fitting in four  $\eta$ -regions. The hits from the eight layers used in the first-stage processing are shown in blue, while hits in layers used in the second-stage processing are shown in green.

**Layers** Choosing which of the available layers to use affects the number of patterns required, the number of false matches and the resolution of the fitted track. The two sides of the double-sided strip staves (petals in the end-caps) are treated as two separate layers while one physical pixel layer is one layer in the pattern bank. Which layers are used depends on the pseudorapidity of the track. For the barrel region, a fixed set of layers is used. For the transition and end-cap regions, multiple sets of layers are defined depending on which layers are hit in the training events. The first set of layers at low  $\eta$  are all barrel layers and progressively going to higher  $\eta$  more end-cap layers are substituted in each subsequent set. At high  $\eta$ , tracks do not leave hits in the strip end-cap layers, hence multiple pixel end-cap hits must be used. Figure 13.3 shows layers used for first- and second-stage fitting in four  $\eta$ -regions.

**Creating the pattern bank** The pattern banks are generated from the full simulation using training muons. It is planned to speed up the pattern generation in the future using samples with multiple training muons. In each of the  $\phi \times \eta = 0.2 \times 0.2$  regions up to 100M muon events are processed. Clusters from ITk layers hit by the training muon are recorded. For each event, only clusters with a bar-code indicating that they come from the primary interaction are considered. Each cluster is converted to a superstrip with a width chosen for its layer (for pixel clusters this is done for both coordinates). The superstrip is then stored in the appropriate logical layer in a pattern. If the pattern already has a hit stored for the logical layer, the new hit is ignored, i.e. patterns are not made with all the possible combinations. If there are enough layers hit in the training event, the pattern is inserted into a

set. If the insertion fails because there is already a pattern with this sequence of superstrips, the usage count of the existing pattern in the set is incremented to indicate that it has been found again. At the same time, patterns with superstrips of twice and four times the width are also made and inserted into pattern sets. The relationship between the stored patterns for different widths of superstrips is stored for later creation of pattern banks with variable width superstrips. The usage count of a pattern is later used to reduce the size of the pattern bank to fit the available space in hardware. A pattern with low usage count is a rare pattern typically created by a scattered track. Removing patterns with low usage count has a very small effect on signal efficiency but a large effect on space required to store patterns.

**Variable-width superstrips and “Don’t Care” bits** Once all the patterns have been made, similar patterns (tracks A and B in Fig. 13.4) can be combined by setting some of the least significant bits of the superstrip number to be ignored in some patterns. Such bits are called “Don’t Care” (DC) bits. The use of DC bits offers better efficiency and lower pattern matching rate for a fixed number of patterns. The DC bits are assigned with the following procedure. Using the relationship between the finest granularity pattern bank and a corresponding bank with double the superstrip size, a list is built of all the “child” patterns shared by each “parent”. These lists are iterated over, comparing the patterns to see if they can be combined into a single pattern by ignoring the least significant bit of the superstrip number. This process can set at most 1 ignore bit per layer, since the parent list has double the superstrip width. To add more DC bits on a single layer, the process is repeated with a parent created after doubling again the superstrip width etc. The number of DC bits that may be set can be specified on a per layer basis.

Finally, the remaining patterns are sorted in decreasing “popularity” and written to the stored pattern bank. The “popularity” of a pattern is the number of muons that would create this same pattern.

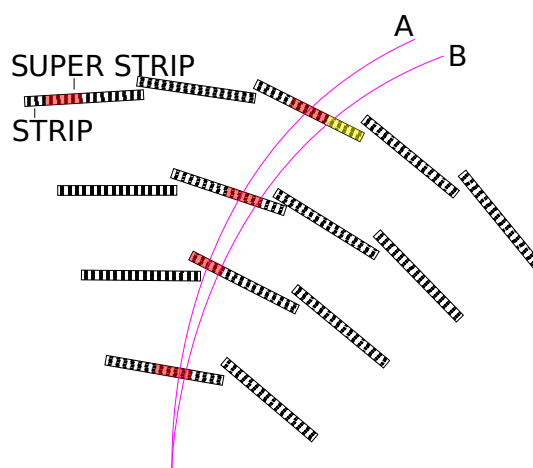


Figure 13.4: Illustration of tracks traversing layers divided into superstrips in a tracker.

**Wild cards** When a pattern has fewer than eight layers hit, the missing layers are marked as wild cards in the pattern. In the pattern matching stage, wild cards are always considered to match. The missing hits in the training muon tracks is usually due to the particle passing through an inactive part of the detector. Figure 13.5 shows where there are missing hits in the silicon strip sensors on both sides of a stave as a function of  $\eta$  and  $z_0$ . In the left hand plot, the gaps between the silicon strip sensors can be clearly seen, since the gaps are aligned for tracks travelling almost vertical away from the beamline in this  $\eta$  range. The pattern bank for this range has around 150k out of a million patterns with two wildcards. In the right-hand plot, no correlation can be seen since tracks at the larger  $\eta$  that go through a gap on one side of the stave hit the edge of the module on the other side. The pattern bank for this range has only 17k out of a million patterns with two wildcards.

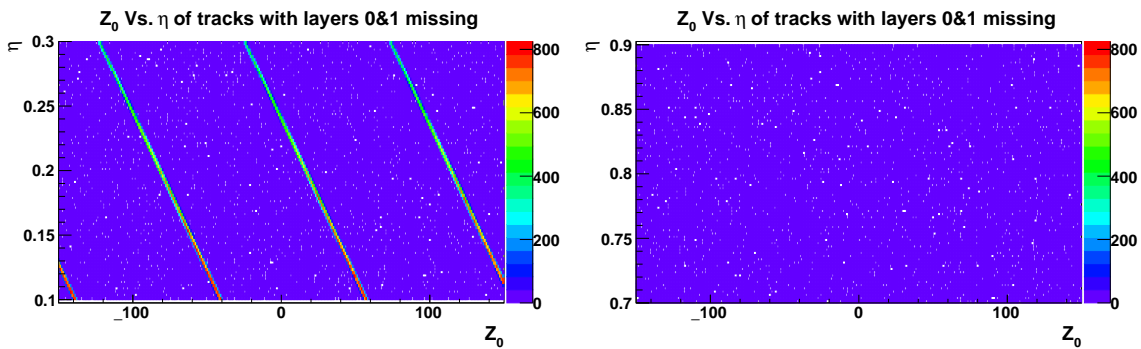


Figure 13.5: Correlation between missing hits on opposite sides of the same stave in the *ITk* strips in the  $\eta$  range 0.1-0.3 on the left and 0.7-0.9 on the right.

Since missing hits due to non-geometrical inefficiencies are rare, patterns with accidental DC bits would have a low “popularity” and would be unlikely to make the cut on the number of patterns described below.

Wild cards are only used to account for inefficiencies due to detector geometry, not to reduce the size of the pattern bank required. For the latter, variable width superstrips are used as described above.

### 13.4.3 Track fitting

**First-stage track fitting** Track fitting is implemented in firmware and performed in an FPGA in the PRM. The firmware implementation will be discussed 13.6. It takes the full-resolution hits from the roads passed by the pattern matching and calculates the track parameters and  $\chi^2$  of the fit. The track parameters  $p_i$  are calculated using a linear interpolation:

$$p_i = \sum_{j=1}^N C_{ij}x_j + q_i, \quad (13.1)$$



where  $x_j$  are the full-resolution local cluster coordinates and  $(C_{ij}, q_j)$  are constants that are unique for each sector, where a sector is defined as a combination of one module from each of the 8 layers used in the first-stage processing. The constants are determined from a large sample of simulated muon tracks with the same parameter ranges and distributions as those used in generating the patterns (see section 13.4.2). The quality of the fit is evaluated using a linearised  $\chi^2$  method which is fast to compute with [FPGAs](#):

$$\chi^2 = \sum_{i=1}^{N-5} \left( \sum_{j=1}^N A_{ij}x_j + k_i \right)^2, \quad (13.2)$$

where  $A_{ij}$  and  $k_i$  are additional constants needed per sector. Several thousand sectors, each with its one sets of constants, are required for fitting a  $\eta \times \phi = 0.2 \times 0.2$  region. Each sector consists of multiple sets of constants to cover possible wildcard configurations in the pattern banks. Without any compression or optimisation, a  $0.2 \times 0.2$  region requires about 40 million coefficients, which need to be stored either in external memories on the [PRM](#) or in the internal [FPGA](#) memory. Preliminary studies to reduce this number indicate that significant savings are possible.

**Second-stage track fitting** The second-stage fitting is performed in a [FPGA](#) on the [TFM](#). For each track the [TFM](#) calculates the 5 helix parameters and the  $\chi^2$  of the fit. The [TFM](#) receives through the [SSTP](#) the 8-layer tracks from 6 [PRM](#) cards and all the hits from the detector layers not used by the [PRM](#). There are two main functions carried out on the [TFM](#): the Extrapolator and the Track Fitter. The Extrapolator finds the hits on the additional silicon layers that are close to a [PRM](#) track, starting from the clusters associated to the first-stage track. The Track Fitter fits the hits on the [PRM](#) track with each combination of hits on the other layers and applies a  $\chi^2$  cut. Those track candidates passing the cut are sent to the [SSTP](#) where duplicate track removal is carried out before sending the tracks to the Event Filter.

The first step in the Extrapolator process is the Data Organiser, which is a database built on the fly that stores the hits for each detector layer so that those near a [PRM](#) track can be rapidly retrieved without having to scan the full hit list. The Extrapolator uses the hits in the [PRM](#) track, along with the detector sector the track is in, to define regions in the other layers in which to search for hits. Those hits along with the [PRM](#) track are sent to the Track Fitter. For each track candidate, the Track Fitter carries out linear calculations to determine each of the helix parameters and the  $\chi^2$  of the fit (Equations 13.1 and 13.2). The inputs to the calculation are the hit coordinate in each detector layer and a set of pre-stored constants that differ for each sector of the detector.

### 13.4.4 Duplicate removal

The removal of duplicate tracks in an event is done with the HitWarrior algorithm run in the [TP](#). The HitWarrior is a part of both track fitting stages. To identify duplicates, the HitWarrior first forms groups of tracks that share more than a given number of [ITk](#) hits,  $N_{min}$ . The parameter  $N_{min}$  can be adjusted to tune the HitWarrior to achieve high efficiency and low duplicate rate. Tracks in a group are compared and those with fewer [ITk](#) hits are removed. Amongst tracks with the same total number of [ITk](#) hits, the one with the lowest  $\chi^2$  is kept. The selection criteria could easily be replaced by a more sophisticated set in the future.

For [gHTT](#), the fake rate after second state fitting is  $<0.2\%$ , where fake tracks are defined as sharing less than 80% of its clusters with a truth track.

## 13.5 HTT Performance Studies

### 13.5.1 Associative Memory (AM) Pattern-Matching Performance

The performance studies described in this section were carried out predominantly for  $p_T > 4$  GeV using full detector simulation in four representative  $\eta - \phi$  HTT-regions of  $0.2 \times 0.2$  within the acceptance of the [ITk](#) strips. In one of these regions ( $0.7 < \eta < 0.9$ ), detailed studies all the way down to  $p_T = 1$  GeV were also performed and used to extrapolate/compare to the performance in other regions and/or the entire detector. As will be shown, the performance vs.  $\eta$  is rather uniform, lending confidence that the results are valid across the entire [ITk](#) coverage. In the pattern matching studies, one million patterns with the highest “popularity” were used for  $p_T > 4$  GeV, two million patterns for  $p_T > 2$  GeV, and four million patterns for  $p_T > 1$  GeV. The four million patterns per [HTT](#)-region matches the overall proposed size of [HTT](#).

The superstrip widths and selection of [DC](#) bits were chosen to ensure high efficiency within the one million patterns.

Adding more [DC](#) bits or increasing the superstrip widths would increase the number of fake matches due to combinatorics, indicating that the current working point is close to optimal. The track-finding efficiency was studied for single muons (and in some cases electrons and/or pions), with a flat  $1/p_T$  spectrum between 4 (or 1) and 400 GeV, and the overall performance, including all dataflow estimates, was estimated from minimum bias events with  $\langle \mu \rangle = 200$  pile-up (referred to simply as pile-up) and from dijet events filtered so that a parton with  $p_T > 30$  GeV points to one of the  $\eta - \phi$  regions under study (referred to as “jets” throughout this section).

Using a configuration with 7 strip layers and a single pixel layer (the outermost barrel layer), and a “7-out-of-8” matching logic, an efficiency of  $\approx 99\%$  can be achieved, with

Table 13.1: Pattern-matching performance for the AM step simulated on minimum bias  $\langle \mu \rangle = 200$  pile-up events ( $\eta \times \phi = 0.2 \times 0.2$  regions,  $p_T > 4$  GeV). All numbers are average per event. The superstrip dimensions and DC bits configuration (per layer) is also given (see text). The efficiency and the number of matches refers to the number of roads found during the AM processing. The found roads are input to the next processing steps.

$\eta$ range	muon eff.	mean matches pile-up	99% interval matches in pile-up	Superstrip width			DC bits
				pixel	barrel	end-cap	
$0.1 < \eta < 0.3$	99.1%	31	151	33/402	40		21111122
$0.7 < \eta < 0.9$	99.2%	21	93	33/402	40		21111122
$1.2 < \eta < 1.4$	98.8%	42	159	33/402	40	20	21111122
$2.0 < \eta < 2.2$	98.7%	10	56	16/200		10	21111122

a rate of matched patterns that can be handled by the fitter, as shown in Table 13.1 for  $p_T > 4$  GeV. The inner and two outer layers use a maximum of two DC bits and the rest a maximum of one DC bit. The rates of matched patterns for  $p_T > 2$  GeV and  $p_T > 1$  GeV are shown in Tables 13.3, 13.4, 13.5, and 13.6.

**Pattern matching at high  $\eta$**  Beyond  $|\eta|$  of 3.0 the tracks do not hit the ITk strip endcap disks at all; hence, only pixel information is available to make tracks. A pattern bank for the  $\eta$  region 3.0-3.1 using only the pixel endcaps was created. Since this region is half of the size of the other regions under study, the target pattern bank size was 0.5M patterns but even without using DC bits the resulting bank only had 0.4M patterns. The performance is summarised in table 13.2.

Table 13.2: Pattern-matching performance in the  $\eta \times \phi = 0.1 \times 0.2$  region, for  $3.0 < \eta < 3.1$  and  $p_T > 4$  GeV. The pattern bank uses pixel endcap layers only. All numbers are average per event.

$\eta$ range	muon eff.	mean matches		Superstrip width pixel
		muon	pile-up	
$3.0 < \eta < 3.1$	95.6%	25	8	16/200

**Performance with inefficient detector channels** As a crude way of simulating a detector with only  $x\%$  efficient channels, the pattern matching has been run with a random  $100 - x\%$  of clusters ignored. The results of this study are shown in Fig. 13.6. The simulation shows that by using 6/8 matches the efficiency will be 95% or higher (at a cost of increased pattern matches) and even with the baseline 7/8 matching logic, the pattern matching efficiency is above 95% for cluster efficiency of 97%. This is very much a worst case simulation, since whole clusters are dropped, not single hits. Moreover, large inefficiencies are not expected because there are more layers in the ITk than used for pattern matching. In the

case of failures where single detector elements or layers stop working, the patterns can be regenerated using redundant elements or layers.

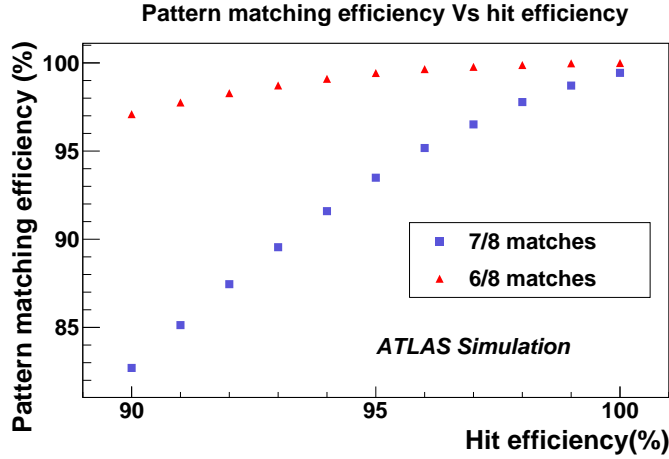


Figure 13.6: Effect of random inefficient channels on pattern-matching efficiency.

### 13.5.2 Track-Fitting Performance

The simulated track-fitting performance at 99% track-finding efficiency (for single muons) in  $0.7 < \eta < 0.9$  region for muon, jets and minimum bias events at  $\langle \mu \rangle = 200$  and for three  $p_T$  thresholds is shown in in Table 13.3. The jets used in the studies are weighted to give a  $p_T$ -spectrum corresponding to Level-0. The table shows per event the average number of fits required to fit all the hit combinations found in pattern matching, the average numbers of tracks found after the quality cut, the average numbers of tracks after duplicate removal (HitWarrior) and average number of sets of first-stage fit constants<sup>2</sup> required to fit the event. The presented numbers are used to extract the requirements for the system.

The fitting performance for first-stage track fitting in rHTT for single jets at  $\langle \mu \rangle = 200$  is shown for four  $\eta$ -regions in Table 13.4. The results shown for  $p_T > 2$  GeV have been extrapolated from studies done at  $p_T > 4$  GeV using weights from the  $0.7 < \eta < 0.9$  region (extracted from Table 13.3).

The same detector configuration as in the pattern matching studies was used. The number of fits is significantly reduced with the HitWarrior algorithm, resulting in a corresponding reduction on the number of tracks found.

<sup>2</sup> A set of fitting constants includes all the pre-calculated values used to calculate the  $\chi^2$  and the track parameters.

Table 13.3: First-stage track fitting performance for  $0.7 < \eta < 0.9$  region at  $\langle \mu \rangle = 200$ . All numbers are averages per event.

particle	min $p_T$	Eff. (%)	# roads	# fits	# tracks $\chi^2 < 40$	# tracks HitWarrior	# fit constants
muon	1 GeV	99.5	144	1115	55	4.6	73
muon	2 GeV	99.1	79	586	23	1.9	40
muon	4 GeV	99.2	48	313	16	1.2	23
jets	1 GeV		195	1519	77	6.2	97
jets	2 GeV		104	804	29	2.4	52
jets	4 GeV		51	344	13	1.1	26
min-bias	1 GeV		110	842	38	3.6	58
min-bias	2 GeV		48	359	6	0.8	27
min-bias	4 GeV		21	133	1	0.2	12

Table 13.4: First-stage track fitting performance in jets at  $\langle \mu \rangle = 200$  ( $p_T > 2$  GeV studied in  $0.7 < \eta < 0.9$ , and extrapolated from 4 GeV to 2 GeV in other regions). All numbers are averages per event.

$\eta$ range	# roads	# fits	# tracks $\chi^2 < 40$	# tracks HitWarrior	# fit constants
$0.1 < \eta < 0.3$	170	1521	60	3.3	75
$0.7 < \eta < 0.9$	104	804	29	2.4	52
$1.2 < \eta < 1.4$	170	1402	71	4.8	90
$2.0 < \eta < 2.2$	65	240	64	3.7	20

Results for first-stage fitting in [gHTT](#) is shown for jets in [Table 13.5](#) and for minimum bias in [Table 13.6](#). The performance is extrapolated from 4 GeV to 1 GeV in the same way as for the first-stage track fitting in [rHTT](#).

Table 13.5: First-stage track fitting performance in jets at  $\langle \mu \rangle = 200$  ( $p_T > 1$  GeV studied in  $0.7 < \eta < 0.9$ , and extrapolated from 4 GeV to 1 GeV in other regions). All numbers are averages per event.

$\eta$ range	# roads	# fits	# tracks $\chi^2 < 40$	# tracks HitWarrior	# fit constants
$0.1 < \eta < 0.3$	314	2874	159	8.5	138
$0.7 < \eta < 0.9$	195	1519	77	6.2	97
$1.2 < \eta < 1.4$	324	2649	189	12	167
$2.0 < \eta < 2.2$	125	454	171	10	37

## 13.5 HTT Performance Studies

Table 13.6: First-stage track fitting performance for minimum bias ( $p_T > 1 \text{ GeV}$  extrapolated from  $4 \text{ GeV}$ ) at  $\langle \mu \rangle = 200$ . All numbers are average per event.

$\eta$ range	# roads	# fits	# tracks $\chi^2 < 40$	# tracks HitWarrior	# fit constants
$0.1 < \eta < 0.3$	166	1481	76	7.2	82
$0.7 < \eta < 0.9$	110	842	38	3.6	58
$1.2 < \eta < 1.4$	218	1639	192	16	111
$2.0 < \eta < 2.2$	53	196	102	9.0	19

### 13.5.3 Muon and Electron Track-finding Efficiencies

The tracking efficiency for muons and electrons is defined as the fraction of events with at least a track with  $\chi^2 < 40$ , only for events in which an offline track is reconstructed. This is then the efficiency measured with respect to offline. The corresponding turn-on curves for first-stage track fitting, estimated on single muons and electrons not embedded in pile-up, are shown in Fig. 13.7. The efficiency for muons is quite flat along the full  $p_T$  range, while electrons show a slower turn-on below  $10 \text{ GeV}$ . The reason of this lower efficiency at low momentum is motivated by the high radiation probability of these particles, which can reduce  $p_T$  below the minimum  $4 \text{ GeV}$  threshold set in the rHTT system. This effect has been verified by excluding from the sample all the electrons below  $10 \text{ GeV}$  and having a recovery of the efficiency in that  $p_T$  range.

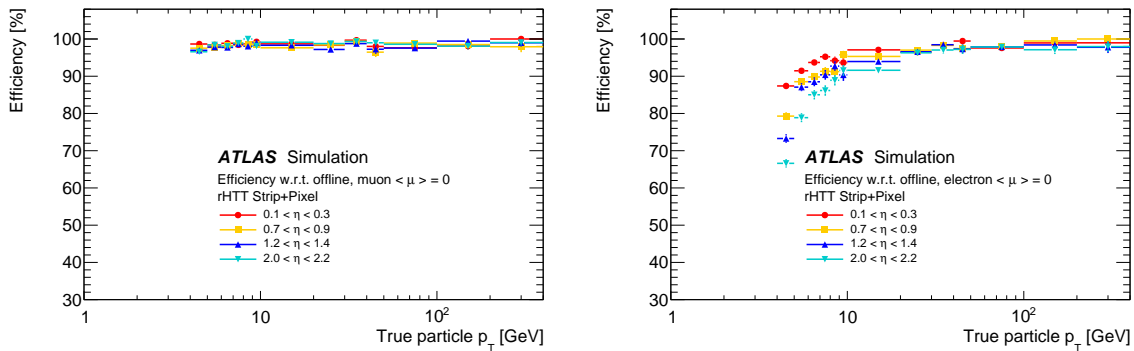


Figure 13.7: First-stage muon and electron track-finding efficiencies in  $\eta$  regions for muons (left) and electrons (right) for  $p_T > 4 \text{ GeV}$ .

### 13.5.4 Resolutions of Track Parameters

As a result of the fit procedures, the resolution on each parameter from first stage fitting is measured in single lepton sample with a spectrum flat in  $1/p_T$  and with  $\langle \mu \rangle = 200$ . The resolution is measured as the RMS of the 95% of the residual distributions (called  $rms_{95\%}$ ),

Table 13.7: First-stage track fitting resolutions ( $rms_{95\%}$ ) for electrons  $p_T > 4 \text{ GeV}$ .

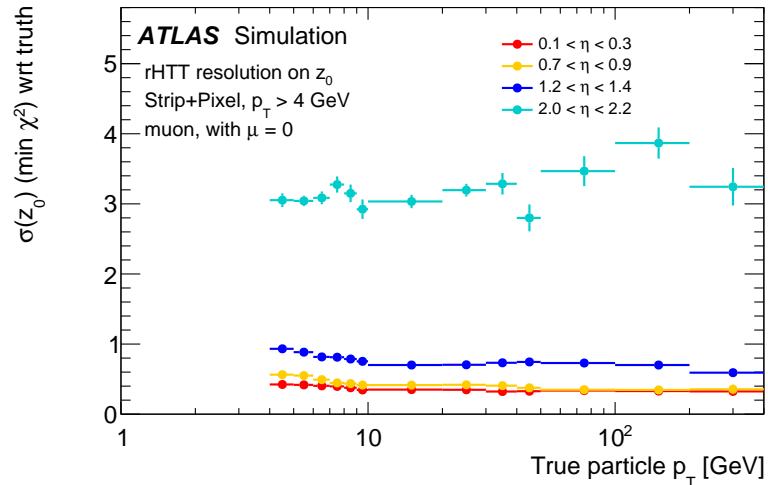
$\eta$ range	$\eta$	$\phi$	$q/P_t [\text{GeV}^{-1}]$	$d_0 [\text{mm}]$	$z_0 [\text{mm}]$
$0.1 < \eta < 0.3$	0.004	0.003	0.021	0.42	2.9
$0.7 < \eta < 0.9$	0.004	0.003	0.031	0.52	4.5
$1.2 < \eta < 1.4$	0.011	0.013	0.048	0.87	19.3
$2.0 < \eta < 2.2$	0.014	0.012	0.059	1.03	22.1

Table 13.8: First-stage track fitting resolutions ( $rms_{95\%}$ ) for muons ( $p_T > 4 \text{ GeV}$ ).

$\eta$ range	$\eta$	$\phi$	$q/P_t [\text{GeV}^{-1}]$	$d_0 [\text{mm}]$	$z_0 [\text{mm}]$
$0.1 < \eta < 0.3$	0.002	0.001	0.003	0.20	0.8
$0.7 < \eta < 0.9$	0.002	0.001	0.003	0.20	0.8
$1.2 < \eta < 1.4$	0.004	0.003	0.007	0.33	3.8
$2.0 < \eta < 2.2$	0.004	0.003	0.014	0.71	7.1

in which 2.5% of each side tail is removed, and the obtained RMS is normalised to the expected gaussian width. The residuals are calculated from the difference of each parameter between the minimum  $\chi^2$  candidate and the truth track parameter. In Tables 13.7 and 13.8 compare the values measured for electrons and muons respectively, for all the eta-regions considered. The resolution in  $2.0 < \eta < 2.2$  region is expected to significantly improve if the barrel layer used in this study is replaced by a disk (the disk geometry was not available at the time of the study). The resolutions from these studies has been used for the trigger performance studies of muons, electrons and tau leptons presented in Chapter 6.

Figure 13.8 shows how the  $z_0$  resolution for muons change as a function of track  $p_T$  for first-stage fitting.

Figure 13.8: Comparison of  $z_0$  resolutions for muons as a function of track  $p_T$  for first-stage fitting.

## 13.5 HTT Performance Studies

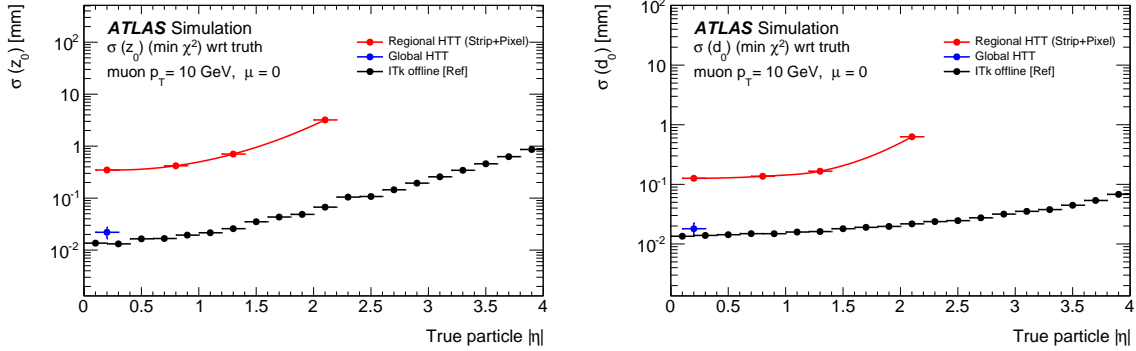


Figure 13.9: Comparison of the  $z_0$  (left) and  $d_0$  (right) resolution for first- and second-stage fitting and offline.

Figure 13.9 shows the  $z_0$  and  $d_0$  resolution for 10 GeV muons as a function of  $\eta$  for the HTT first- and second-stage processing and offline tracking resolution. Only the central region was studied for the second-stage fitting.

### 13.5.5 Simulated HTT Data Size

The data size has been simulated to allow data bandwidth calculations in the hardware. The requirements and estimations in Section 13.6 are derived from these simulations. The fraction of data in ITk, separated by layer, for different trigger objects is shown in Table 13.9. The average RoI fraction per event to be processed by the regional HTT is 2.3% and the average data fraction per layer is ranging between 3.5% and 6.8% for the barrel strip and pixel layers used in rHTT. This data is used for optimising the partitioning of the HTT system and for dataflow and bandwidth calculations.

The number of clusters per ITk layer in a  $0.2 \times 0.2$   $\eta - \phi$  region for  $0.7 < \eta < 0.9$ , for pattern banks with different  $p_T$  thresholds (corresponding to gHTT, rHTT and L1Track), is shown in Fig. 13.10 (left). The number of clusters for pattern banks with  $p_T > 4$  GeV in four  $0.2 \times 0.2$  regions across  $\eta$  is shown in Fig 13.10 (right). The layers shown are those used in the first-stage processing, where Layer 0 is the outermost ITk pixel layer and the others are strip layers. The result is shown for jets in  $\langle \mu \rangle = 200$  pile-up. The cluster occupancy is dominated by minimum bias; events with a jet pointing in the given  $\eta - \phi$  region only have about 10% higher occupancy. This data is used for calculating FPGA resources and power.



Table 13.9: The fraction of data read out in a given RoI for different trigger objects.

Trigger	Object Multiplicity	RoI size	$\eta - \phi$ Fraction	Pixel Layer Data Fraction					Strip Layer Data Fraction			
				0	1	2	3	4	0	1	2	3
single $e$	1	0.2	0.13%	4.8%	3.8%	1.8%	1.2%	1.0%	1.3%	0.8%	0.6%	0.4%
single $\mu$												
single $\gamma$												
forward $e$	1	0.2	0.13%	4.8%	3.8%	1.8%	1.2%	1.0%	1.3%	0.8%	0.6%	0.4%
di- $\gamma$												
di- $e$	2	0.2	0.25%	9.6%	7.6%	3.6%	2.4%	2.0%	2.6%	1.6%	1.2%	0.8%
di- $\mu$	2	0.2	0.25%	9.6%	7.6%	3.6%	2.4%	2.0%	2.6%	1.6%	1.2%	0.8%
$e - \mu$	2	0.2	0.25%	9.6%	7.6%	3.6%	2.4%	2.0%	2.6%	1.6%	1.2%	0.8%
single $\tau$												
di- $\tau$	2	0.2	0.25%	9.6%	7.6%	3.6%	2.4%	2.0%	2.6%	1.6%	1.2%	0.8%
single jet												
large- $R$ jet												
four-jet	5.0	0.8	10.2%	60.0%	52.0%	34.0%	27.5%	23.0%	25.5%	20.0%	17.0%	14.5%
$H_T$	5.0	0.8	10.2%	60.0%	52.0%	34.0%	27.5%	23.0%	25.5%	20.0%	17.0%	14.5%
$E_T^{miss}$	2.8	0.8	5.7%	33.6%	29.1%	19.0%	15.4%	12.9%	14.3%	11.2%	9.5%	8.1%
VBF	2	0.8	4.1%	24.0%	20.8%	13.6%	11.0%	9.2%	10.2%	8.0%	6.8%	5.8%
Support			0.2%	1.6%	1.4%	0.8%	0.6%	0.5%	0.6%	0.5%	0.4%	0.3%
Average per Event			2.3%	17.7%	15.0%	9.1%	7.1%	6.0%	6.8%	5.1%	4.3%	3.5%

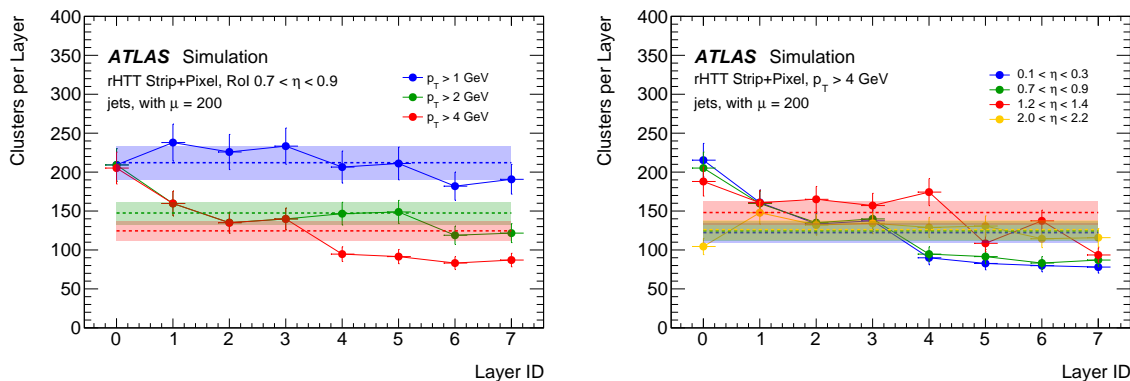


Figure 13.10: The average number of clusters as a function of layer for three pattern bank  $p_T$  thresholds in an  $\eta - \phi$  region with  $0.7 < \eta < 0.9$  (left) and the number of clusters in four  $0.2 \times 0.2$   $\eta - \phi$  regions as a function of detector layer (right). The average number of clusters ( $1\sigma$ ) are shown in the shaded bands.

### 13.6 Description of the HTT Hardware and Firmware

The **HTT** core components shown in Fig. 13.2 will be custom electronic boards that receive **ITk** data from the **EF** via the **HTTIF**, perform the pattern recognition and return track candidates to the **EF**, again via the **HTTIF**.

The **TP** is the main **HTT** processing unit implemented using an **ATCA** main board, where the computing power is distributed among mezzanines plugged into it. The function of the **TP** is to receive input data, perform clustering, share data with other **TPs**, then map clusters to logical layers and send them for processing to the mezzanine cards. Once tracks are received back from the mezzanines, duplicate track removal is performed on the **TP** and then tracks are returned to the **EF** via the **HTTIF**. The physical **TP** boards will be used to implement two distinct functions: the first is the **AMTP** which will perform first-stage tracking using **ITk** data from up to 8 layers, and the second function is the **SSTP** which performs the second-stage processing in the case of **gHTT**, using all the remaining **ITk** layers. The **AMTP** module will house two **PRM**. The **SSTP** module will house two **TFM**.

For intra-crate communication a full-mesh backplane will be used to allow point-to-point connection, while inter-crate communication and communication with the rest of **TDAQ** will be handled through optical links.

The **TP** design features two **FPGAs** that will share the data transmission and control the data traffic through the board. For the Phase-II requirements, today's **FPGAs** can provide transceivers with enough bandwidth to handle the expected data traffic. Each board in the shelf will have the ability to communicate to any other board in the same **ATCA** shelf.

The **PRM** mezzanine is expected to occupy about half of the height of the **AMTP** and a depth of about 15 cm, which holds two **PRMs** per **AMTP**. The **PRMs** will perform **AM**-based pattern recognition and track fitting in **FPGA**.

Distributing most of the computing in the mezzanines allows the same **TP** to be used for the **SSTP**. Second-stage processing will be implemented with the **TFM** which will be mechanically and electrically compatible with the **PRM**.

The connection between each mezzanine and the mainboard will use a high-speed connector for data and possibly a separate connector for power. The power distribution to the **AM ASICs** is an important part of the design. To simplify this the DC-DC converters attached to the core of the **AM ASICs** will be installed directly on the **PRM** and controlled by it.

This modular design provides uniformity of the main boards, flexibility for processing in the mezzanines, and efficient management of the different components.

### 13.6.1 Dataflow requirements

The dataflow through the **HTT** system has implications for the size of the system and the required performance of each module. The bandwidth for the input **ITk** data is calculated separately for 1<sup>st</sup> stage layers and 2<sup>nd</sup> stage layers. The input to the calculation is taken from performance studies. The input parameters, their values and their sources, are given in Table 13.10.

**Data duplication** The model for first-stage calculations on the **HTT**-unit level is shown in Fig. 13.11. The **HTT** system massively parallelises its task by dividing the reconstruction into multiple sub-units, each responsible for delivering tracks with  $(\eta_{\text{track}}, \phi_{\text{track}})$  contained in some rectangular region of the full detector space. Each **HTTIF**, **HTT** unit, **TP** board, **PRM** and **TFM** are responsible for such a region. However, a single silicon detector element may be necessary to reconstruct tracks in multiple such regions. For example, it is necessary to send data from the innermost pixel layer to multiple **HTT** units because the extended length of the beam-spot allows tracks of a given  $(\eta_{\text{track}}, \phi_{\text{track}})$  to impinge upon pixel modules with a large spread in  $z$ . Similarly, bending of low- $p_T$  tracks in the magnetic field causes significant data duplication at large radii. The ‘duplication factor’ of a module is defined as the number of hardware units requiring data from that module.

Simulated samples are used to calculate the number of hardware units (i.e., **TP** boards, **TFMs**, etc.) to which a given silicon module must be sent so that the system is able to reconstruct all tracks in the full detector volume. This number is hence referred to as the ‘duplication factor’ of the module, defined for a specific set of hardware units. For each hardware unit in the system, single muons are used to establish a list of modules that are penetrated by the tracks which the unit is tasked to reconstruct. Next, minimum bias events are processed, with the number of hits per module recorded and multiplied by word size to find the total average event size per module. *Average* duplication factors are then computed by taking the weighted average of duplication factors for each module, with weights given by the per-module event sizes,

$$\text{average duplication} = \frac{\sum_{\text{modules}} (\text{duplication factor}) \cdot (\text{module hits}) \cdot (\text{word size})}{\sum_{\text{modules}} (\text{module hits}) \cdot (\text{word size})}.$$

In this manner, duplication factors are computed at the level of the **HTTIF**, **HTT**, **TP** boards, **PRMs**, and **TFMs** using only the modules necessary for either the 1st or 2nd stage track-finding. Different factors are computed for the **rHTT** (2 GeV tracking) and **gHTT** (1 GeV tracking) systems, as the **gHTT** requires increased duplication in the outer layers to account for tracks with larger curvature. The duplication is more extreme for hardware units corresponding to smaller regions of  $\eta - \phi$  space. The duplication factors are listed in Table 13.11.

## 13.6 Description of the HTT Hardware and Firmware

Table 13.10: *HTT parameters for rate, bandwidth, and internal dataflow calculations.*

Parameter	Value	Comment
Regional detector readout rate	100 kHz equiv.	$10\% \times 1 \text{ MHz}$ : TDAQ TDR, Table 13.9
Global detector readout rate	100 kHz equiv. <sup>a</sup>	TDAQ TDR, Table 13.9.
ITk strip data volume	0.5MB	ITk strip TDR, Table 16.2
ITk pixel data volume	2.4MB	ITk pixel TDR <sup>b</sup>
HTT-region $\eta \times \phi$	$0.2 \times 0.2$	region size used for performance studies
HTTIF coverage $\eta \times \phi$	$2.66 \times \pi/4$	
HTT unit coverage $\eta \times \phi$	$2.66 \times \pi/16$	
SSTP coverage $\eta \times \phi$	$2.66 \times \pi/16$	
TFM coverage $\eta \times \phi$	$1.33 \times \pi/16$	$\sim 6.6 \times (\eta \times \phi)$ HTT-region
AMTP coverage $\eta \times \phi$	$0.44 \times \pi/16$	
PRM coverage $\eta \times \phi$	$0.22 \times \pi/16$	$\sim 1.1 \times (\eta \times \phi)$ HTT-region
Max clusters/strip layer <sup>c</sup> HTT-region (Regional)	200	Fig. 13.10
Max clusters/strip layer <sup>c</sup> HTT-region (Global)	260	Fig. 13.10
Max clusters/pixel layer <sup>c</sup> HTT-region (Barrel L4)	210	Pixel TDR Table 2.7 and Fig. 13.10
Max clusters/pixel layer <sup>c</sup> HTT-region (EC L4))	210	Pixel TDR Table 2.7 and Fig. 13.10
# of roads/HTT-region (Regional)	170	Busiest HTT-region, Table 13.4
# of roads/HTT-region (Global)	270 <sup>d</sup>	Busiest HTT-region, Tables 13.5, 13.6
# of 1 <sup>st</sup> stage fits/HTT-region (Regional)	1500	Busiest HTT-region, Table 13.4
# of 1 <sup>st</sup> stage fits/HTT-region (Global)	2250 <sup>d</sup>	Busiest HTT-region, Tables 13.5, 13.6
# of fit constants/HTT-region (Regional)	90	Busiest HTT-region, Table 13.4
# of fit constants/HTT-region (Global)	140 <sup>d</sup>	Busiest HTT-region, Tables 13.5, 13.6

<sup>a</sup> Combination of full-scan and large coverage tracking.

<sup>b</sup> Calculated from information therein.

<sup>c</sup> Layer defined as a silicon surface giving a cluster with full  $\phi$ -coverage.

<sup>d</sup> Average of jet and minimum bias events.

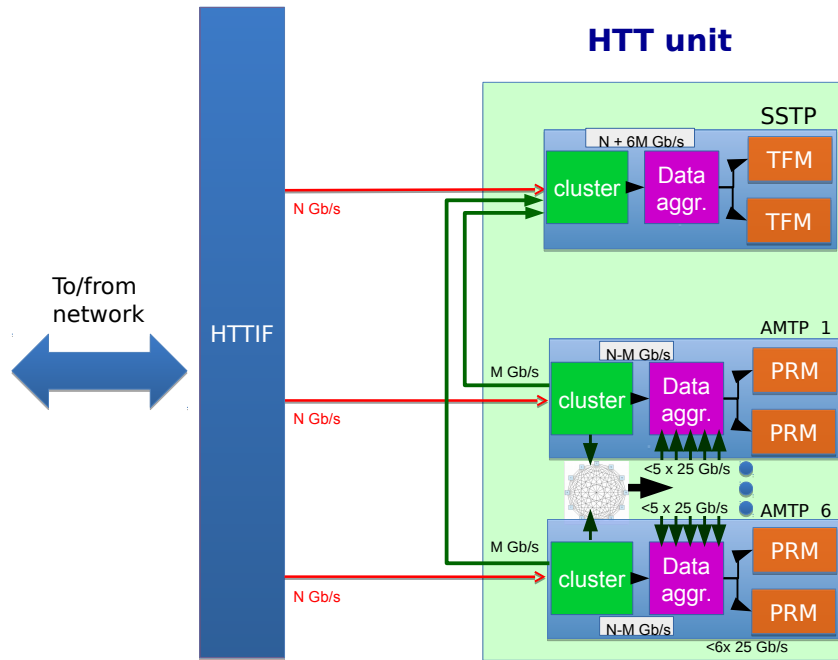


Figure 13.11: Diagram of a *HTT* unit showing data duplication and resulting maximum data flow between different components in the unit. The use of a full mesh backplane for sharing clusters between the *AMTPs* in a unit is indicated with the sketch between the two bottom green boxes.

Table 13.11: Duplication factors used in the bandwidth calculation

Parameter	Value
HTTIF 1st stage w/ 1 GeV tracks	1.4
HTTIF 1st stage w/ 2 GeV tracks	1.3
HTTIF 2nd stage w/ 1 GeV tracks	2.4
HTT Unit 1st stage w/ 1 GeV tracks	2.7
HTT Unit 1st stage w/ 2 GeV tracks	2.2
AMTP w/ 1 GeV tracks	4.4
AMTP w/ 2 GeV tracks	3.4
PRM w/ 1 GeV tracks	6.3
PRM w/ 2 GeV tracks	4.8
SSTP w/ 1 GeV tracks	4.6
TFM w/ 1 GeV tracks	6.4

For the 1<sup>st</sup> stage layers the data rate is expected to be 1.6 Tb/s for the 100 kHz full detector equivalent (10% of 1 MHz) of regional tracking. The same 1.6 Tb/s bandwidth will be used for the 1<sup>st</sup> stage of global tracking at 100 kHz. The 2<sup>nd</sup> stage layers used only for global tracking will use a bandwidth of 1.2 Tb/s. The bandwidth split between 1<sup>st</sup> and 2<sup>nd</sup> stage layers depends on the assignment of layers. Studies show that a baseline assignment of layers that keep data sharing under control for an *HTT* crate is in  $\eta - \phi$   $2.66 \times \pi/16$ . This

## 13.6 Description of the HTT Hardware and Firmware

Table 13.12: *Input data flow in HTT. The payload bandwidth (BW) per unit represents the payload data that the HTTIF receives and forwards per HTT unit.*

	rHTT	gHTT 1 <sup>st</sup> st.	gHTT 2 <sup>nd</sup> st.
event size	1.4 MB	1.4 MB	1.5 MB
event rate	$0.1 \times 1$ MHz	100 kHz	100 kHz
raw data rate	1.1 Tb/s	1.1 Tb/s	1.2 Tb/s
raw data rate w/ board dup.	3.8 Tb/s	4.9 Tb/s	5.6 Tb/s
raw data rate w/ unit dup.	2.4 Tb/s	3.0 Tb/s	5.6 Tb/s
raw data rate w/ HTTIF dup.	1.4 Tb/s	1.6 Tb/s	2.9 Tb/s
HTTIF input BW	60 Gb/s	67 Gb/s	121 Gb/s
unit input BW	25 Gb/s	31 Gb/s	58 Gb/s
board input BW from HTTIF	4 Gb/s	5 Gb/s	58 Gb/s
equalised input BW from HTTIF	4 Gb/s	12 Gb/s	16 Gb/s
input payload BW per HTTIF	248 Gb/s		
input payload BW per unit	112 Gb/s		

assignment is used for the dataflow calculations, neglecting the  $\eta$  dependence. The input dataflow from HTTIF to HTT is summarised in Table 13.12. The equalised input bandwidth from HTTIF is the per board bandwidth after equalising the bandwidth between AMTPs and SSTPs. The last two lines summarise the input bandwidth per HTTIF and per HTT unit. They are the sum of the corresponding values for rHTT, gHTT first stage, gHTT second stage. The HTT shall allow for some flexibility in the assignment of layers to first-stage and second-stage processing. In this regards, part of the bandwidth shown in Table 13.12 might be later reassigned between the first and second stages. As this section describes the HTT modules are limited by internal processing power rather than I/O bandwidth. A bandwidth of 10 Gb/s per link is assumed for all fibre links as well as for ATCA backplane links or on-board links.

**Data buffering** In order to study the buffer occupancy of the off detector HTT hardware an initial discrete event simulation has been implemented including reasonable latency estimates for the processing time of the different hardware components, data sizes and events rates.

From the initial results from this simulation, estimates of the buffer occupancy in terms of events and number of bits have been obtained.

The input rate to the simulation is 200 kHz divided equally between 10% of 1MHz from the regional tracking, and 100 kHz from the global tracking. The 100  $\mu$ s potential input skew has not been included - it is assumed that events will be aligned, so any additional buffering to achieve this has been assumed.

For the input to the Clustering Finder with 75 Kb per event, approximately 40 events will need to be buffered, which includes a factor of two contingency with respect to what is required to achieve containment of 98% of events. This results in a 3 Mb buffer in total.

For the data Aggregator, directly following the Cluster Finder, 10 events will be required, to give a 750 Kb buffer. For the PRM input, 35 events will be required, again with a factor of two contingency, to yield a 2.7Mb buffer. The Duplicate Removal stage would require a 12 event buffer, for a total of 0.9 Mb. In total, the AMTP buffering would need to be approximately 8 Mb in size.

### 13.6.2 Tracking Processor (TP)

The TP is an ATCA board designed to host two FPGA Mezzanine Card (FMC) double mezzanines, which can be either PRM or TFM cards. The main functionalities implemented in the TP hardware are, broadly speaking, off-board I/O, on-board I/O, clustering of raw pixel data, time-alignment of the input data, interfacing to the ATCA crate, and monitoring of the internal firmware block dataflow. Off-board I/O will include receiving of raw detector data, the sharing of data (hits, clusters and tracks) with other TPs via the RTM or the ATCA mesh, transferring tracks to the HTTIF and/or SSTP boards. Data will then need to be time-aligned and exchanged with the on-board mezzanine cards, either to perform pattern recognition and track fitting or second-stage processing. Hits will be mapped onto logical layers in preparation for their processing by the PRM or TFM. The track fitting output will then be fed back for trigger the decision through the HTTIF.

Ancillary functionalities will include interfacing with the ATCA mesh, data switching for efficient sharing of the Pixel clustering results and monitoring of the dataflow at the I/O boundaries RTM, ATCA and FMC and between functional blocks within the TP.

The current preliminary design is based on two symmetric FPGA devices, sharing the input bandwidth and simultaneously connected to all the FMC mezzanines (Fig. 13.12). This design splits the bandwidth requirements among multiple FPGAs running concurrently, and could allow early tests to run on prototypes with a reduced number of FPGAs. Figure 13.13 breaks down these functions in logical blocks to be implemented within the electronics board.

As shown in Fig. 13.12 and 13.13, dataflow monitoring and debugging is included in the design, allowing monitoring and - possibly - injection of data in the various functional blocks to monitor and test them individually.

**Major Specifications** The functionality of the TP will be implemented in two FPGA devices. The two critical resources are the number of I/O pins required for the interface described above (on-board, off-board and monitoring) and the number of logic cells required to implement the data handling and clustering. To quantify these resources, the dataflow requirements summarised in Table 13.10 are used for calculating the dataflow in the TP. The

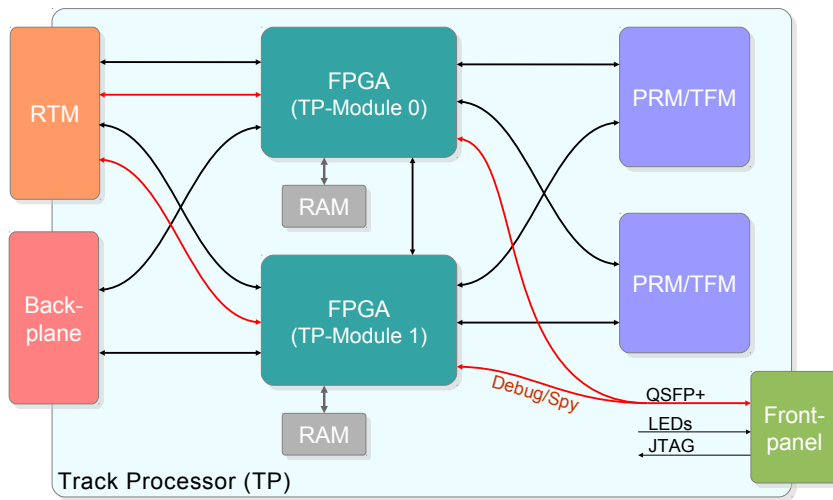


Figure 13.12: Main diagram of the *TP*, illustrating the external interfaces and the dataflow with the on-board *FPGA* devices. The read lines are for monitoring and debugging purposes.



Table 13.13: *AMTP and SSTP FPGA resources required in order to handle the dataflow described in Table 13.10, compared to the resources available in a potential choice of FPGA. It is assumed that each TP board houses two identical FPGAs, with all resources here corresponding to a single FPGA.*

Quantity		AMTP	SSTP	XCKU085	XCKU115
Links	RTM, data input	4	3		
	RTM, cluster I/O	1	6		
	RTM, track I/O	1	6		
	Mezzanine	4	12		
	ATCA Mesh	1×13	1×13		
	FPGA-to-FPGA	4	4		
	Monitoring	2	2		
	Total Links	29	46	56	64
Logic cells (kCells)	Pixel clustering	490	490		
	Switching	70	70		
	Total	560	560	1088	1451
Memory [Mb]	HTTIF input buffer	1.6	2.5		
	Cluster input buffer	0.3	1.9		
	Cluster finder buffer	0.2	0.1		
	Aggregator buffer	0.1	0.2		
	HitWarrior buffer	0.1	0.2		
	Total	2.3	4.8	57	76

assumption is that input data is fed to the system asynchronously, but with sufficiently low time spread within an individual HTT unit consisting of six AMTPs interfaced with one SSTP through shelf-to-shelf optical links. The top block of the table describes the study's assumed input parameters and the bottom block shows the corresponding estimated requirements at the board/unit level. It is assumed that the detector data is duplicated at the input to the system so that all needed data is sent to each HTT unit. Within a HTT unit, in order to balance FPGA resource usage, pixel clustering is shared among the TP boards with the clusters sent to other TPs via the backplane.

The resulting dataflow estimates from Table 13.10 and Fig. 13.11 can be translated into resource requirements for the on-board FPGA logic. The underlying assumptions used are that I/O links will be implemented at a bandwidth of 10 Gb/s and pixel clustering requires ~ 90 FPGA kcells per Gb/s of input to be processed. Table 13.13 summarises the resulting resource requirement per AMTP and SSTP board. Please note that debugging and monitoring features are not accounted for in these estimates.

The resources listed are compatible with the implementation of the logic in two FPGAs. Each FPGA carries out all of the TP functions, but with half of the total data flow.

An early implementation of the AMTP set-up is being studied using pre-prototype hardware based on the Data Formatter board [13.5], and a FMC mezzanine card carrying 12 of

the current generation (v.6) **AM ASICs** (each with 128k patterns) and one Kintex Ultra-scale 060 **FPGA** handling dataflow on the mezzanine, controlling the **AM ASICs** and doing first-stage track fitting. The same **FMC** can also be interfaced to a standard PCI Xilinx **FPGA** evaluation board for **ATCA**-independent tests and performance studies. Figure 13.14 shows these components.

### 13.6.3 Pattern Recognition Mezzanine (PRM)

The **PRM** will contain 12 **AM ASICs** and a large **FPGA**. The **FPGA** will control the **PRM**, prepare the input data to the **AM ASICs**, receive the matched roads, and perform the first-stage fitting. The 12 **AM ASICs** will provide 4.6M patterns per **PRM**.

The **PRM** design will provide a direct connection between the **FPGA** and **AM ASICs**. The **FPGA** will be the controller and interface for the **ASICs**. There will be two groups of 6 **AM ASICs** each. Each group could be controlled independently, receiving different data and possibly working on separate events. For each road found by the **AM ASICs**, the road ID will be converted back to the 8 **SSIDs** by accessing a high bandwidth external **RAM**.

Figure 13.15 shows the processing functions implemented in the **PRM**. The first step in **PRM** processing will be converting **ITk** clusters received from the **AMTP** into **SSIDs** that will be sent to the **AM ASICs** for pattern matching. Next is the **Data Organiser (DO)** function, which is an on-the-fly database that stores clusters allowing their fast retrieval based on the roads found. The **SSIDs** are also sent to the **DO** for use as addresses in storing the clusters. This operation mode is called **DO** write mode. The **DO** uses three sets of on-chip memory for each detector layer: the **Hit List Memory (HLM)**, the **Hit List Pointer (HLP)**, and the **Hit Count Memory (HCM)**. The **HLM** sequentially stores each hit received. The **HLP** stores the **HLM** address of the first hit stored for each **SSID**. The **HCM** stores the number of hits in each **SSID**. All clusters for a given **SSID** will have to arrive sequentially. This function is performed in parallel for the 8 logical layers.

Once all clusters for one event have been received and processed, the loading to the **AM ASICs** will be complete and readout of matched roads can start. Each road IDs will be sent to the **DO** and at the same time be used as an address into the external **RAM** to convert the road ID into a set of 8 **SSIDs** and a sector number. The **SSIDs** are the pointers that the **DO** will use to retrieve the full resolution clusters from memory. For each road found, the **DO** will output in parallel, in a single clock cycle: the road ID, sector ID and one hit per layer. Multiple hits in a layer will require additional clock cycles.<sup>1</sup> This operation mode for the **Data Organiser** is called **DO** read mode.

Both the **DO** and the **AM ASICs** have two-event buffers. While the first event is being processed reading out **AM** roads with the **DO** in read mode, the second (next) event is

---

<sup>1</sup> For a road with at most one hit per layer a single clock cycle will be needed. For a road with 3 hits in the layer with the most hits 3 cycles will be needed.

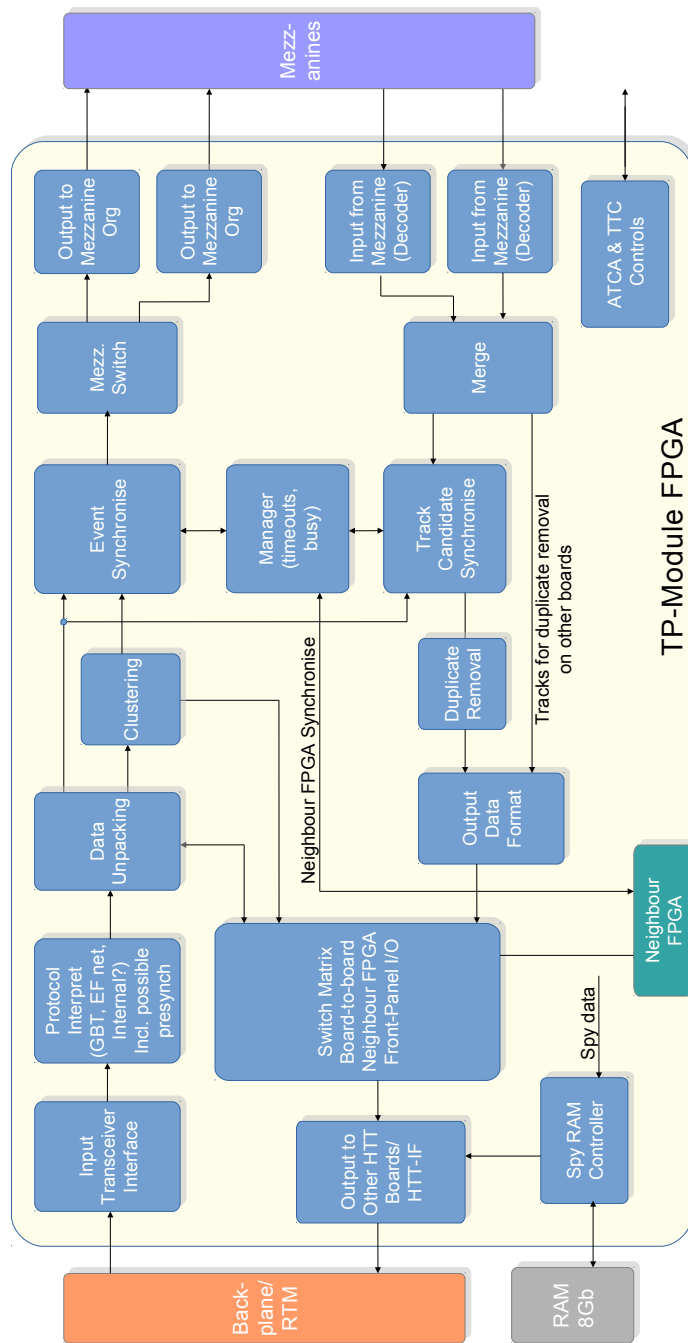


Figure 13.13: Main diagram of the TP, illustrating the major functional blocks to be implemented in the board.

### 13.6 Description of the HTT Hardware and Firmware

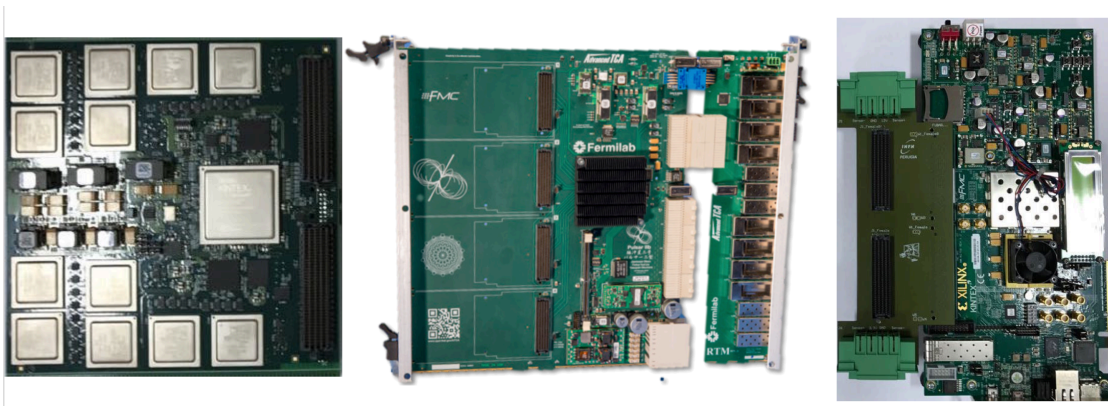


Figure 13.14: Hardware demonstration prototypes based on current-generation devices: *AM mezzanine with Kintex Ultrascale FPGA (left), ATLAS Data Organiser ATCA carrier for FMC mezzanine cards based on a Virtex 7 FPGA (middle), and a Virtex 7 FPGA evaluation board equipped with a mechanical adapter to host the AM mezzanine card.*

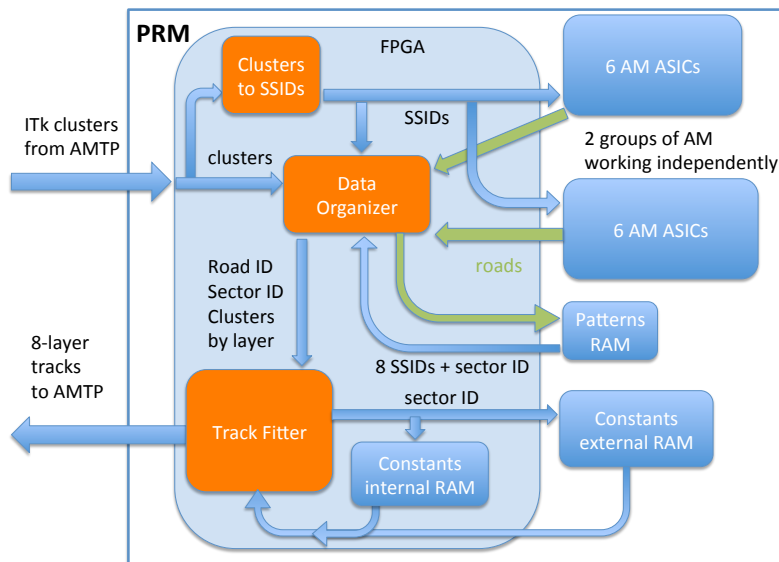


Figure 13.15: Main diagram of the PRM.

loading [SSIDs](#) to the [AM ASICs](#) with the [DO](#) in write mode. The [PRM](#) can proceed to process the third event once the first event has completed read mode and the second event has completed write mode.

The road+hits packet from the [DO](#) will be sent to the Track Fitter module and buffered. For each road+hits packet, the Track Fitter will fit all combinations of hits looping over all combinations of one hit per layer. If one of the layers does not have a hit, then the combinations of one hit on each of the other 7 layers are used. For each combination, the  $\chi^2$  is calculated in linear fits using the hit positions and a set of precalculated fit constants. A  $\chi^2$  cut is applied to reject most fake tracks. For combinations passing the  $\chi^2$  cut 5 helix parameters are calculated again in linear fits using the hit positions and additional constants. For both the  $\chi^2$  and helix parameters, the fit constants are extracted from memory using the sector number as a pointer. The track fitting constants are stored in a combination of the internal memory and in an external [RAM](#) that are accessed in parallel.

Two types of fitters will be implemented. If one of the layers has no hit, then the Majority track fitter is used. One can still use the same set of constants to calculate the  $\chi^2$  and helix parameters provided that there is an estimate of where the hit in the empty layer would have been. The missing hit coordinate is estimated using a linear function of the other hit coordinates. (This equation comes from minimising the  $\chi^2$  over the position of the hit in the empty layer.) If there are hits in all layers, the nominal fitter is used.

The good tracks coming out of the track fitters are sent back to the [AMTP](#). A track packet consists of the hit on each layer, the  $\chi^2$ , the 5 helix parameters, the track road, the sector number, and the hit map indicating which layers had real hits.

The [PRM](#) design aims at maximising the number of stored patterns and the bandwidth to the external memories. The limit is the number of I/O pins in the [FPGA](#). This means that a large [FPGA](#) package is needed.

The two external [RAMs](#) are needed to reading [SSIDs](#) for the found roads, and constants for the fit. The baseline external [RAM](#) device is a reduced latency RAM chip (RLDRAM3) with a 1 Gbit size. The two external [RAMs](#) will use approximately 170 [FPGA](#) pins. The size to store the [SSIDs](#) is 4.6M patterns times 176 bits, which is just above 800 Mb. The size of the fit constants is assumed to be 6k bits, half of which are needed for the  $\chi^2$  calculation and half for the calculation of the helix parameters. With the proposed device 160k sets of track fitter constants can be stored. The 3k bits required for the  $\chi^2$  calculation need to be accessed at high rate. The maximum rate allowed is 30 MHz, which is a good match to our requirements. The internal [RAM](#) can also be used to store fit constants. This can provide additional bandwidth for the access to most frequently used constants.

The main connections between the [FPGAs](#) and each group of 6 [AM ASICs](#) are shown in Fig. 13.16. The input data to the [AM](#) and the majority of control lines will be sent from the [FPGA](#) to the first [AM ASIC](#). Each [AM ASIC](#) will send the same data to the next [AM ASIC](#) until all 6 in a group receive the data. Within each group of 6 [AM ASICs](#), a pair will share

a single output channel to the **FPGA**. The total number of pins used is 127 for each **AM** group, as indicated in Table 13.17. The total number of pins needed for external **RAMs** and **AMs** is 424.

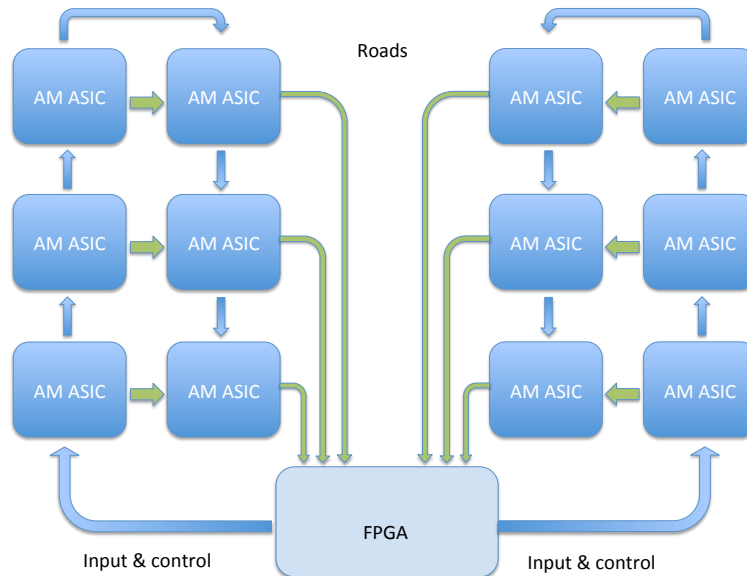


Figure 13.16: Diagram with the main connections between the **PRM FPGA** and the 2 groups of **AM ASICs**.

The **PRM** input bandwidth from the **AMTP** is chosen to be enough to feed both groups of **AM ASICs** at the maximum rate to cover peak **AM ASIC** use for the **L1Track** application. For each group of **AM ASICs** the maximum input bandwidth is 64 Gb/s, which is calculated as 8 layers times 250 MHz, which is the **AM ASIC** clock frequency, times 32 bits word size for one cluster.<sup>2</sup> Given that the two groups of **AM ASICs** will receive shared data, we can allocate 120 Gb/s. For the usage of **HTT** in the baseline **TDAQ** 40 Gb/s between the **TP** and each **PRM** will be sufficient.

Table 13.14 reports rough estimates for the resource usage compared with resources available in a Xilinx KU085 **FPGA** chosen as an example **FPGA** for the **PRM**.

The main parameters for the **PRM** card are reported in Table 13.15. Several parameters are driven from **AM ASIC** performance, as discussed in Section 13.6.4. The main parameters that affect the processing speed are the cluster processing rate, road processing rate, fit constants access, and track fitting rate. These parameters are based on the following as-

<sup>2</sup> The word size for each cluster on the **HTT** modules is assumed to be 32 bits. These will be condensed to 16 bits for **AM** processing. The extra information is used for track fitting.

Table 13.14: *PRM FPGA resource usage compared with a Xilinx KU085.*

resource	required	available	usage %
High speed links	12	48	25%
Logic cells	N/A	1088k	N/A
Block RAM	40 Mb	56 Mb	70%
DSP	1870	4100	50%
IO lines	60 + 400	104+520	80%

Table 13.15: *Main parameters for the PRM card*

parameter	value	comment
Total number of patterns	4.6M	in 2 independent gr. of 2.3M patterns
Peak cluster processing rate	250 MHz	in parallel for the 8 logical layers
Max average cluster processing rate	60 MHz	average over event and layers
Road readout rate	200 MHz	in parallel from each pair of <a href="#">AM ASICs</a>
Road to 8 <a href="#">SSID</a> conversion	500 MHz	from the external RAM
Road processing rate	200 MHz	assuming one cluster per layer. <sup>a</sup>
Fit constants access	30 MHz	from external RAM
Track fitting rate	1 GFit/s	using internal DSP slices

<sup>a</sup> The number of cycles required is equal to the number of clusters found in the busiest layer.

sumptions: up to 12 coordinates are used in the fit<sup>3</sup>, integer multiplications with up to 18 bits for the local coordinate and up to 27 for the constants, and a constant size of 100-200 values of 27 bits. At most 2 missing coordinates are allowed from the sum of missing hits in the event and wildcards in the stored pattern.

Pre-prototypes of the [PRM](#) for R&D have been produced as part of the INFN RDfase2 activity. The last prototype produced was the PRM06 [13.6] that is shown in Fig. 13.17. The PRM06 is a generic R&D prototype for phase-II that uses available AM06 [ASICs](#). Two high capacity and high bandwidth external memories (MT44K32M36) are installed.

The PRM06 is a demonstrator for the first prototype of the [PRM](#) card. The main structure could be similar, while the main change will be to upgrade to [AM08 ASICs](#). The PRM06 has a Xilinx Kintex Ultrascale [FPGA](#). A [FPGA](#) of the same family is a possible choice for the next prototype.

<sup>3</sup> With 12 coordinates possible first-stage layer choices are: 1 pixel and 7 strips layers (using 9 coordinates), 2 pixels and 6 strip layers, or more pixel layers for up to 4 pixel and 4 strip, or 6 pixel layers that could be used in the forward region. Up to 8 pixel layers can be used discarding one coordinate in 4 layers or increasing the size of the constants.



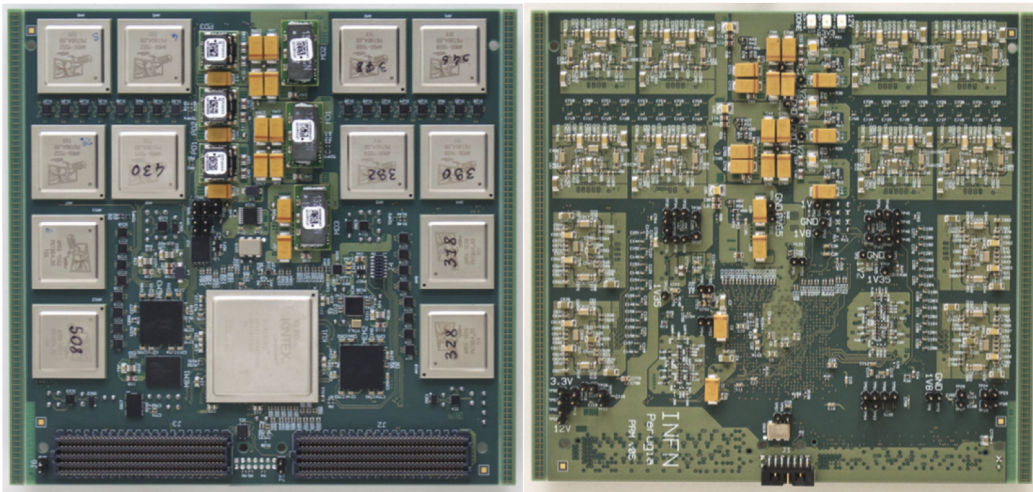


Figure 13.17: Pre-prototypes of the *PRM* card using *FTK* AM06 *ASICs*.

### 13.6.4 Associative Memory (AM) for Phase-II

The core component of the hit pattern recognition will be the *AM ASICs*. This chip is a massively parallel system able to perform bit-wise comparison of incoming data with the pre-stored templates. It has been shown in reference [13.7] that the *AM* does better hit filtering than *FPGA* implementing Hough transform. The production version of the *ASIC* will be called *AM09*. As introduced in previous section, the relevant features for this device are: the number of patterns per chip, the input and output bandwidths, power consumption, and the number of layers, and bits per layer used to describe a pattern.

**Requirements** The Phase-II scenario will require more powerful *AM ASICs*, with respect to the current state of the art *AM ASICs* represented by the *FTK Fast-Track Associative Memory ASIC (AM06)*. The current *AM07* prototype [13.8] (based on 28 nm CMOS process) has been fabricated in August 2017 (Fig. 13.18) and is currently under test. The chip is fully functional and operates up to 250 MHz. LVDS drivers and receivers have been measured with good results up to 1.1 Gb/s. Power dissipation (TX+RX) is about 8 mW at 1 Gb/s. The design for *AM08 ASICs* is already ongoing.

With the aim to achieve the *HTT* requirements, key improvements in the next generation *AM ASICs* will be needed with respect to *AM06*:

- **a higher clock speed and bandwidth:** the next version of the *AM ASIC* will also increase the internal core-matching clock frequency, from 100 MHz to 250 MHz, and the core-readout clock frequency: from from 100 MHz to 200 MHz . With respect to *FTK*, the number of used layers for the pattern recognition are not expected to change hence the same number of layers will be available both in *AM08* and *AM09*. For each



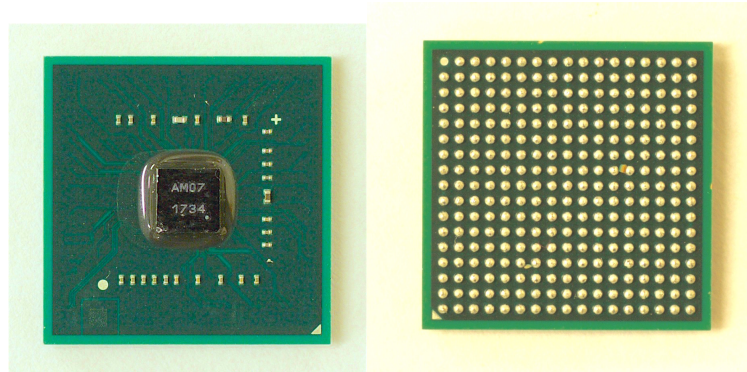


Figure 13.18: *The current prototype of the AM ASIC*

of the 8 layers, 16 bit words can be entered at full clock speed with an input bandwidth of 4 Gb/s per layer. We decided that one patterns is composed by 16 bits  $\times$  8 words. The connection with the FPGA is expected to run at up to 1.6 Gb/s limited by the Low-Voltage Differential Signaling (LVDS) protocol speed on general purpose FPGA pins. LVDS driver and receivers working at 1.8 V will be designed for I/O ports with a slightly higher bandwidth target of 2 Gb/s for usage with future generation FPGAs;

- **a larger density of patterns:** The cell size is expected to be about a factor 3 smaller,  $675 \text{ } \mu\text{m}^2$  to  $186 \text{ } \mu\text{m}^2$  per pattern, allowing to achieve the goal of storing (i.e., 384k patterns) 3 times more patterns in the same area: AM08 will contain 16k patterns and AM09s will contain  $3 \times 128\text{k}$  pattern. The suggested block size is 32k patterns, which requires 12 bits for a  $3 \times 128\text{k}$  patterns ASIC. We plan to allocate 16 bits to cover also the aggressive goal of  $4 \times 128\text{k}$  patterns;
- **a low energy usage per operation:** to achieve this requirement a new CAM cell has been designed: the KOXORAM [13.9]. Simulation of the new CAM cell predicts better energy efficiency: 0.30 fJ/comparison/bit compared to the 0.80 fJ/comparison/bit for the XORAM used in AM06. Other contributions to the AM ASIC power dissipation are to be accounted, resulting in an expected energy per operation for the device of 1 fJ/comparison/bit. A non-core power consumption should not exceed the 500 mW. In similar way, the idle core power consumption should not exceed the 500 mW;
- **a reduction of current peaks:** an implementation of a decoupling scheme in the Power Distribution Network (PDN) will be given to smooth the power peaks that caused minor system issues for the AM06.
- **a reduced latency:** to cope with the Level-1 needs in the evolved scenarios, and to avoid the integration of multigigabit transceivers: The maximum pass-through latency (last hit in AM input, first road at AM output) must not exceed the 100 ns for AM09 and the (50 ns for AM08);
- **a more flexible layer threshold:** Since many application will use the chip, the quorum logic thresholds set will include all combination ( $\{0; 1; 2; 3; 4; 5; 6; 7; 8; 9\}$ );

- **a more flexible blocks enabling:** the new chip will have the possibility to enable/disable blocks of patterns on an event-by-event basis<sup>4</sup> thus saving power since disabled blocks would not make any comparison.

The **HTT** system also requires some specific **AM ASIC** global parameters that does not change with respect to **AM06**:

- based on **28 nm CMOS process**;
- the **junction working temperatures** ranges from 0 °C to 120 °C;
- **AM ASIC** must be **mountable on both sides** of the PCB.

**Work flow** The workflow toward the final **ASIC** is organised in three steps

- **AM08 prototype:** small area Multi Project Wafer (MPW) prototype to silicon proof all the features that will be integrated in **Associative Memory 09 pre-production ASIC (AM09pre)**.
- **AM09pre pre-production:** the full area **ASIC** prototype will be fabricated with a reduced-cost production run<sup>5</sup>. The **AM09pre** will be developed starting from the **AM08** prototype design and extending the memory area, therefore the specification of both versions must be compatible. The timing performance of the final production devices is ensured exploring the transistor performance variability with three “engineering corner runs”.
- **AM09 production:** A second production run<sup>6</sup> will be performed to include in the **AM09** design refinements before going to industrial production.

**AM09 design** In this section we describe the design technique that we choose to adopted to reach the **HTT** requirements:

- **Clock speed, bandwidth, and interface:** Two clocks will be used to synchronise the hit input and output road data: the CLK Hit Fast (CLKHF) running at 1 GHz, and the CLK Readout Fast (CLKRF) running at 800 MHz. The read-out latency could be further optimised exploiting a direct connection between the **AM ASICs** and the **FPGA**, performing first stage fitting within the **PRM**. The chip will receive detector hits through 8 input buses (**HIT\_IN**), propagation within the chip to find correspondence with the pre-stored patterns. The results of the elaboration will be delivered through the **ROAD\_OUT** buses. With the aim to simplify the routing a cascading with a daisy chain architecture will be implemented at **PRM** level. For this reason, these data interfaces will be also copied in output on the **HIT\_OUT** buses and in input on the **ROAD\_IN** buses. Synchronicity will be obtained thanks to: CLKHF for the **HIT\_IN** buses; CLKRF for all **ROAD\_IN** buses; CLKHF\_OUT for the **HIT\_OUT** buses;

---

<sup>4</sup> Implementation using a specific k-word at the beginning of the event

<sup>5</sup> Cost will be limited employing a **Multi-Layer Masks (MLM)** full-mask set pilot run

<sup>6</sup> still based on a **MLM** full-mask set pilot run

Table 13.16: Frequency and serialisation capabilities for AM08 and AM09 signals

SYMBOL	BITS / WORD	SERIAL. FACT.	WORK. FREQ. (MHz)	BW (Gbps)	WORDS	AURORA?
<b>ADDRESS</b>						
ADDR_IN	20	1:8	125	1	1	Yes
<b>DATA</b>						
HIT_IN	16	1:4	250	2 (DDR)	8	Yes
HIT_OUT	16	4:1	250	2 (DDR)	8	Yes
ROAD_OUT	32	8:1	200	1.6	4	Yes
ROAD_IN	32	1:8	200	1.6	4	Yes
<b>CONTROLS</b>						
CTRL_IN	16	1:8	125	1	1	
CTRL_OUT	16	8:1	125	1	1	
<b>CLOCKS</b>						
CLKHF			1000			
CLKHF_OUT			1000			
CLKRF			800			
CLKRF_OUT			800			

CLKRF\_OUT for all ROAD\_OUT buses. The ADDR\_IN buses will be used to give in input the write/read address of memory bank and/or control registers. The chip will be essentially controlled by means of different channels: 1) a fast and sync control technique; 2) an adhoc fast control mechanism; 3) a slow control commercial protocol (i.e, Serial Peripheral Interface – SPI). Tables from 13.16 to 13.17 summarise the interface specifications for the AM08 and AM09.

## 13.6 Description of the HTT Hardware and Firmware

Table 13.17: Pin count for *AM08* and *AM09* signals and connection via FPGAs

SYMBOL	AM08	AM09	FPGA lines		
	#bumps (#LVDS pairs)	#bumps (#LVDS pairs)	1-AM09	6-AM09	HP/HR
<b>ADDRESS</b>					
ADDR_IN	8(4)	8(4)	8	8	HP
<b>DATA</b>					
HIT_IN	64(32)	64(32)	64	64	HP
HIT_OUT	64(32)	64(32)	0	0	
ROAD_OUT	32(16)	32(16)	8	24	HP
ROAD_IN	32(16)	32(16)	0	0	
<b>CONTROLS</b>					
CTRL_IN	4(2)	4(2)	4	0	
CTRL_OUT	4(2)	4(2)	4	0	
SPI	4	4	4	4	HR
MODE	6(3)	6(3)	6	6	HP
HOLD_IN	4	4	1	3	HR
HOLD_OUT	3	3	0	0	HR
<b>CLOCKS</b>					
CLKHF	2(1)	2(1)	0	2	HP
CLKH_OUT	2(1)	2(1)	2	0	
CLKRF	2(1)	2(1)	0	2	HP
CLKR_OUT	2(1)	2(1)	2	2	HP
<b>TOTAL</b>	<b>265/268</b>	<b>265</b>	<b>103</b>	<b>127</b>	

- **Larger density of patterns:** The chips will be organised into cores: an independent unit able to perform parallel pattern recognition tasks. There are three different “core” concepts in the device.
  - **Independent-output-cores:** the unit which contains 1/4 of total patterns. For the AM09, it contains  $3 \times 32k$  for baseline ( $4 \times 32k$  for aggressive). For the AM08, it contains 4k patterns.
  - **Enable-core size:** the unit of pattern-match enable/disable. It contains 32k (4k) patterns for the AM09 (AM08).
  - **Design-core:** the design unit with 4k patterns.

All cores have to be compatible with the new I/O scheme. AM08 will be composed of four “design-cores”: The Table 13.18 shows the order of priority in term of designing efforts for AM08:

Table 13.18: Core features and options

Feature for each Core	1	2	3	4
<b>Control technology</b>				
Variable row number (typ = 64 rows: optimisation of skews vs # row number)				x
SRAM-like reading			x	x
Quorum fullcustom		x	x	x

The core 2 is the baseline and the core 1 will be placed to reduce risks if Quorum fullcustom does not work. The core 3 and 4 will be used to test the new feature of SRAM-like reading. If this feature will not consume more than 5% w.r.t. core 1 and 2 will be integrated in AM09. The core 4 will implement the variable row number architecture. In AM07 and AM06 we used block of 64 rows. However, this is not clear that is the optimal solution. For this reason, different blocks with several rows number will be placed to find the optimal solution in terms of power, timing and area. As HTT impose strict requirements in terms of pattern density and accurate study on the area budget has been performed. The reduced-cost run have an area limit of  $154 \mu m$ , an estimation of area for the AM09 has been given. Fig. 13.19 shows the pie chart. Estimation has been obtain with core 2.

An open point of AM07 design was that the Quorum circuits (previously called "majority") occupy more silicon area than memory cells because they are based on standard cells. We estimate that by developing a full custom design, a factor  $\times 5$  will be saved in area devoted to Quorum logic.

Unlike previous versions, AM08 and AM09 allow the possibility to fast read the bit content of CAM. The approach mimics the SRAM architecture and it a nice feature to fast debug the system.

- **Low energy usage per operation:** At the same time, we see in previous requirement that a lower energy usage per operation with respect to AM06 have to be obtained. For this reason, an extrapolation study for AM09 power consumption has been given.

## 13.6 Description of the HTT Hardware and Firmware

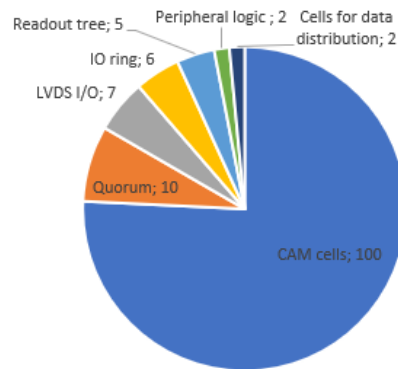


Figure 13.19: *Area Budget for the AM09*

The **AM ASICs** are low power devices that use very little energy per operation. The energy per operation is in the scale of few femto-Joule. For comparing a 16 bit word with one pattern the **AM09** will use 16 fj.

More **AM** devices can be placed with high density on a relatively small printed circuit area, thanks to their low power. The unit cost at production of an **ASIC** is relatively low, with respect to medium or large **FPGAs**. In this context, the power consumption becomes a key parameter because it determines the maximum number of **AM09**, and thus the maximum **AM** processing capability, for a given printed circuit area.

The **AM08** and **AM09** will operate from a single core power supply of 1.0 V. The IO voltage will be 1.8 V with all data inputs and outputs fully compatible with LVDS18. The power distribution must keep voltage at transistors within the simulated corners (Table 13.19). Half of the allowed ripple from the simulation corners is allocated for power distribution on the **PRM** card, the other half for the distribution in the package and die.

Table 13.19: *Voltage settings for AM ASIC.*

	sim. corners	nom. V	max ripple <sup>a</sup>
Digital core power (VDD)	0.9 V – 1.1 V	1.0 V	±50 mV
Digital I/O power (VDDPST)	1.62 V – 1.98 V	1.8 V	±90 mV
Analog core power (VDDA)	0.9 V – 1.1 V	1.0 V	±50 mV
Analog I/O power (VDDAPST)	1.62 V – 1.98 V	1.8 V	±90 mV

<sup>a</sup> The max ripple at balls is the required voltage stability on the PCB.

The **AM09** power consumption consists of a baseline component of 1 W per **ASIC** (the sum of I/O power and baseline internal power), plus a data-driven contribution equal to rate of bits compared times 1 fj/comparison/bit. The value of 1 fj/comparison/bit assumes an average bit flip rate of 50%. The last term is proportional to the average rate at which data words are presented at the **AM09** inputs. For example comparing

16 bits words with an average 50% bit-flip for data sent at an average of 50 MHz on all 8 buses requires a power of 2.5 W. This means that the data-driven power consumption is 0.05 W/MHz times the average word input rate. The formula for total AM09 power is:

$$P(\text{AM09}) = 1W + \langle \text{inputrate} \rangle * 0.05W/\text{MHz} \quad (13.3)$$

Since these chips contribute to the major part of HTT power consumption, an accurate dimensioning of power system can be achieved with chip information.

To ensure the bandwidth and power consumption requirements, the chip will be packaged in a Ball Grid Array (BGA). The AM08 will be packaged in a 17×17 BGA package (289 pins), and the AM09 will be packaged in a 23×23 BGA package (529 pins).

- **a reduction of current peaks:** In order to make the power usage more uniform in time, DCO will be used to implement a digital Phase-Locked Loop (PLL) able to spread in time the commutations of internal signals through a whole clock period. In order to implement this, different block of Associative Memory logic will be driven by clocks with different phase. The change between clock domains will happen between two near by registers with a very short propagation time. The DCO is a fully-digital oscillator module capable of producing clock signals with frequencies ranging from 0.8 GHz to 4 GHz in most typical cases. To be functional, DCO requires a current reference in order to obtain an output current with a low sensitivity against Power Supply and Temperature. The nominal output current is 1 μA with a nominal variation close to ±50 nA in a temperature range from 0°C to 120°C.
- **A controlled junction temperature:** A temperature sensor will be designed. It will be based on the measurements of the base-emitter voltage in the diode-connected bipolar from the current reference in a range from 0°C to 130 °C with a resolution of 2 °C. It will require a calibration procedure along the process variation. The analog to digital conversion is based on the time-to-digital technique. It is based on a current reference respectively based on a current mode architecture, in order to obtain an output current with a low sensitivity against power supply and temperature. The nominal output current is 1 μA with a nominal variation close to ±50 nA in a temperature range from 0°C to 120°C.

### 13.6.5 Track Fitter Mezzanine (TFM)

The Track Fitter Mezzanine, which sits on the SSTP main board, finds tracks on all 13 ITk logical layers and calculates the 5 track helix parameters and the  $\chi^2$  of the fit. The functional diagram of the card is shown in Fig. 13.20

There will be 192 TFMs in the system, each containing two mid-range FPGAs. The inputs to the board, 11 serial links running at 10 Gb/s, come through the TFM-SSTP mezzanine connector. One link carries the 8-layer tracks found by the 6 PRMs that feed one TFM.



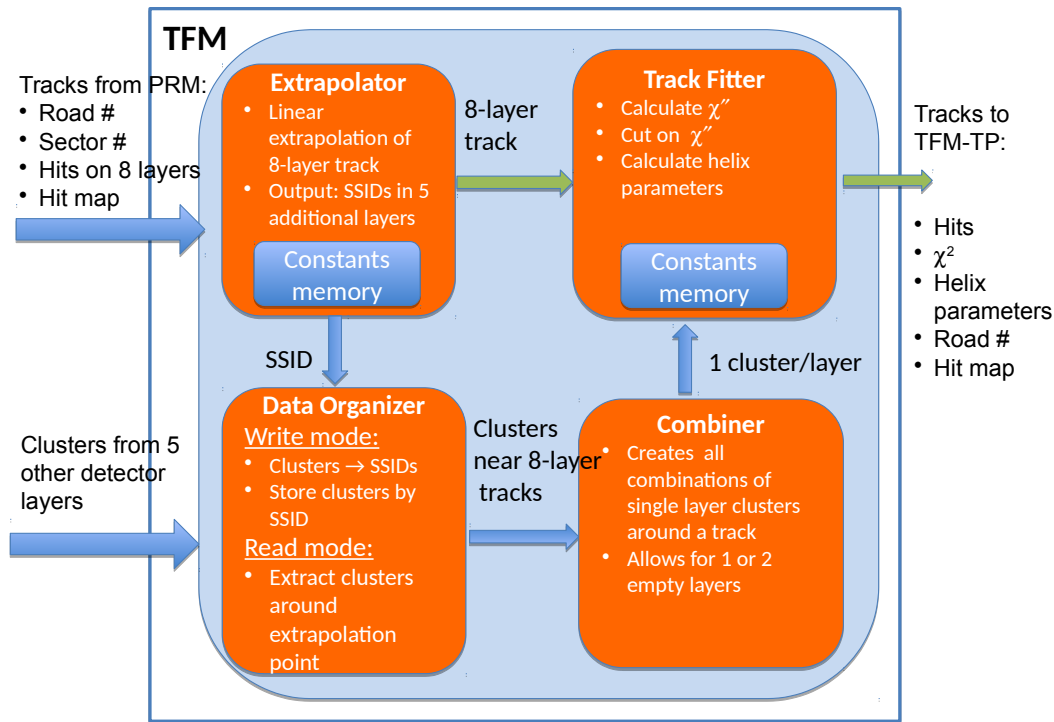


Figure 13.20: The functional diagram of the TFM.

The track data consist of the clusters on the track, the pattern recognition road number, the sector number used to look up fitting constants, and the cluster map which indicates which of the 8 layers had a cluster. There are 10 serial links, 2 each carrying the clusters from one of the detector layers not used in the PRM.

The first step in TFM processing is the Extrapolator, which finds the clusters on the 5 other detector layers that are close to a PRM 8-layer track. The incoming clusters enter the DO. The 5 layers are processed in parallel, with the clusters on each sorted so that those in the same SSID are forwarded sequentially. The DO uses three sets of on-chip memory for each detector layer: the HLM, the HLP, and the HCM. The HLM sequentially stores each cluster received. The HLP stores the HLM address of the first cluster stored for each SSID. The HCM stores the number of clusters in each SSID.

For each 8-layer track, the Extrapolator calculates the most likely cluster position in each of the other 5 layers. A matrix calculation uses the 9 track cluster coordinates and a set of 10 constants per layer extracted from external memory. The most likely cluster positions are converted to SSID numbers, a check is made that there are clusters in at least 4 of the 5 layers, and if so the clusters are extracted.



The track candidates from the Extrapolator are sent to the Track Fitter. A candidate consists of an 8-layer track with its clusters and all of the nearby clusters on each of the other 5 layers. All of the combinations of one cluster per layer on the 5 layers are fit. If one of the layers does not have a cluster, then the combinations of one cluster on each of the other 4 layers are used. For each combination, the  $\chi^2$  and the 5 helix parameters are calculated in linear fits using the cluster positions and a set of constants extracted from external memory. A  $\chi^2$  cut is applied to reject most fake tracks.

Three types of fitters will be implemented. If one of the 5 non-PRM detector layers has no cluster, then the Majority track fitter is used. The same set of constants as is used when there are no missing hits can be used here to calculate the  $\chi^2$  and helix parameters provided there is an estimate of where the cluster in the empty layer would have been. The missing cluster coordinate is estimated using a linear function of the other cluster coordinates. (This equation comes from minimising the  $\chi^2$  over the position of the cluster in the empty layer.) If there are clusters in all 5 layers, the Nominal fitter is used. If the  $\chi^2$  for a nominal fit fails, it is possible that one of the layers contains a random cluster. The Recovery fitter sequentially drops the cluster on one of the 5 layers and refits the remaining clusters as a Majority fit. If at least one of the recovery fits passes the  $\chi^2$  cut, the track with the lowest  $\chi^2$  is passed on.

The good tracks coming out of the track fitters are sent back to the SSTP. A track packet consists of the cluster on each layer, the  $\chi^2$ , the 5 helix parameters, the track road, the sector number, and the cluster map indicating which layers had real clusters.

The tentative choice is two moderate priced Xilinx FPGAs per TFM (XCKU085) along with high speed DDR3 memory chips (Micron MT44K16M36). This is driven by the processing requirements for the board. Simulation indicates that the PRMs will send 12 8-layer tracks per event to a TFM for a minimum  $p_T$  of 4 GeV. Scaling with curvature gives an expected 48 tracks per event per TFM for  $p_T$  above 1 GeV, or 24 tracks per event per TFM FPGA. With the 100 kHz event rate in gHTT, that corresponds to 2.4 Mtracks/s, or 1 track per 400 ns. There will be 1 or 2 connections (sets of extrapolation constants) per track, with each set containing 100 constants. The chosen FPGA and memory allow for this rate of memory data transfer.

The expected occupancies in the 5 layers being added to the 8-layer track indicate that there will be on average 5.4 combinations of hits in those layers to be fit with the hits from a PRM track. This assumes an extrapolation window of  $\pm 3\sigma$  in the 8-layer track angular resolution. The result is that a fit must be done on average in 75 ns. To obtain the  $\chi^2$  and helix parameters, 18 calculations must be done, each consisting of 19 multiplies and adds. Each of those calculations would be done in a different set of DSPs and as a result they can be completed within the allowed time. One set of 342 fitting constants has to be loaded for each connection. With the number of memory chips that an FPGA can support, the constants can be loaded within the available time. An extrapolation of the logic resources used in the FTK Second-Stage Board, which carries out similar calculations, shows that the

Table 13.20: *TFM FPGA resource usage compared with a Xilinx KU085.*

	Needed per event	Capability per event
High-speed links/FPGA	12	44
Logic cells/FPGA	0.4M	0.8M
DSPs/FPGA	2500	4100
I/O lines/FPGA	260	676

chosen [FPGAs](#) have the needed logic capacity. An example of resources needed and the capability available in the example [FPGA](#) is given in Table 13.20.

### 13.6.6 The HTT Interface (HTTIF) Infrastructure

The [HTT](#) is connected to a commodity network through a [HTT](#) interface ([HTTIF](#)). The [HTTIF](#) will translate the network-specific data into a lower-level protocol to distribute the hit data to individual [AMTP](#) cards. The system will use QSFP links links giving a 10 Gb/s connection between the [HTTIF](#) and the [AMTP](#) card. The hardware used for the [HTTIF](#) will be based on [FELIX](#) with dedicate [HTTIF](#) firmware. A single [HTTIF](#) server is connected to a group of [AMTP](#) and [SSTP](#) cards. The connection to an [ATCA](#) card is through a rear transmission module ([RTM](#)). The [RTM](#) is passive with mostly point-to-point connections. Utilisation of [RTMs](#) will allow us to reduce the mechanical stress on the fibre-optical links and the number of insertions. An additional advantage is housing the fibres in the rear of the rack where room is normally available for this purpose.

The bandwidth may depend on the pseudo rapidity of the trigger tower and on the needed data duplication. Data duplication is needed because of track bending in the magnetic field and the size of the luminous region (beam spot). To reduce the data duplication in the [HTTIF](#) the [ATCA](#) backplane is used to share strip and pixel clusters between [AMTP](#) cards residing in the same shelf. The second-stage track processor boards [SSTP](#) will receive track data from [AMTPs](#) and hit data directly from [HTTIF](#). All the reconstructed tracks are sent back from the track processor boards to the [HTTIF](#).

The [HTTIF](#) server will process data requests received from the [EFPU](#) over the network. Each request will be received as a single network message in order to have all relevant [ITk](#) data arrive at the same time. Each [HTTIF](#) server is equipped with two [FELIX](#) I/O cards. For each tracking request, data with a granularity of one detector element will be routed to one or both [FELIX](#) cards and then to the [ATCA](#) cards. The [HTTIF](#) will send out data to the cards of a given [HTT](#) unit with a maximum skew of 100 microseconds.

Table 13.21: *HTT* 1<sup>st</sup>-stage dataflow summary. The *gHTT* output bandwidth and event size refer to the *HTT* internal data flow between *AMTP* cards and *SSTP* cards.

	rHTT/event	gHTT/event	HTT rate	available
# Cluster /PRM (layer average)	200	260	46 MHz	60 MHz
# Roads/PRM	170	270	45 MHz	400 MHz
# Constants read/PRM	90	140	23 MHz	30 MHz
Fits/PRM	1500	2250	400 MHz	1 GHz
Tracks after $\chi^2$ /PRM	80	280	36 MHz	
Tracks after HitWarrior/AMTP	10	35	4.5 MHz	
rHTT output bandwidth /AMTP	640 Mb/s			
<Tracks after HitWarrior>/AMTP	7	20	2.6 MHz	
Total output bandwidth	250 Gb/s	750 Gb/s	1 Tb/s	
Average event size	30kB	900kB		

### 13.6.7 Dataflow summary

The internal and output dataflow for 1<sup>st</sup>-stage processing is summarised in Table 13.21. The required rates are based on studies in Section 13.5 and the *HTT* rate is the sum of *rHTT* and *gHTT* rates. For *rHTT* the values are from the jets samples with  $p_T > 2$  GeV. For *gHTT* the values are the average of the jets and minimum bias samples with  $p_T > 1$  GeV. For both cases the worst case  $\eta$  region is considered and the numbers are scaled up by a factor of 1.1, because each *PRM* covers a size  $\sim 1.1$  times larger than the region size used for the performance studies as indicated in Table 13.10.

The output bandwidth numbers assume 160 bits for track parameters and track identifiers, such as road ID and sector ID, plus 480 bits for clusters on tracks for a total of 640 bits/track. The lines “Tracks after  $\chi^2$ ” and “Tracks after HitWarrior” refer to the worst case  $\eta$  regions. The line “<Tracks after HitWarrior>” refers to the average number of tracks over all  $\eta$  regions. The output bandwidth and event size are based on this average.

Table 13.21 reports an event size of 30 kB for *rHTT* and an average event size of 250 kB for *gHTT*. The *gHTT* average event size naively accounts for the fact that only 100 kHz of the 400 kHz *EF* pre-selected events will undergo global tracking. Table 13.22 summarizes the estimated dataflow for second stage fitting.

Figure 13.21 summarises the payload data bandwidth for 1/8 of *HTT*.

### 13.6.8 Size and power consumption of the HTT system

Overall, the *HTT* system consists of 576 *AMTPs*, 1152 *PRMs* each with 12 *AM ASICs* (13824 total), 96 *SSTPs*, 192 *TFMs*, and 24 *HTTIFs*. The *AMTPs* are located in 48 *ATCA* shelves, and the *SSTPs* are in 8 *ATCA* shelves. Each group of 6 *AMTP* shelves will preferentially

Table 13.22: *HTT 2<sup>nd</sup>-stage dataflow summary*

	Needed/event	Capability/event
# of clusters/ <b>TFM</b> (max layer)	580	5000
# of clusters/ <b>TFM</b> (average)	380	5000
Cluster rate/ <b>TFM</b> (max layer)	58 MHz	500 MHz
Cluster rate/ <b>TFM</b> (average)	38 MHz	500 MHz
# of 1 <sup>st</sup> -stage tracks/ <b>TFM</b>	60	120
# of constant sets read/ <b>TFM</b>		
Extrapolator	240	500
Fitter	125	230
Fits/ <b>TFM</b>	325	600
<Tracks after HitWarrior>/ <b>SSTP</b>	20	
Total output bandwidth	150 Gb/s	
Average event size	200kB	

Table 13.23: *Size of the HTT system*

	HTT
Number of <b>HTTIF</b>	24
Number of <b>ATCA</b> shelves for <b>AMTPs</b>	48
Number of <b>AM/SSTP</b> boards per shelf	12
Total number of <b>AMTPs</b>	576
Number of <b>PRMs</b> per <b>AMTP</b>	2
Total number of <b>PRMs</b>	1152
Number of <b>AM ASICs</b> per <b>PRM</b>	12
Total number of <b>AM ASICs</b>	13824
Number of <b>ATCA</b> shelves for <b>SSTPs</b>	8
Total number of <b>SSTPs</b>	96
Number of <b>TFMs</b> per <b>SSTP</b>	2
Total number of <b>TFMs</b>	192

be located near a shelf of **SSTPs** and the corresponding 6 **HTTIFs** to ease fibre routing. The number of **HTTIF** is estimated assuming 115 Gb/s payload bandwidth for each **HTT** unit and 250 Gb/s for each **HTTIF**, see Table 13.12.

The system size is summarised in Table 13.23. The **rHTT** and **gHTT** functions are performed by a single overall **HTT** system that will respond to both type of requests in a transparent way, i.e. processing regional or global data.

The system physical size is determined by processing power density per module. The modules are expected to make a balanced use of **Printed Circuit Board (PCB)** area and power dissipation to optimise processing density per module. An initial power budget is shown in Table 13.24. The power allocation for **FPGA**, DC/DC converters and other components is estimated from similar cards. The power for the **AM ASICs** is directly proportional to the

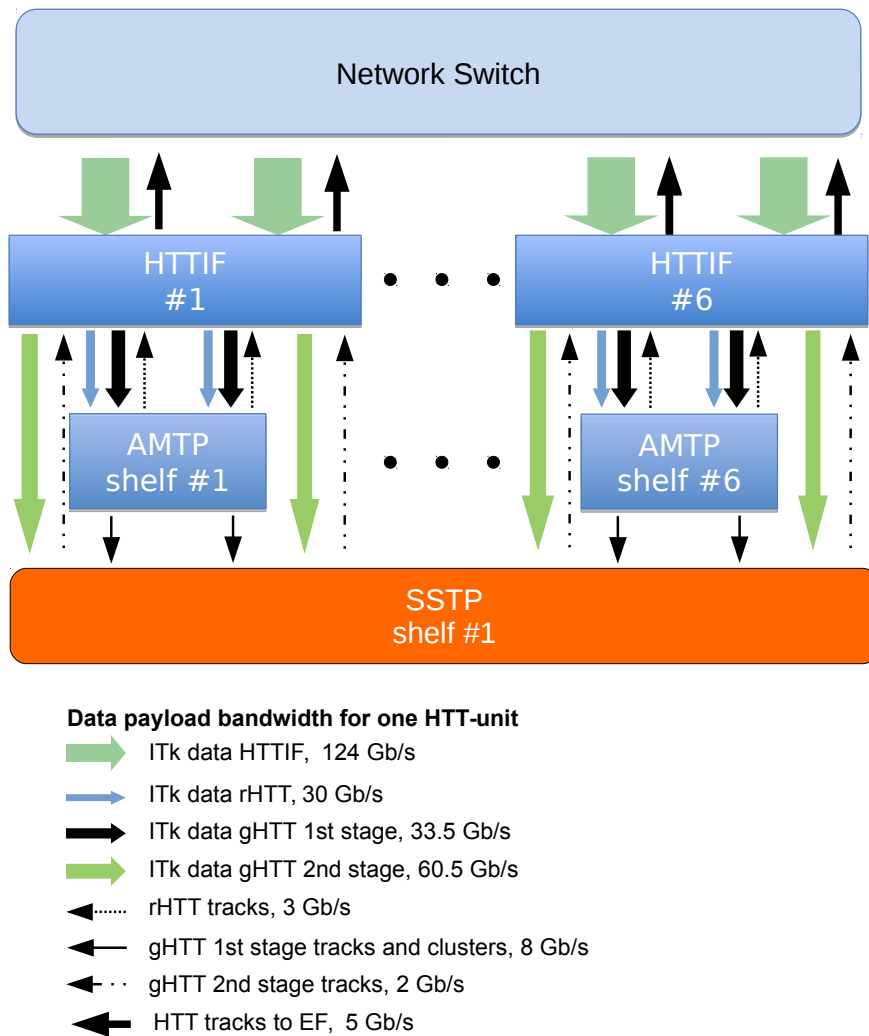


Figure 13.21: A summary of data flow in 1/8 of HTT. The blue and pink boxes are ATCA crates with AMTP or SSTP boards. The numbers are for 50% link utilisation

amount of data sent for processing, as taken from simulation, and to the average number of bit flips among consecutive data words. The power budget is calculated with a AM ASIC energy per operation of 1 fJ per comparison per bit plus a baseline of 1 W per ASIC at an average cluster processing rate of 60 MHz, see Equation 13.1.

## 13.7 Project Milestones

Table 13.24: *Summary of the estimated power usage*

AMTP main card (including DC/DCs)	100 W
PRM FPGA	30 W
PRM 12 AM ASIC	50 W
PRM others (RAMs, IO fanout, DC/DC)	25 W
Total/AMTP	310 W
SSTP main card (including DC/DCs)	100 W
TFM 2 x FPGA	75 W
TFM others (RAMs, IO fanout, DC/DC)	25 W
Total/SSTP	300 W

Table 13.25: *HTT project high-level milestones. Milestones are characterised as software (SW), firmware (FW), or hardware (HW).*

Date	Type	Milestone
Q2/18	FW	Update of FPGA resource usage tables from TDR for SSR
Q2/18	HW	System Specification Review (SSR)
Q4/18	HW	AM08 ASIC submission
Q2/19	SW	Baseline processing configuration defined (including post-TDR optimisation)
Q2/19	HW	Preliminary Design Review
Q3/20	FW	Firmware that explores one by one all critical functions of main FPGAs
Q4/19	HW	AM09pre ASIC submission
Q4/20	HW	AM09 ASIC submission
Q3/21	SW	Offline SW providing online SW with configuration data loaded on the slice
Q4/21	FW	Slice processing demonstrated on prototypes
Q1/22	HW	Final Design Review
Q4/22	SW	Large scale Monte Carlo production demonstrated
Q1/23	FW	HTT processing and EF-HTT integration demonstrated with 2 HTT units
Q2/23	HW	Production Readiness Review
Q2/25	HW	Hardware installation

## 13.7 Project Milestones

A detailed schedule for the HTT construction project is given in a separate Work Break-down Structure document. A summary of the most important milestones is presented in Table 13.25.

## References

- [13.1] ATLAS Collaboration, *Fast Tracker (FTK) Technical Design Report*, Tech. Rep. CERN-LHCC-2013-007. ATLAS-TDR-021, CERN, Geneva, Jun, 2013.  
<https://cds.cern.ch/record/1552953>.

- 
- [13.2] S. Belforte et al., *The CDF trigger silicon vertex tracker (SVT)*, IEEE Trans.Nucl.Sci. (1995) 860.
- [13.3] J. Adelman et al., *The Silicon Vertex Trigger upgrade at CDF*, NIM (2007) 361.
- [13.4] CDF Collaboration, *The CDF detector: an overview*, [Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment](#) **271** (1988) no. 3, 387 – 403. <http://www.sciencedirect.com/science/article/pii/0168900288902987>.
- [13.5] Okumura Y., Olsen J., Liu T., Yin H., *Prototype performance studies of a Full Mesh ATCA-based General Purpose Data Processing Board*, [Nuclear Science Symposium and Medical Imaging Conference \(NSS/MIC\)](#), 2013 IEEE (2013) .
- [13.6] C. Gentsos et al., *Track finding mezzanine for Level-1 triggering in HL-LHC experiments*, in *2017 6th International Conference on Modern Circuits and Systems Technologies (MOCASST)*, pp. 1–4. May, 2017.
- [13.7] J. Gradin, M. Martensson, and R. Brenner, *Comparison of two hardware-based hit filtering methods for trackers in high-pileup environments*, submitted to JINST . <https://arxiv.org/abs/1709.01034>.
- [13.8] A. Annovi, *A Low-Power and High-Density Associative Memory in 28 nm CMOS Technology*, in *6th International Conference on Modern Circuits and Systems Technologies (MOCASST)*. 2017.
- [13.9] A. Stabile, *Italian patent, 102016000077445 (UA2016A005430) del 22.07.2016 (Ns. rif.: P1637IT00)*, .





## 14 Evolution Scenario

The single hardware-level trigger discussed throughout this document is the selected configuration of the trigger architecture design for the Phase-II [TDAQ](#) System and will be the nominal design for the start of Run 4. However, two main risks have been identified that may impact the eventual performance of the system at the ultimate [HL-LHC](#) luminosity (see [Table 1.2](#)). The first risk is the uncertainty in the projected trigger rates for hadronic objects at  $\langle\mu\rangle = 200$ . The second risk is the uncertainty on the occupancy in the inner pixel detector layers at the ultimate [HL-LHC](#) conditions; should this occupancy greatly exceed expectations, the increase in event size would exceed the bandwidth specification for the pixel detector readout chips. The architecture scheme described in this chapter mitigates both risks by adding a second hardware-level trigger (Level-1) which uses regional tracks built from hits in the [ITk](#) strip detector and the outer layer of the pixel detector. This additional trigger stage would provide for a rudimentary primary vertex selection for multijet triggers, allowing for additional hadronic background rejection and a lower overall readout rate for the inner detector pixel layers. Additional details regarding the motivation and criteria for such an evolution are presented in [Section 14.1](#), followed by the requirements for the evolved system in [Section 14.2](#) and a description of the resulting “evolved” architecture design in [Section 14.3](#). The implications for the individual [TDAQ](#) sub-systems are described in [Section 14.4](#), including the design considerations incorporated into the baseline system design that would allow for evolution and an estimate of the latency of the evolved system. Finally, the opportunities for a variety of physics signatures provided by the inclusion of hardware-based regional tracking in the Level-1 trigger are presented in [Section 14.5](#). A timeline for the eventual commissioning of the evolutionary system is outside the scope of this document.

### 14.1 Criteria for Evolution

The two main criteria for the evolution to the split-level hardware trigger configuration are the hadronic trigger rates and the inner pixel detector layer occupancies. If either or both are higher than expected, the baseline [TDAQ](#) architecture would not survive the ultimate [HL-LHC](#) running conditions. In this case, the hardware of the baseline trigger system is designed so it can be used in the evolved [TDAQ](#) system with adjustments to the firmware as described later in this chapter.

## 14.1 Criteria for Evolution

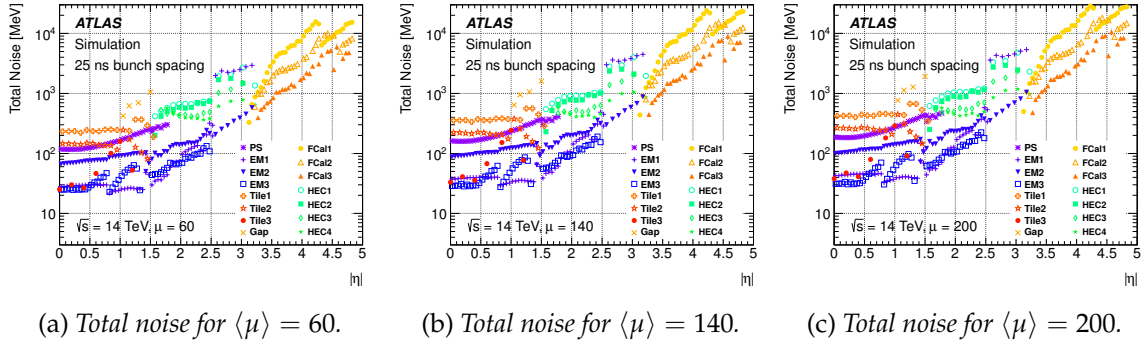


Figure 14.1: Quadratic sum of simulated electronics and pileup noise per calorimeter cell for each calorimeter layer as function of pseudorapidity. Each subfigure corresponds to a different level of pileup noise; the total noise at each  $\mu$  is used in the topological clustering algorithm to reconstruct hadronic events.

### 14.1.1 Uncertainty in Hadronic Trigger Rate Estimates

In the nominal HL-LHC trigger menu for the single-level trigger architecture (see Table 6.4), the trigger rates are based on simulated data with  $\mu = 200$  interactions per crossing. Lepton trigger rate projections have proven reliable. However, projections for hadronic rates are extremely sensitive to the contribution from stochastic jets (accidental overlaps of uncorrelated particles).

Hadronic activity is reconstructed using the topological clustering algorithm described in Section 6.2. The noise setting in the algorithm depends on the expected number of interactions per crossing and varies as a function of calorimeter layer and  $|\eta|$ , as shown in Fig. 14.1. These settings are then used in simulation to form topological clusters that are the inputs to the anti- $k_t$  jet-finding algorithm in order to determine the hadronic trigger rates as a function of  $\mu$ . An illustration of this is the trigger rate per unit  $\mu$  for a hadronic trigger requiring four EM-scale jets above 30 GeV, as shown in Fig. 14.2. Several features can be extracted from this figure. First, adjusting the effective number of interactions per crossing by  $\pm 10$  around each pileup setting results in an increasing variation in the trigger rate. Second, the central values of the rates at  $\mu = 60$ ,  $\mu = 140$ , and  $\mu = 200$  are not linearly increasing. Third, there is an overall uncertainty on these central values for  $\mu = 140$  and  $\mu = 200$  since we cannot yet measure the hadronic trigger rate in data at these high values of pileup. Thus, a reasonably small underestimate of the stochastic jet contribution can result in a large increase in the four-jet trigger rate. Based on the HL-LHC plan, the pileup conditions are expected to reach  $\mu = 140$  at the start of Run 4. Running at the so-called nominal HL-LHC conditions will already give us important feedback as to the accuracy of the simulation in the ultimate HL-LHC conditions. If the hadronic trigger rates at  $\mu = 140$  are higher than predicted by the simulation, a move to the evolved architecture would be needed to reject additional background in the Level-1 hardware trigger.

### 14.1.2 Uncertainty in ITk Pixel Detector Occupancies

The **ITk** pixel detector is made up of five layers in the barrel (layers 0 – 4, where layer 0 is the innermost pixel layer) and consists of five rings for each endcap [14.1].

The occupancy (number of hits per chip) has been studied in simulated  $t\bar{t}$  events at  $\mu = 200$ , as described in Ref. [14.1] and shown for the barrel region in Fig. 14.3. The highest occupancy is expected in Layer 0 of the barrel of the **ITk** pixel detector, where the average number of hits per chip is expected to be 137 in the nominal inclined geometry. The highest average occupancy with respect to  $z$  is 247 hits per chip, which corresponds to a maximum inner pixel channel occupancy of 0.16% and a maximum hit density of  $0.64 \text{ mm}^{-2}$ .

The occupancy drives the readout data rate for the **ITk** pixel detector. The data rate as a function of pixel layer (for both inclined barrel and endcap) is shown in Table 14.1. In the single-level trigger scheme, the innermost pixel layer in the barrel has the highest expected average data rate per link at  $\mu = 200$ : 3.97 Gb/s. The readout is designed for an average data rate of 70% of the maximum bandwidth, or 3.6 Gb/s, to allow for event-by-event fluctuations. The maximum data rate is 5.12 Gb/s per **FE** chip. In the evolved scheme, the front-end chips for the inner barrel and endcap layers incorporate logic to remove hits from memory that do not belong to an event accepted by the Level-0 Trigger (the so-called “fast clear”), bringing the rates well below the design specification. Aggregation causes rates near the soft limit in the outer layers as well in the evolved scheme.

However, if the innermost pixel occupancy is dramatically higher than predicted by simulation at the start of Run 4, the event size would increase and the data rate could exceed the

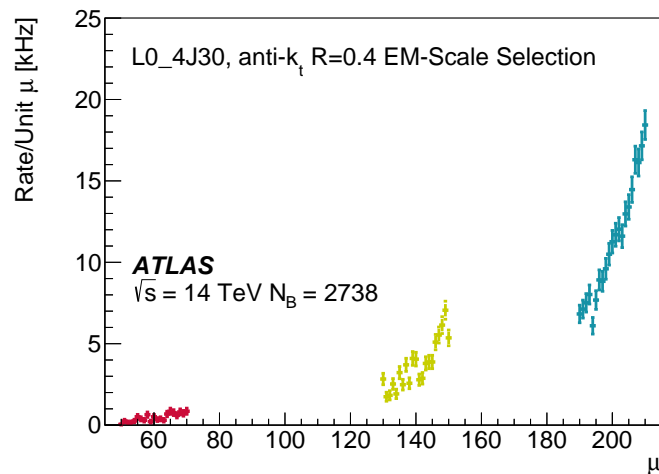
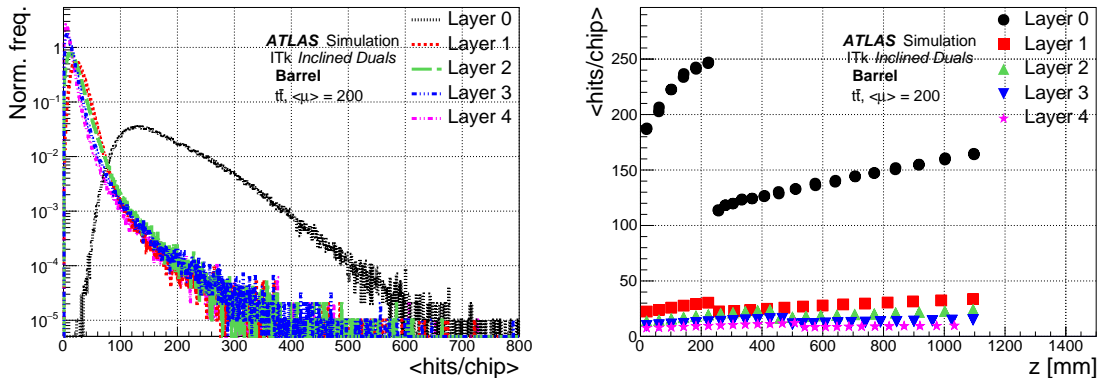


Figure 14.2: The four-jet trigger rate per unit  $\mu$  as a function of  $\mu$ . Three pileup values are used for the noise setting in the topological clustering algorithm:  $\mu = 60$  (red),  $\mu = 140$  (yellow), and  $\mu = 200$  (blue).

## 14.2 Requirements for the Evolved System



(a) Occupancy distribution for the *ITk* barrel pixel detector as a function of layer.

(b) Average occupancy vs.  $z$  for the *ITk* barrel pixel detector as a function of layer.

Figure 14.3: Distribution of chip occupancies over all events (minimum single hit) and layers for the *ITk* pixel barrel in the inclined dual layout.

design value of 5.12 Gb/s at the ultimate HL-LHC luminosity. Furthermore, the development of the 5.12 Gb/s serial output of the pixel detector aggregator chip carries a moderate technical risk [14.1]. In either scenario, a lower readout rate of 600 – 800 kHz (compared to the 1 MHz readout rate for the single-level TDAQ system) or other measures to reduce bandwidth consumption might be needed for the pixel detector. The lower readout rate can be accomplished in the evolved architecture through the additional rejection provided by the Level-1 hardware trigger. In this scheme, the longer latency can be accommodated by the pixel readout system.

## 14.2 Requirements for the Evolved System

The main requirements for the evolved system are as follows:

- Calorimeter and muon data shall be sent to the Level-0 hardware trigger every bunch crossing, as in the baseline system.
- To minimise the amount of effort and avoid problems in the commissioning phase, the Level-0 latency shall be kept to the value of the baseline Level-0-only hardware trigger system (10  $\mu$ s).
- Regions of interest for tracking shall be provided by the [Global Trigger](#) sub-system.
- Upon a LOA, regional *ITk* data (no more than 10% of the detector data volume) shall be provided to the Level-1 hardware-based trigger system, at an event rate of at least 2 MHz. For the *ITk* strip detector, the regional data will be produced by FE electronics then sent out via FELIX, while for the *ITk* pixel detector outer layers (Layers 2-4) and the forward rings, the regional data will be extracted from the full Level-0 data stream

Table 14.1: Expected data rates per link for Pixel (inclined) barrel and endcap layers for the single-level trigger (1 MHz) and evolved (4 MHz) L0A rates. The rates are based on simulation and are averages per event; for each layer the value for the module with the highest data rate is indicated. The maximum data rate per front-end chip is 5.12 Gb/s [14.1].

Layer	Baseline Data Rate Per Link (Gb/s)	Evolved Data Rate Per Link (Gb/s)
Barrel Layer 0	3.97	-
Barrel Layer 1	1.78	-
Barrel Layer 2	2.08	2.08
Barrel Layer 3	1.28	2.56
Barrel Layer 4	0.88	3.52
Endcap Layer 0	2.15	-
Endcap Layer 1	2.14	-
Endcap Layer 2	2.60	2.60
Endcap Layer 3	1.56	3.12
Endcap Layer 4	1.08	4.32

in FELIX, then sent out. Other ITk pixel layers and rings will not be used in the Level-1 decision.

- The HTT hardware shall be designed to minimise the effort required to become part of the real-time trigger path.
- Regional charged tracks with  $p_T > 4$  GeV shall be reconstructed in the Level-1 Trigger system within the Level-1 latency budget.
- A rejection factor of 2.5 (7) shall be provided by the Level-1 Trigger to achieve a L1A rate of 800 kHz (600 kHz).
- Upon a L1A, all of the remaining ITk data shall be read out at a rate of 800 kHz assuming a L0A rate of 2 MHz or 600 kHz assuming a L0A rate of 4 MHz.
- The two inner pixel layers (Layers 0-1) shall be read out to the DAQ system at the L1A rate.
- The total latency of the Level-1 hardware-based trigger system (including the Level-0 latency) shall be less than 30  $\mu$ s (35  $\mu$ s) for the 4 MHz (2 MHz) L0A scheme.

The required data rates for the evolved system are summarised in Table 14.2. The parameters for the 4 MHz scheme are driven by the ITk strip detector readout constraints, while the parameters for the 2 MHz scheme are driven by NSW detector readout constraints. These and other considerations with respect to the readout of the ATLAS detectors are described in the following sub-sections.

## 14.2 Requirements for the Evolved System

---

Table 14.2: *The relevant data rates and Level-1 latencies for the 4 MHz and 2 MHz L0A rate schemes. The regional data fraction is assumed to be 10% of the L0A rate, and the total allowed data rate is 1 MHz. The data rate for the full readout of the ITk detectors is set by the difference between these two quantities.*

	4 MHz L0A Scheme	2 MHz L0A Scheme
Maximum Level-1 Latency	30 $\mu$ s	35 $\mu$ s
Fraction of Regional Data	10%	10%
Effective Regional Readout Rate	400 kHz	200 kHz
Full ITk Detector Readout Rate	600 kHz	800 kHz
Total Data Rate	1 MHz	1 MHz

### 14.2.1 ITk Pixel Detector

The design of the FE chip for the inner pixel layers (0-1) accommodates the reception of both Level-0 and Level-1 signals, and incorporates logic to remove hits from memory that do not belong to an event accepted by the Level-0 Trigger (described as the “Fast Clear” logic in Ref. [14.1]). This design dictates that if the TDAQ L0A latency is within 10  $\mu$ s and the L1A signal arrives less than 35  $\mu$ s later, the hit losses will be < 1%. The total memory in the pixel FE corresponds to a maximum latency of 12.5  $\mu$ s.

The outer pixel layers (2-4) shall be read out at the maximum Level-0 rate of 4 MHz, with the installation of additional cables and fibres<sup>1</sup>. In the case of the outer pixel layers, the selection of data for the Level-1 hardware-based trigger system will be done off-detector, in the Readout System, based on Regional Readout Requests received from the Global Trigger via the corresponding FELIX units.

### 14.2.2 ITk Strip Detector

The ITk strip detector readout shall be able to accommodate a 4 MHz Level-0 rate without any design modifications, provided that the RoIs requested cover at most 10% of the detector. This maximum readout rate is derived from the 128-event depth of the FE buffer. At a 4 (2) MHz Level-0 rate, and 600 kHz (800 kHz) Level-1, Level-1 accept signals shall arrive at the FE chip within 20 (40)  $\mu$ s after the Level-0 signal. Thus the ITk strip detector FE buffer determines the maximum Level-1 latency (30  $\mu$ s) in the 4 MHz scheme.

---

<sup>1</sup> The installation of these additional cables and fibres leads to an increased material budget of the order of a few percent; this additional material will not have a sizeable impact on the detector performance.

### 14.2.3 NSW

The evolved scheme differs from the specifications used to design the NSW system; an in-depth study has been carried out to understand the implications. In the Level-0/Level-1 scenario the NSW would only be read out on L1A, effectively increasing the bandwidth headroom. A Level-1 latency of  $25 \mu\text{s}$  (and potentially up to  $35 \mu\text{s}$ , depending on the dead-time incurred by the NSW front-end electronics) will not lead to any data loss. The dead-time incurred by the NSW front-end electronics determines the maximum latency ( $35 \mu\text{s}$ ) for the 2 MHz L0A scheme.

## 14.3 Overview of the Evolved System Architecture

If the criteria outlined in Section 14.1 are met, the TDAQ architecture will evolve as shown in Fig. 14.4, to a two-level Level-0/Level-1 system with at least 2 MHz and up to 4 MHz L0A rate. The data rates for these two schemes are summarised in Table 14.2.

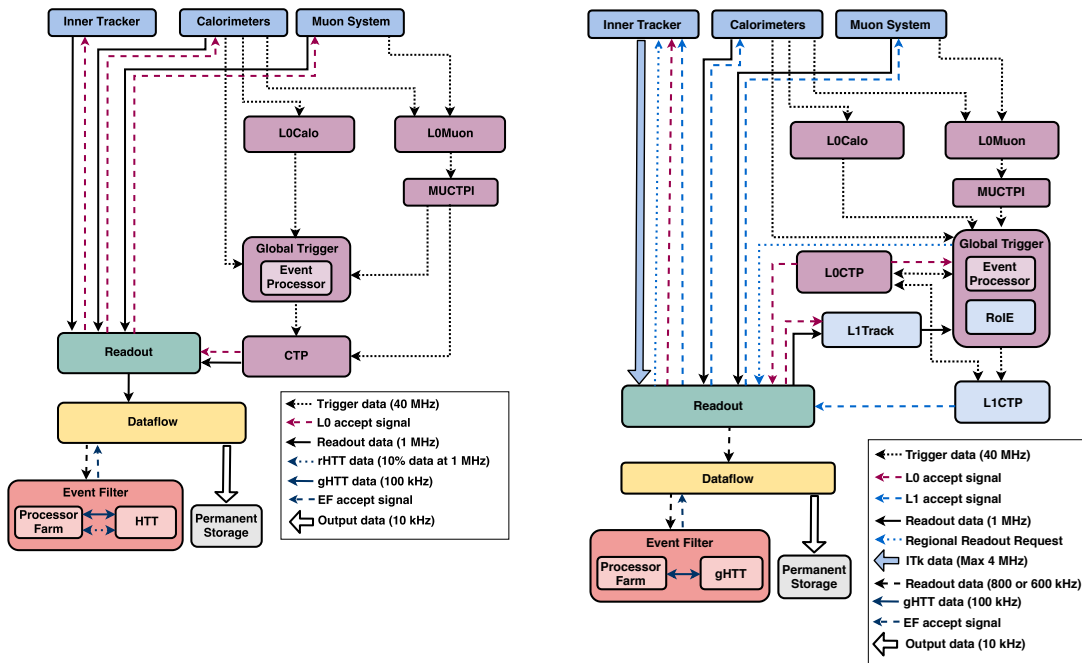
The functional overview of the evolved TDAQ system is shown in Fig. 14.4. Unlike the single-level trigger configuration, the hardware trigger is split into a two-level hardware trigger system, where the HTT performs the primary reduction of the Level-0 rate for an affordable EF farm size. The basic Level-0 Trigger functionality is unchanged. However, a Region of Interest Engine (RoIE) is added to the Global Trigger to calculate RoIs upon a L0A. The RoIs are used to generate R3s for the ITk strip detector and the equivalent off-detector data selection for the outer layers of the ITk pixel detector. In the case of the ITk strip detector, the R3s are distributed to the front-end electronics via FELIX, initiating the transfer of data for use in L1Track. Due to bandwidth limitations on the front-end electronics, the selected data cannot represent more than about 10% of the total detector data on average. In the case of the ITk outer pixel layers, the regional readout requests are used in the Readout System to select the relevant data for L1Track. L1Track is a hardware-based regional tracking system that reconstructs tracks in these RoIs; it is composed of the same hardware and firmware as the HTT components, but is subject to an additional latency constraint of  $10 \mu\text{s}$ .

The Global Trigger receives the regional charged track objects and combines these with the same calorimeter- and muon-based objects as foreseen in the single-level scheme. Additional algorithms to ensure jets in hadronic events arise from the same primary vertex may be implemented. The output is sent to the Level-1 CTP, which is responsible for the final Level-1 trigger decision and for distributing the L1A signal to the detectors to transmit their full event data. The overall latencies are summarised in Table 14.2.

The Level-0-only TDAQ design is capable of operation at 1 MHz and luminosities up to  $\mathcal{L} = 7.5 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ , but all so-called “hooks” necessary to evolve to a Level-0/Level-1 system at a later stage are implemented in the design. The implications for the design of the individual TDAQ sub-systems are described in Section 14.4.



### 14.3 Overview of the Evolved System Architecture



(a) Baseline *TDAQ* System with a single-level hardware trigger. (b) Evolved *TDAQ* System with a two-level hardware trigger.

Figure 14.4: Diagram of the evolved *TDAQ* System in Phase-II with a two-level hardware trigger (b), compared to the single-level hardware trigger configuration (a). The additional components in the evolved system are shown in light blue: the *RoIE* within the *Global Trigger*, *L1Track*, and *L1CTP*. For the *ITk* strip detector, the regional data will be produced by *FE* electronics then sent out via *FELIX*, while for the *ITk* pixel detector outer layers (Layers 2-4) and the forward rings, the regional data will be extracted from the full Level-0 data stream in *FELIX*, then sent out. The remaining full detector data is sent after a *L1A*. Direct connections between each Level-0/Level-1 trigger component and the Readout system are suppressed for simplicity.



---

## 14.4 Design Implications for the TDAQ Sub-systems

### 14.4.1 Level-0 Calo Trigger

The change from the single-level hardware trigger to a two-level system is transparent for Level-0 Calorimeter trigger hardware. It might imply, however, a few small changes in the firmware depending on the readout scheme. In the single-level system, the readout will be issued upon receiving the [L0A](#) signal. In a two-level system it seems advantageous to issue the readout after the [L1A](#) signal at lower rate and potentially higher readout bandwidth. Moving from one readout scheme to the other requires minor changes in the firmware for all [FEX](#) systems as well as the [Hub](#) module and its daughter card.

### 14.4.2 Level-0 Muon Trigger

For the evolution from the single-level to the two-level hardware trigger architecture, no change is required for the Level-0 muon trigger system. The hit data for [RPC](#), [TGC](#), [NSW](#), and [MDT](#) are transmitted to [FELIX](#) after a [L1A](#) signal in the two-level architecture. The [RPC](#), [TGC](#), and [MDT](#) hit data are streamed off the chambers to the off-detector electronics; these buffers are compatible with the Level-1 latency of the evolved architecture. The [NSW](#) hit data are stored in the buffers in the on-detector electronics before the readout; these buffers are also compatible with the Level-1 latency of the evolved architecture.

### 14.4.3 Global Trigger

The [Global Trigger](#) will evolve in the two-level architecture to incorporate additional inputs from the [L1Track](#) trigger system. This will allow the [GEP](#) modules to include tracking information when refining trigger objects for Level-1 trigger decisions, which has the potential for large rate reductions of pile-up and noise. The multiplexed [L1Track](#) data will be provided to the [GEP](#) in identical fashion as the other data inputs so that the track information can be included in extended object processing. A new [RoIE](#) will receive the pattern of active trigger inputs contributing to the [L0A](#) decision ([TIC](#)) and use this to suitably seed regional readout of the [ITk](#) strip tracker and selection of outer [ITk](#) pixel data for transmission to [L1Track](#).

To enable this the [Global Trigger](#) will have three additional major functions in the two-level architecture over and above the Level-0 system described in Chapter 9:

- At the end of the Level-0 trigger process, after an [L0A](#), the [Global Trigger](#) must identify the [RoIs](#) corresponding to accepted trigger hypotheses that require tracking information from the [L1Track](#) regional hardware-based tracking system.

## 14.4 Design Implications for the TDAQ Sub-systems

---

- The **Global Trigger** must initiate the generation of the **R3** in the **ITk** strip tracker to seed the transfer of data to the **L1Track**, transferring this information to the **ITk** Readout System along with the corresponding off-detector data selection for the outer **ITk** pixel layer that is buffered in the **FELIX** system.
- The **Global Trigger** shall store Level-0 objects until the **L1A**; these objects may be updated with **L1Track** information.
- When the **L1Track** data are available, the **Global Trigger** shall use this information to update the Level-0 trigger hypotheses to form the input to **L1CTP** for the Level-1 trigger decision and transfer this information to the **L1CTP** system.

These functions will be implemented by updating the **Global Trigger** as follows:

- the number of **GEP** nodes will remain the same, with the **GEP** node that processed an event at Level-0 processing it at Level-1, and the **GEP** firmware will be updated to include these additional steps:
  - produce the compact data structure mapping the **TIP** onto the **TOBs** which caused those **TIP** bits to be fired every event as the **TIP** are being transferred to the **CTP Interface**, which in the single-level architecture is only produced after the **L0A** decision, and transfer this to the new **RoIE** described below
  - buffer the Level-0 event information pending the receipt of tracking information
  - when the Level-1 track information is available, use this to update the Level-0 trigger hypotheses and transfer the resultant Level-1 **TIP** to the **CTP Interface**
- additional **MUX** modules will be added to receive the tracking information from **L1Track** in a manner analogous to the Level-0 **MUX** modules
- additional **MUX** variant modules will be added, running the new **RoIE** firmware, as described in more detail below
- the **CTP Interface** will be updated to transmit the Level-1 **TIP** to **L1CTP**

The updated system processing and trigger decision flow including these new steps within the **GEP** and the new modules are illustrated in Fig. 14.5.

### Impact on the GEP

The step to produce the data structure mapping the **TIP** onto the **TOBs** which caused those **TIP** bits to be fired will not significantly impact on the resource usage on the **GEP** as this functionality is required in the single-level scheme to generate the readout data - it just moves from running after the **L0A** to during the **CTP Interface** and **L0CTP** processing.

The same **GEP** node that processed an event at Level-0 will process the event at Level-1. The buffering step will therefore consist of the **GEP** node streaming the event information into its local **RAM** after the **L0A** and retrieving it after a suitable interval for amalgamation with the tracking information from **L1Track**.

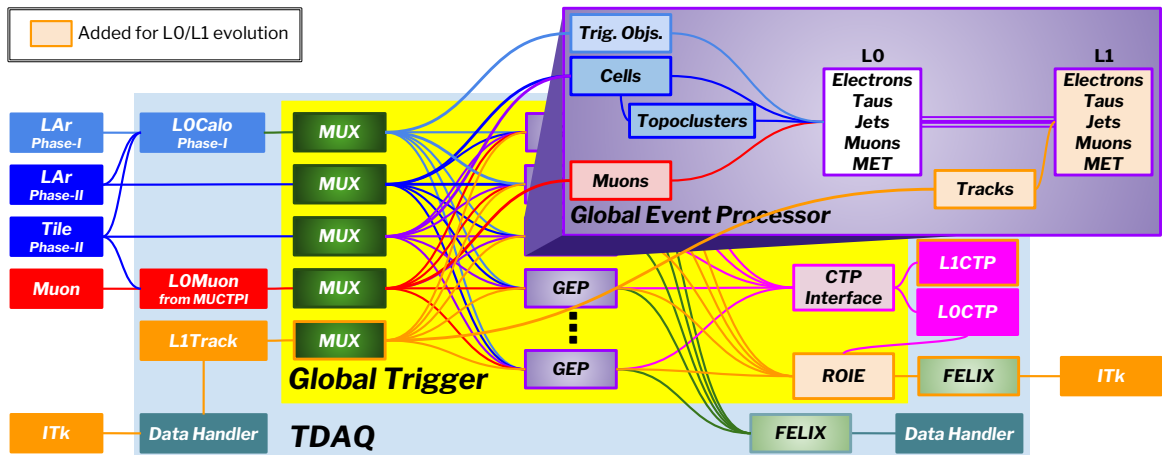


Figure 14.5: Illustration of the *Global Trigger* system processing and trigger decision flow in the evolution to a two-level architecture.

In the single-level scheme the *GEP* node would have streamed its readout information to the Zynq chip after *L0A*; this step now occurs on the retrieved information after *L1A*.

The algorithms and hypotheses using the tracking information are not expected to be as intensive as the jet-finding algorithm and have a nominal allocation of 10% of the device, taking resource utilisation to 60%.

### The RoIE

The *RoIE* consists of a set of *MUX* nodes with point-to-point links to each *FELIX* module in the *ITk* strip and outer pixel readout. Every bunch crossing the *RoIE* receives the *TIC* trigger decision from *LOCTP* and identifies those *TOBs* requiring tracking information by comparing this with the complete information from the respective *GEP* node about those *TOBs* contributing to accepted trigger items. It then transmits the requisite information down the correct point-to-point links to generate the *R3* in the *ITk* strip tracker and corresponding off-detector data selection in the outer *ITk* pixel readout to seed the transfer of data to the *L1Track*.

The *RoIE* firmware essentially consists of two lookup steps, *TIC* to *TOBs* followed by *TOB* to *R3* seed, which are well suited to *FPGA* implementation.

## 14.4.4 Trigger, Timing, and Control

The *TTC* paths for the evolved system are shown in Fig. 14.6. In comparison to Fig. 5.3, a *GBT* connection between the Level-1 Trigger Processors and *FELIX* and a connection

between the **RoIE** in the **Global Trigger** and **FELIX** are added. Additionally, the **CTP** is divided into a **LOCTP** and **L1CTP**, as described in Section 14.4.5.

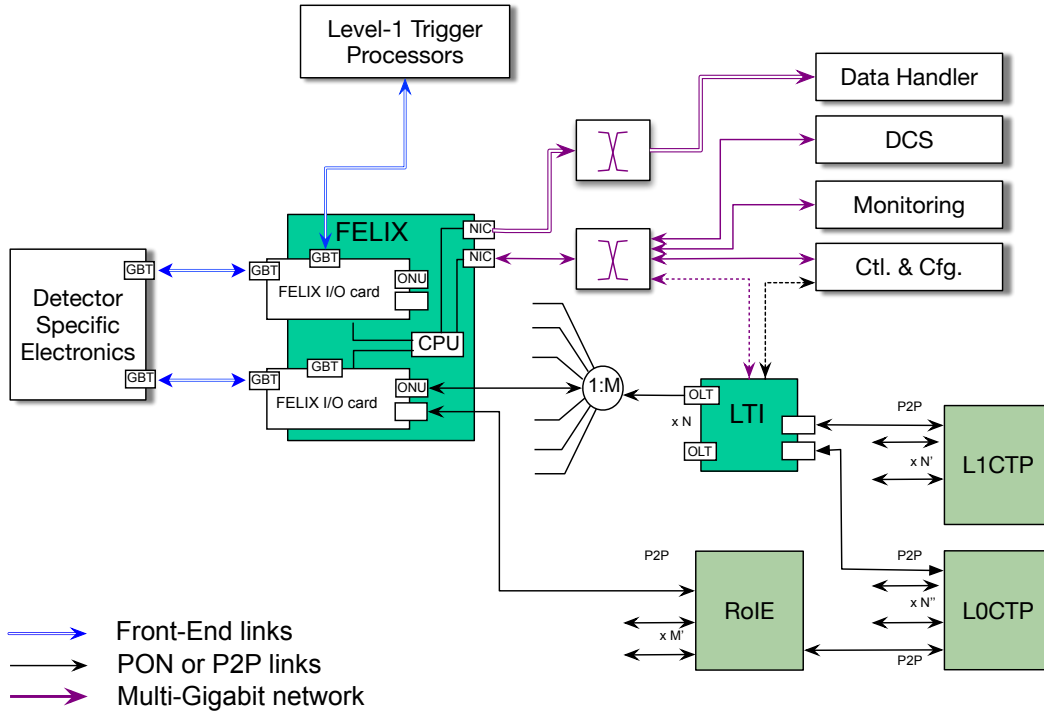


Figure 14.6: The evolutionary two-level system timing and control paths.

### 14.4.5 Central Trigger System

There is one **CTP** for the Level-0 trigger system (**LOCTP**). For the evolved two-level trigger system, it is planned to use an additional **CTP** dedicated to the Level-1 trigger system (**L1CTP**). The **L1CTP** would use the same hardware as the **LOCTP**, but with different firmware for the Level-1 functions.

Figure 14.7 shows the Central Trigger part of the evolved Level-0/Level-1 architecture. Compared to the single-level scheme, the **L1CTP** is added in the same **ATCA** shelf. The **LOCTP** calculates, for every **LOA**, the pattern of active trigger inputs contributing to the **LOA** decision (**TIC**). This is transmitted to the **RoIE** of the **Global Trigger**, where it is used in identifying regions of the detector to be used in the Level-1 **rHTT**. For sending the **TIC** pattern (along with associated information), a 12-way ribbon fibre transmitter module [14.2] is foreseen which allows to connect 12 serial optical links with line rates of 9.6 Gb/s. In addition, it has to provide, for each **LOA**, the pattern of active trigger items after veto (**TAV**) and transmit it to the **L1CTP**. For sending the **TAV** pattern to the **L1CTP**, another 12-way ribbon fibre transmitter module is foreseen. The inputs from **LOCTP** can be used for triggers that

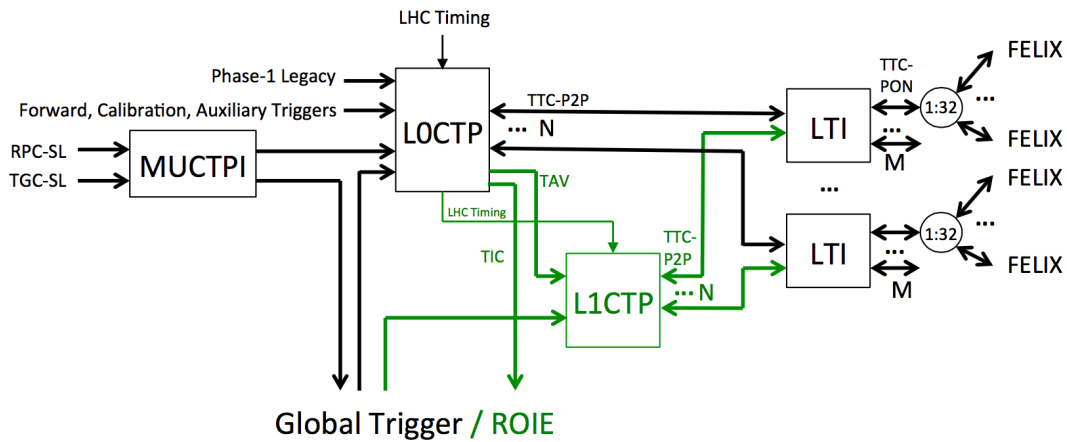


Figure 14.7: Context diagram showing the *LOCTP* and *L1CTP* systems interfacing to the *MUCTPI*, the *Global Trigger*, the *RoIE*, and the *FELIX* systems (via the *LTI* modules). Only the trigger path is depicted; control, monitoring, and readout are not shown.

do not require validation at Level-1 and to provide pre-scaled triggers without the Level-1 requirement for monitoring and efficiency measurements.

Figure 14.8 shows the *L1CTP* in more detail. It uses a CTPIN module to receive the *TAV* word from the *LOCTP*. A CTPCORE module with Level-1-specific firmware receives the Level-1 trigger inputs from the *Global Trigger* via one of its two 12-way ribbon fibre receiver modules, while the other receiver module is used to receive the Level-0 *TAV* word via the CTPIN. The signals from the *Global Trigger* are synchronous with the *LHC* clock and arrive with a fixed latency after the collision. The inputs from *LOCTP* will be delayed in the CTPIN to match the later arrival time of the *Global Trigger* inputs. The trigger processing in the CTPCORE includes a synchronisation and time alignment step, the application of trigger logic and pre-scaling. If necessary, veto gating and deadtime generation can be applied. The resulting *L1A* and associated *TTC* information will be fanned out to the *LTI* modules via the *TTC* outputs of the CTPCORE board. The second SFP input of the *LTI* is used for this purpose. The *LTI* merges the Level-1 *TTC* information with the Level-0 information and makes it available on its *TTC-PON* outputs.

Table 14.3 summarises the estimated latency contributions of the critical path through the central trigger system.

For the evolved Level-0/Level-1 architecture, the *TIC* word calculation will take an additional 200 ns, which is counted in the Level-1 latency. The link for the *TIC* word from the *CTP* to the *RoIE* is estimated to 200 ns, including a 15 m long fibre. The processing step of the *L1CTP* is estimated to be 0.275 ns. The uncertainty on the latency estimates is quoted as contingency in the rightmost column of Table 14.3.

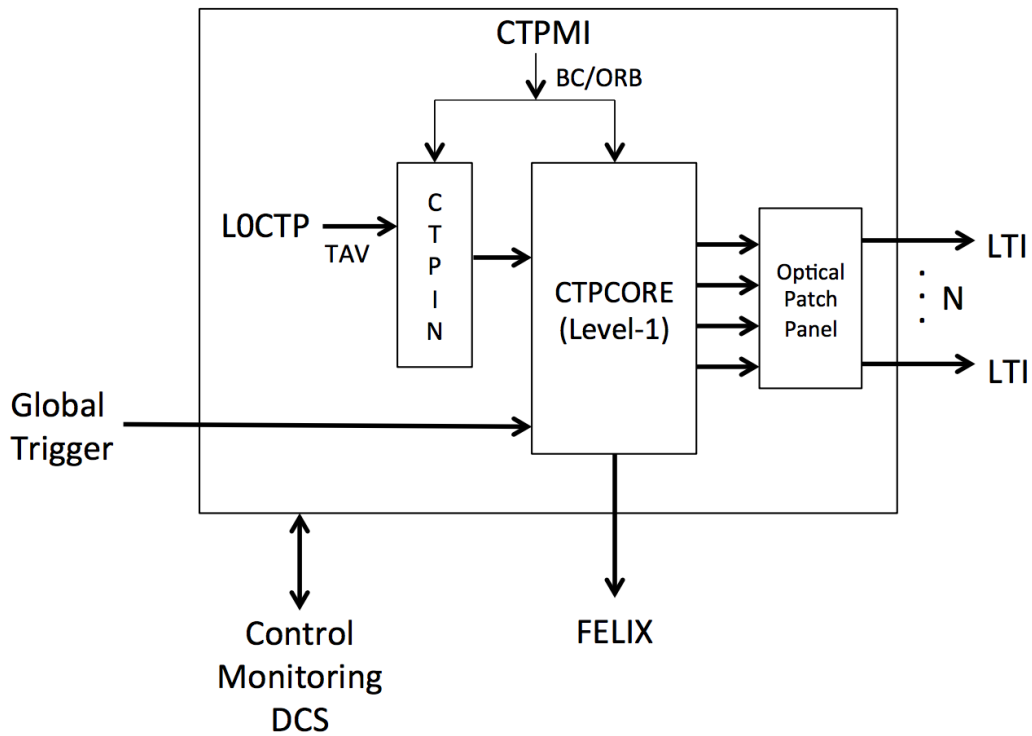


Figure 14.8: Basic architecture of the *L1CTP*, showing its main components, inputs, outputs, and the main information flow.

Table 14.3: Summary of the expected latency of the *CTP* in the evolved two-level system.

Item	Partial Current Best Estimate ( $\mu s$ )	Uncertainty/Contingency ( $\mu s$ )
Calculation of TIC in the <i>LOCTP</i>	0.200	0.050
Link TIC to <i>RoIE</i> (5 BC + 15 m fibre)	0.200	0.025
<i>L1CTP</i> Processing	0.275	0.100

### 14.4.6 Summary of Hardware Trigger Latency Estimates in the Evolved System

The latency constraints for the evolved system, as defined by the time elapsed between collision time and the arrival of the trigger decision at the [FELIX](#) outputs. The Level-0 latency is  $10\ \mu\text{s}$  for the 4 MHz rate, and is constrained by the readout of the inner pixel layer. Given a Level-0 rate of 4 MHz (2 MHz), the corresponding Level-1 rate is 600 kHz (800 kHz) and the maximum latency is  $30\ \mu\text{s}$  ( $35\ \mu\text{s}$ ), which is constrained by the [ITk](#) strip ([NSW](#) and [ITk](#) pixel) readout.

An early attempt in estimating of the [L1A TDAQ](#) Latency is presented in Table 14.4. The integral Maximum Possible Value of  $25.5\ \mu\text{s}$  is within the worst case requirement of  $30\ \mu\text{s}$ .

### 14.4.7 Data Acquisition

The overall architecture of the [TDAQ](#) system places several dependencies on the design of the [DAQ](#) system. In general terms these can be categorised using the throughput and propagation of trigger data. The specific impact of a change depends on the trigger rates which will need to be supported by the system. For the purposes of this assessment two scenarios have been used: a Level-0 rate of 2 MHz with a Level-1 rate of 800 kHz and a Level-0 rate of 4 MHz with a Level-1 rate of 600 kHz.

#### Readout

Moving to a two-level [TDAQ](#) architecture, with a hardware track trigger ([L1Track](#)) running as part of Level-1, would significantly increase the amount of input data arriving at [FELIX](#), as well as the number of links to be supported. The main contributor to this will be [ITk](#), where an extra  $\sim 7000$  links will be needed to readout the strip and outer Pixel layers at 2 or 4 MHz. There will also be  $\sim 3500$  links from to [L1Track](#) from strips and Pixel via [FELIX](#). The result of this will be an increase in the number [FELIX](#) systems needed, as presented in Table 14.5. The [FELIX](#) implementation will also become more complicated because of the need to service regional readout requests for [ITk](#) data and to transfer the corresponding data via low-latency point-to-point links to [L1Track](#). The number of Data Handler systems will also have to increase to satisfy the increasing number of readout links, as will the scale of the readout network between the these components and [FELIX](#). Each Data Handler will have to support larger local buffers to store data between a [L0A](#) and a [L1A](#). One reduced requirement compared to the single-level system would be the data rate between the Data Handlers and the Storage Handler. This would drop from 1 MHz to either of the two lower Level-1 rates. The structure of the Readout system with the split-level hardware trigger, along with the new interfaces, is presented in Fig. 14.9.

#### 14.4 Design Implications for the TDAQ Sub-systems

Table 14.4: Phase-II Hardware Trigger Level-1 Latency for the Evolutionary scenario; values are expressed in  $\mu\text{s}$ .

System	Item	BEV	Unc.	$\int$ BEV	$\int$ MPV
CTP	L0CTP Processing			6.1	8.0
	Calculation of TIC in the L0CTP	0.200	0.05		
	Link TIC to RoIE (5BC + 15m fibre)	0.200	0.03		
RoIE	RoIE processing	0.225	0.05		
	Link to FELIX (5BC + 20m fibre)	0.225	0.03	6.8	8.9
FELIX R3	R3 processing	0.200	0.05	7.0	9.2
	GBT TX	0.100	0.03		
	R3 to ITk Stave (110m fibre)	0.500	0.05	7.6	10.0
ITk Strips	GBT RX	0.125	0.05		
	HCC receive and decode	0.125	0.03		
	R3 Request to Last hit leaves HCC	3.000	0.50		
	GBTX	0.125	0.05		
	Stave to FELIX (110m fibre)	0.550	0.03	11.6	15.2
FELIX L1	GBT-FPGA RX	0.050	0.03		
	transit time	0.100	0.10		
	Link to L1Track (5BC + 20m fibre)	0.225	0.03	11.9	15.8
L1Track	Header unpacking	0.100	0.10		
	Processing	4.800	0.60		
	Result aggregation	0.250	0.25	17.1	22.9
Global Trigger	Processing	0.400	0.20		
	Link to L1CTP (5BC + 15m fibre)	0.225	0.03	17.7	24.0
CTP	L1CTP processing	0.275	0.10		
	link to LTI (5BC + 30m fibre)	0.275	0.03		
	LTI internal processing	0.100	0.03		
	Link to FELIX (5BC + 10m fibre)	0.175	0.05	18.5	25.2
FELIX	Processing	0.100	0.05		
	GBT-FPGA TX	0.100	0.03	18.7	25.5
	L1A at FELIX output (L1 Latency)			18.7	25.5

Table 14.5: Additional Readout units needed to serve a split level system. The increased network throughput may also require an increase in the number of NIC ports, but the precise number would be a subject of further study.

Component	Additional Number of Units
FELIX I/O Cards	236
FELIX Servers	120
Data Handlers	260



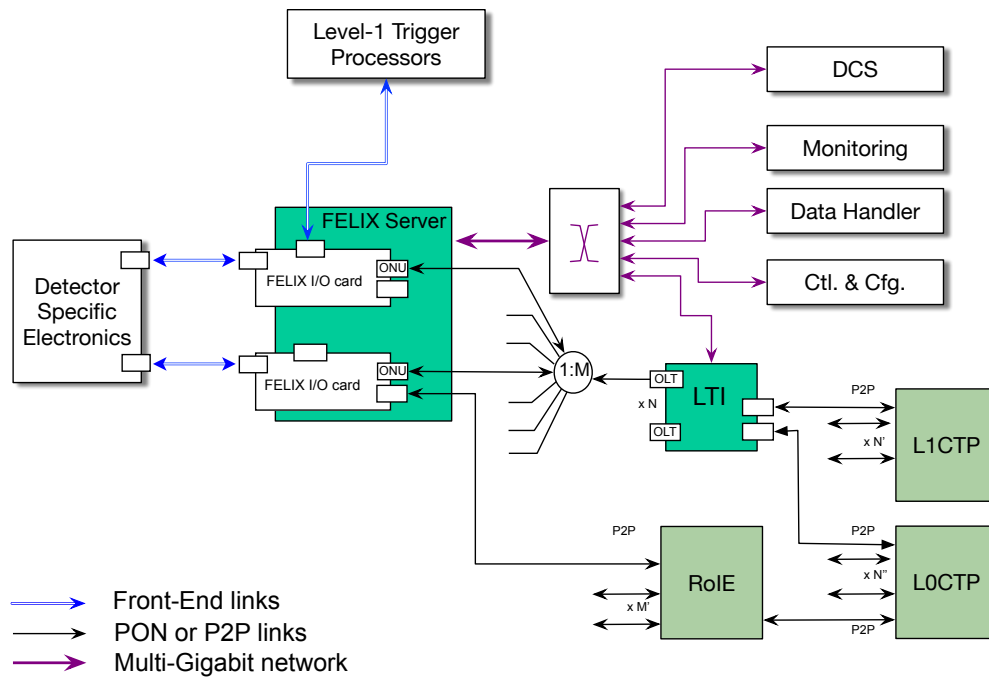


Figure 14.9: Readout system architecture for Phase-II in the Evolved System.

## Dataflow

The structure of the Dataflow system will not have to change in order to satisfy the requirements of a split-level system. **L1Track** and **rHTT** have a comparable rejection power, but the higher input rate in the evolved system means that the bandwidth requirements for transfer of data to the **EF** will be larger, with the need to transfer 600 or 800 kHz of full events instead of 400 kHz in the single-level scheme. The reduced input bandwidth to the system overall means that for a **L1A** of 600 kHz the total throughput required of the Storage Handler is reduced by 1.4 TB/s but for a **L1A** of 800 kHz the total throughput increases by 1.8 TB/s. The rates and data volumes to be transferred to the Event Aggregator and **gHTT** remain unchanged from the single-level case. A summary of the bandwidth requirements for each component of the system is presented in Table 14.6.

## 14.4.8 Event Filter

Evolution of the **TDAQ** architecture affects the **EF** in two fundamental ways: first it changes the input rate to the **EF** and potentially calls for a modified trigger strategy to be applied to

## 14.4 Design Implications for the TDAQ Sub-systems

Table 14.6: *The Phase-II Dataflow traffic required, assuming a split level system with Level-1 rate of 600-800 kHz. The data volumes coming in to the system will vary strongly by subdetector, with the minimum aggregate rate exceeding that for the single-level trigger scheme.*

Component Connection	4 MHz/600 kHz	2 MHz/800 kHz
Detector Front-ends to <a href="#">FELIX</a>	>6 TB/s	>6 TB/s
<a href="#">FELIX</a> to Data Handlers	>6 TB/s	>6 TB/s
Data Handlers to Event Builder/Storage Handler	3.2 TB/s	4.8 TB/s
Storage Handler to Event Filter	3.2 TB/s	4.8 TB/s
<a href="#">EF</a> to <a href="#">gHTT</a>	563 GB/s	563 GB/s
<a href="#">EF</a> to Event Aggregator/Permanent Storage	60 GB/s	60 GB/s

those events, and second the [rHTT](#) has moved to the Level-1 hardware-based trigger, with the associated event rejection now taking place before the [EF](#).

No specific hardware changes are necessary for the [EF](#) processors, but there is likely to be a significant increase in the level of computing power required due to the higher rate remaining after regional tracking in the evolved system. The quantitative estimate of the additional computing power needed can be determined only with a specific trigger menu and strategy in mind; the purchase cost would strongly depend on the timing of the planned evolution. These issues will be investigated in detail if the system evolution is deemed necessary according to the criteria described in Section 14.1.

Two scenarios have been developed to mitigate readout limitations, resulting in input rates to the [EF](#) of 800 and 600 kHz, in both cases less than the 1 MHz of the single-level trigger architecture. However, the rates are higher than the 400 kHz expected in the single-level scheme once regional tracking ([rHTT](#)) has been used. With regional tracking having already taken place upstream in Level-1 in the evolved architecture, this will place a greater computing burden on the [EF](#) to achieve rejection through software-based regional tracking. This will be kept to a minimum by using the output of [L1Track](#) to seed precision software tracking.

The evolved menu (Table 14.12) shows a scenario in which some hadronic signatures are given lower thresholds to boost acceptance for key physics signals, resulting in an additional [EF](#) input rate of around 170 kHz. These signatures would all require additional processing from the [EF](#) and [gHTT](#). If necessary the input rate to [gHTT](#) can be reduced by pre-selections on some signatures such as the di-tau, and use of regional rather than global tracking in the case of the 4  $b$ -jet signature. The [gHTT](#) can be tuned to handle a greater input rate by raising the  $p_T$  threshold from 1 to 2 GeV, and/or limiting global tracking to a  $z$ -vertex found by [L1Track](#). These more advanced uses of [HTT](#) could be commissioned after the baseline functionality for use with or without the evolved architecture. Some analysis-specific selection may be needed to limit the [EF](#) output rate, but this would not be expected to impose significant computing requirements.

### 14.4.9 HTT

In the evolution of the [TDAQ](#) system to the two-level architecture, the hardware that supports the regional hardware-based tracking system ([rHTT](#)) is relocated from the [EF](#) to the Level-1 trigger system, reconfigured, and renamed [L1Track](#). In order to compensate for the latency requirement on [L1Track](#), the track  $p_T$  threshold is raised to 4 GeV (from 2 GeV in [rHTT](#)). The specifications for the [rHTT](#) and [L1Track](#) are compared in [Table 14.7](#).

Table 14.7: Comparison of specifications for the [rHTT](#) in the single-level trigger scheme and [L1Track](#) in the evolved design.

Trigger	Latency requirement	Level-0 rate [MHz]	Trigger threshold [GeV]
<a href="#">rHTT</a>	No	1	2
<a href="#">L1Track</a>	6.0 $\mu$ s	2–4	4

Regional hits from the three outer layers of the [ITk](#) pixel detector and all layers of the [ITk](#) strip detector are transmitted to [L1Track](#) upon a [L0A](#) based on regions identified by the [Global Trigger](#). These regions can compose up to 10% of the [ITk](#) detector volume for a given [L0A](#) rate scheme. The [L1Track](#) coverage in the forward region relies on a sufficient number of Pixel layers being read out upon a [L0A](#).

In the 2 MHz [L0A](#) rate scheme, 192 of the [AMTPs](#) will be configured as the [L1Track](#) system. For a 4 MHz [L0A](#) rate, with the most simple extrapolation, 384 [AMTPs](#) are configured as [L1Track](#). Further optimisation for the 4 MHz case should be performed to reduce the required [AMTPs](#) modules. The [gHTT](#) function will remain as a separate system that works as a co-processor to the [EFPU](#) in the [EF](#). In the single-level scheme just described, the split of [L1Track](#) and [gHTT](#) functions into two separate system requires duplication of [AM](#) patterns and it is likely that additional hardware will be required. The need for additional hardware can be removed either relaxing the requirements on [gHTT](#) or using the full [HTT](#) system to implement both [L1Track](#) and [gHTT](#) functions together. This latter case would allow to share [AM](#) patterns among the two functions as in the Level-0-only [TDAQ](#) design. Further developments are needed to evaluate the performance of a combined [L1Track](#) and [gHTT](#) system, where [L1Track](#) processing is prioritised. System diagrams for [L1Track](#) and [gHTT](#) are shown in [Figs. 14.10](#) and [14.11](#), respectively.

The simulated track fitting efficiency and the average number of fits for first stage track fitting is shown for single jets embedded in minimum bias of pileup 200 in [Tables 14.8](#). Because [L1Track](#) latency depends on the tail of distributions, both average number of fits and the 99% - point of the distribution are given. Track parameter resolutions for [L1Track](#) are presented in [Section 13.5](#).

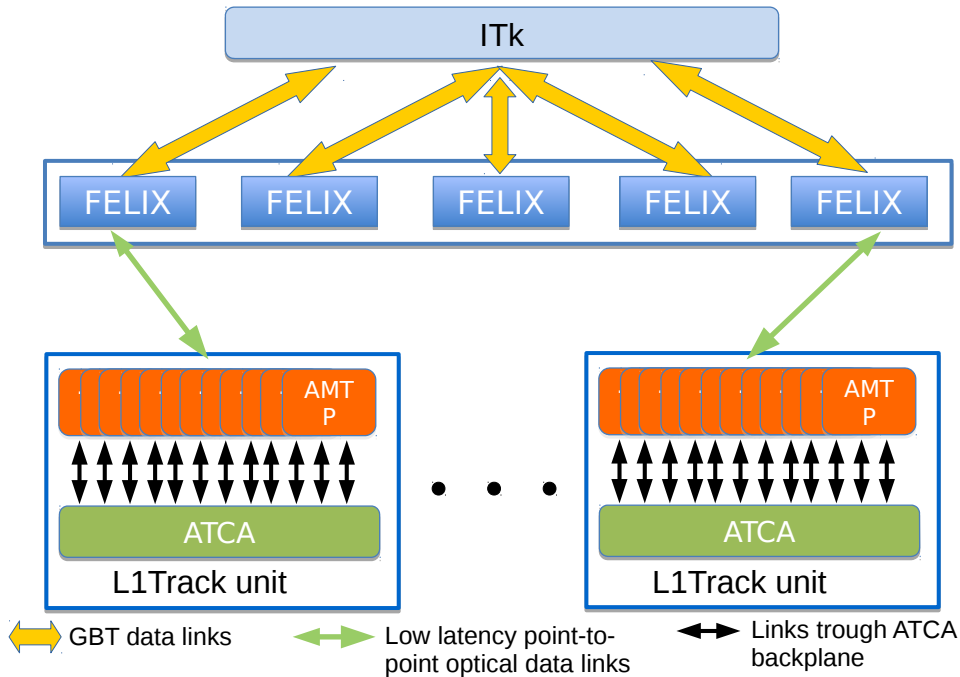


Figure 14.10: Overview diagram of the *L1Track* system showing interconnections within *L1Track* units and with the *FELIX*.

Table 14.8: First stage track fitting performance for single jets ( $p_T > 4 \text{ GeV}$ ) in PU200.

$\eta$ range	# fits (99%)	# fits $\chi^2 < 40$ (99%)	# tracks HW (99%)
$0.1 < \eta < 0.3$	614 (3287)	27 (158)	1.5 (10)
$0.7 < \eta < 0.9$	323 (2094)	19 (129)	1.1 (6)
$1.2 < \eta < 1.4$	581 (3038)	32 (200)	2.2 (13)
$2.0 < \eta < 2.2$	99 (551)	29 (153)	1.7 (7)

### Latency of *L1Track*

The latency of the *HTT* is not a critical parameter in the single-level trigger scheme since data is sent to the *HTT* from storage. In contrast, the latency is a critical parameter for *L1Track*.

The latency of data flow through the *PRM* has been studied using a simple discrete event simulation written in Python using the SimPy package. The simulation assumes a *L1Track* configuration with 2 MHz *LOA* accept rate and a 10% detector volume processed by the regional tracking per *LOA*. Realistic distributions for the number of input hits, the number of roads identified for subsequent processing, and the resulting number of track fits required, are taken from the pattern recognition studies described in Section 13.5.1. The model for

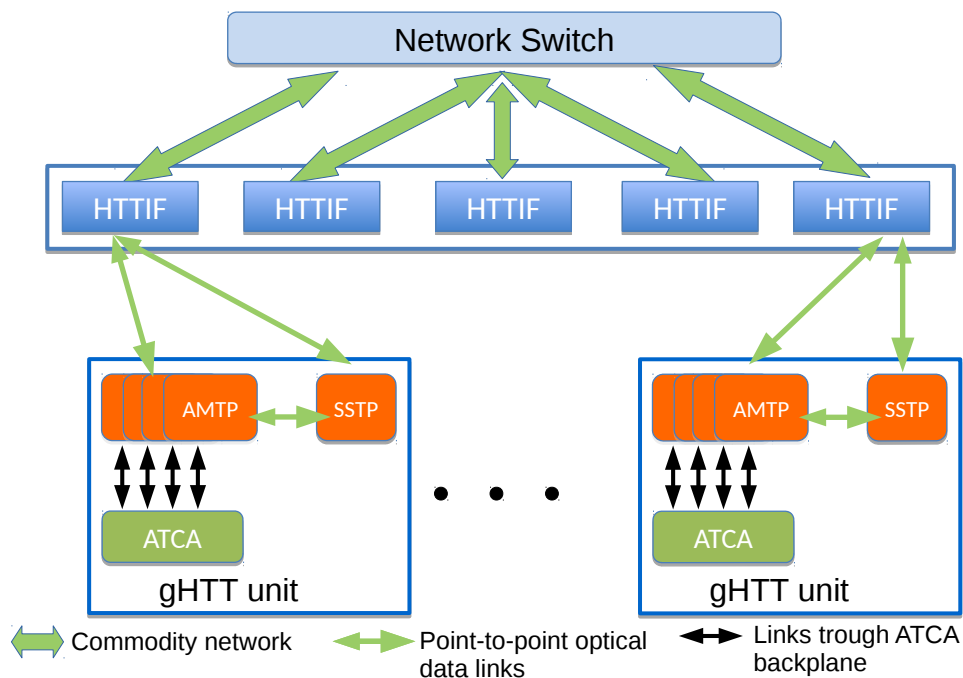


Figure 14.11: Overview diagram of the *gHTT* system showing interconnections within *HTT* units and with the *HTTIF*.

dataflow and the timings used for each stage are shown in Fig. 14.12. The model do not include dataflow on crate level. In future studies the model will be expanded to take the crate level latency into account.

In the simulation, events are generated at random from realistic distributions of time difference between successive *LOA* and with the parameters for the occupancy and fit multiplicities from the distributions mentioned, but taken collectively, to ensure that correlations between number of hits, number of fits and the processing times taken are all properly considered. The simulation generates the individual events at random, and so the simulation of expected queuing times into each stage of the simulation is correctly handled. The timeline for processing the events as explained in 13.6.3 is shown in Fig. 14.13. At  $t_1$ , the *AM* processing for event 1 is finished and *SSIDs* from event 2 can be input in parallel with the output of roads from event 1. At  $t_2$ , *SSIDs* from event 3 can be input in parallel with the output of roads from event 2. However, the input of *SSIDs* for event 4 cannot start until  $t_3$ , when the output of roads from event 2 finishes even though the input and processing of *SSIDs* from event 3 have both finished earlier. The diagram also shows that the reading of hits on roads from the *DO*, track fitting and output happen largely in parallel.

The simulation shows that the latency is primarily determined by the input of the superstrips to the *AM* chips through their shared buses. This is the motivation for having the

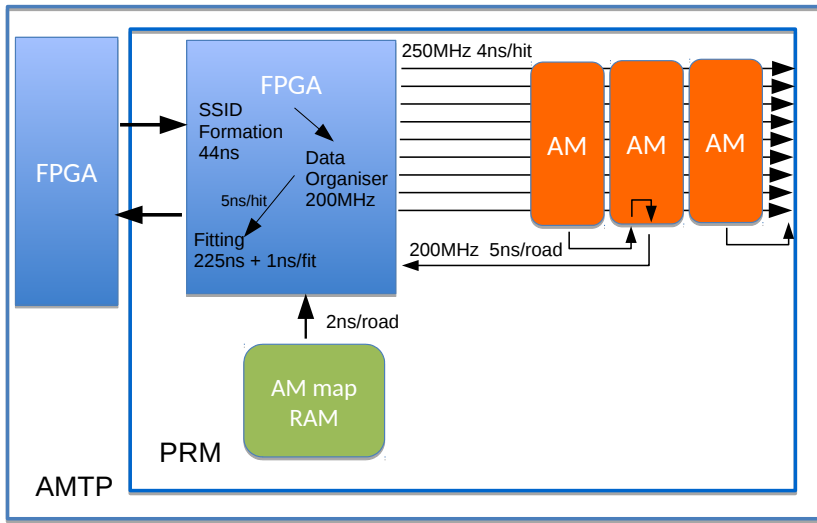


Figure 14.12: The dataflow within the PRM on the AMTP, including with the timing used in the discrete event simulation.

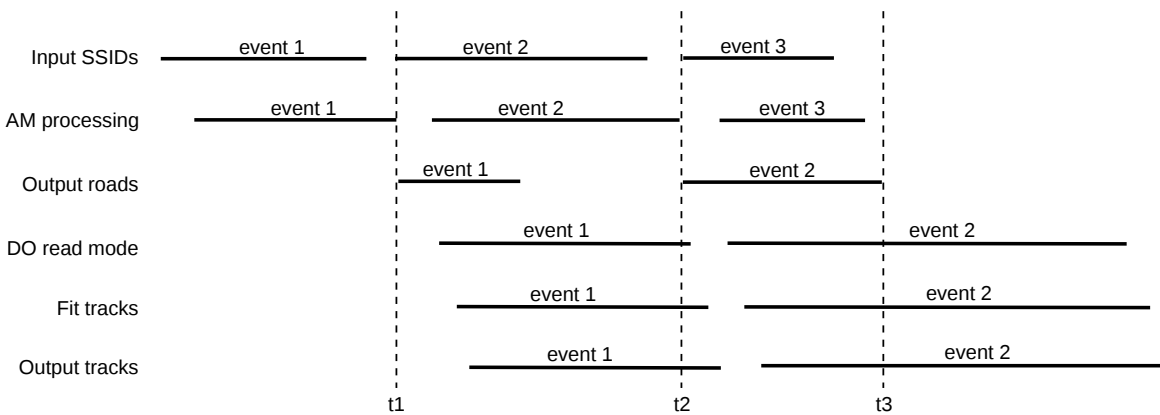


Figure 14.13: Illustrative timeline for event processing in L1Track.

AM chips on a PRM separated into two zones with independent buses for each zone. With two sets of input buses and a cut-off at  $2.5 \mu\text{s}$  for the input stage, the simulation shows that greater than 99% of the super-strips can be read in to the AM chips within the cut-off time. Figure 14.14 shows the time spent waiting for the input bus to become available and the percentage of super-strips that arrive at the AM chips. Current studies indicates that AMTP and PRM algorithms will have sufficiently low latency. More complete post-TDR studies will be need to confirm that the resources on the FPGAs planned for the L0 only TDAQ architecture will be sufficient to achieve the latency goals. At the time of FDR it

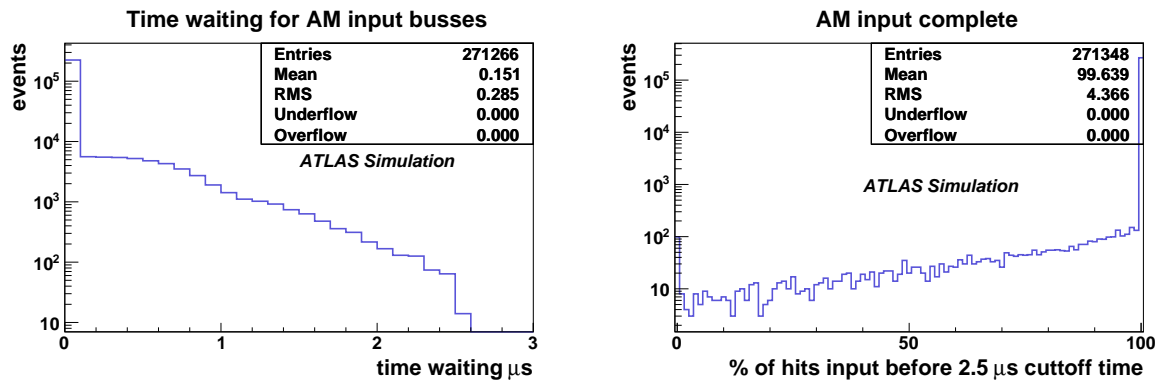


Figure 14.14: Simulation results for *L1Track* running at 2 MHz showing the time spent waiting for the *AM* chip input busses to become free on the left; and, right, the percentage of hits that arrive at the *AM* chips within 2.5  $\mu$ s.

Table 14.9: *L1Track* input dataflow for a *L0A* rate of 2 MHz. See text.

		<b>L1Track 2 MHz</b>	<b>L1Track 4 MHz</b>
	Rate (equivalent)	200 kHz	400 kHz
	total - crate - module	total - crate - module	total - crate - module
<b>L1Track regional</b>	# of shelves	16	32
<b>L1Track regional</b>	no overlap b/w	3200 - 200 - 17 Gb/s	6400 - 200 - 17 Gb/s

should be evaluated which *FPGAs* will be sufficient to implement the single-level and the dual-level *TDAQ* architecture.

### **L1Track** Dataflow and Power Consumption

The estimated input dataflow and first-stage dataflow for *L1Track* in the evolved system is summarised in Tables 14.9 and 14.10. The input dataflow values are approximate because they use the same duplication factors for both 2 MHz and 4 MHz. The 2 MHz configuration in Table 14.10 provides two copies of 0.7M patterns per  $0.2 \times 0.2$  *HTT*-region. The 4 MHz configuration provides two copies of 1.4M patterns per  $0.2 \times 0.2$  *HTT*-region. Further studies optimising the usage of patterns and *FPGA* processing as a function of rate are needed to optimise the *L1Track* system size. In Table 14.10 the relevant parameter for the cluster rate is shown for half a *PRM* since each half *PRM* will perform *AM* pattern matching independently. The calculation of the output bandwidth assumes only 160 bits per track, which include the track parameters and basic information. The transmission of the clusters on tracks uses an additional 480 bits that are not needed for the Level-1 decision but are useful for regional tracks in the *EFPUs*.

The power consumption for *L1Track* operated at both 2 MHz and 4 MHz is similar to that of the baseline *HTT*.

## 14.5 Physics Opportunities with the Evolved System

Table 14.10: *L1Track* first-stage dataflow summary for a *LOA* rate of 2 MHz and 4 MHz.

	<b>L1Track 2 MHz</b>	<b>L1Track 4 MHz</b>
Rate(equivalent)	200 kHz	400 kHz
#PRMs	384	768
# HTT-regions/PRM	3.3	1.7
#clusters/ $0.2 \times 0.2$	210 <sup>a</sup> ( $<150>$ <sup>b</sup> )	210 <sup>a</sup> ( $<150>$ <sup>b</sup> )
#clusters/half PRM (per event)	350 <sup>c</sup> ( $<250>$ <sup>d</sup> )	210 <sup>c</sup> ( $<150>$ <sup>d</sup> )
Cluster rate/half PRM	50 MHz <sup>a</sup>	60 MHz <sup>a</sup>
#roads/PRM(per event)	330	170
Fit rate/PRM	400 MHz	400 MHz
Tracks out per PRM(per event)	8	4
Bandwidth out/PRM	260 Mbps	130 Mbps
Bandwidth out/shelf	4 Gb/s	4 Gb/s
#ATCA shelves	16	32
Total Bandwidth out	100 Gb/s	200 Gb/s

<sup>a</sup> Busiest layer.

<sup>b</sup> Average layer occupancy.

<sup>c</sup> Busiest layer with worst case overlap estimate.

<sup>d</sup> Average layer occupancy with worst case overlap estimate.

Table 14.11: *Power summary for L1Track for 2 MHz and 4 MHz LOA rate schemes.*

	<b>L1Track 2 MHz</b>	<b>L1Track 4 MHz</b>
number of AMTPs	192	384
Main card (including DC/DCs)	100 W	100 W
PRM FPGA	30 W	30 W
PRM 12 AM ASIC	60 W	60 W
PRM others (RAMs, IO fanout, AM core DC/DC)	25 W	25 W
Total/AMTP	330 W	330 W

## 14.5 Physics Opportunities with the Evolved System

The evolved system provides a mitigation strategy for detector readout bandwidth limitations and hadronic trigger rate uncertainties, but it also provides opportunities to reduce thresholds and improve acceptance for physics signatures. One use case of this option is to address a new physics discovery or important theoretical progress. Here we present how the additional rate available could be used to substantially improve acceptance for a range of known Standard Model and *BSM* signatures. Table 14.12 shows the gains achievable by distributing the additional Level-0 rate across four prospective trigger selections:  $E_T^{\text{miss}}$ , di-hadronic tau, 4-jet, and inclusive *VBF* Higgs. The  $E_T^{\text{miss}}$  trigger would deliver  $\approx 2.4 \times$  the acceptance for  $ZH \rightarrow \nu\nu bb$  and  $\approx 2 \times$  acceptance for the example compressed *SUSY* model presented in Section 2.6. The di-hadronic tau trigger would increase the signal acceptance



from 30% to 55% for the **VBF**  $H \rightarrow \tau\tau$  channel and 32% to 54% for the  $HH \rightarrow bb\tau\tau$  channel. The four-jet trigger would improve the limit on  $HH \rightarrow 4b$  from  $1.85 \sigma/\sigma_{\text{SM}}$  to  $1.65 \sigma/\sigma_{\text{SM}}$ . Finally, the inclusive **VBF** Higgs acceptance would increase from 6.6% to 10%. These are the acceptance gains that would be available at the Event Filter, but the corresponding events could not be recorded in the planned 10 kHz Event Filter output limit. Instead, analysis-specific selections would need to be applied. For example, a requirement of a soft lepton in addition to  $E_{\text{T}}^{\text{miss}}$  would be appropriate for the **SUSY** models, and the exotic Higgs searches could require additional soft jets in addition to a  $E_{\text{T}}^{\text{miss}}$  requirement. These are possible in the Event Filter but not in the Level-0 trigger because of the low thresholds of these objects.

Table 14.12: *Prospective additional triggers for an evolved system. The gains for example physics channels are described in the last column.*

Signature	Single-Level Scheme Threshold	Dual-Level Scheme Threshold	Level-0 (kHz)	Level-1 (kHz)	EF before analysis specific cuts (kHz)	Gain
$E_{\text{T}}^{\text{miss}}$	210 GeV	160 GeV	800	80	3	$2\times$ acceptance for compressed <b>SUSY</b> model and $2.4\times$ for $ZH \rightarrow \nu b b$
di- $\tau$	40, 30 GeV	30, 20 GeV	800	80	2.2	increased acceptance from 30% to 55% for <b>VBF</b> $H \rightarrow \tau\tau$ and 32% to 54% for $HH \rightarrow bb\tau\tau$
4 jet w/ 2-btags	65 GeV	55 GeV	800	100	0.4	improved limit in $HH \rightarrow 4b$ from 1.85 to $1.65 \sigma/\sigma_{\text{SM}}$
<b>VBF</b> Higgs	75 GeV + topological	60 GeV + topological	280	40	40	increased acceptance from 6.6% to 10% for inclusive <b>VBF</b> Higgs production
Total			2680	300	-	

The example evolved system menu in Table 14.12 allows 3 MHz more events to be passed from Level-0 to Level-1, but that rate must then be reduced to 600 kHz using regional tracking and algorithms similar to those described in Section 6. The data fraction of an event needed is similar to the fractions shown in Table 6.5. The increased rate into the Event Filter leads to an increased need for precision tracking with **gHTT** for use in  $b$ -tagging and track-based pile-up corrections and calibrations. In order to mitigate this increased need, the **HTT** co-processor can be used in a regional or a global capacity. Because of the length of the beam spot, a large amount of detector data is required to reconstruct charged particles arising from a given vertex. However, in the evolved system charged particles have already

been reconstructed in [L1Track](#), along with the possibility for coarse determination of the z-position of the primary vertex. This can be used to restrict the detector data used so that only tracks that originate from a region along the beam line near the triggered collision are reconstructed. A similar algorithm is used in the Run 2 [HLT](#) for *b*-tagging; first, a vertex is reconstructed with a fast tracking algorithm, and then the precision tracking is restricted to a selected region within the beam spot. This aggressive algorithm is not assumed for the single-level configuration in order to ensure its robustness and flexibility.

## 14.6 Challenges of the Evolved System

The single-level and dual-level trigger configurations chosen for the baseline trigger architecture have been carefully evaluated. The single-level hardware trigger architecture was chosen in order to minimise the complexity of the [TDAQ](#) system and [TDAQ](#)-detector interactions, the cost, and the challenges of commissioning [L1Track](#). These considerations are described in detail in the following paragraphs.

The evolved architecture configuration is considerably more complicated than the single-level architecture. In addition to performing all of the capabilities of the single-level hardware trigger, the dual-level hardware trigger in the evolved system must accommodate the fast collection of [RoI](#) information and the formation of [R3](#) signals, as well as the processing of [L1Track](#) events once they are received by the [Global Trigger](#). The data handling in [FELIX](#) and the [Data Handler](#) is significantly more complex, as is the signalling from the Central Trigger via the [LTIs](#) to the [FELIX](#). More complex signal handling is also used for detectors that make use of the Level-1 signal before reading out, notably the strip and inner-pixel front-end [ASICs](#).

There are additional costs associated with the evolved scheme, including those associated with reading out the detectors at the higher Level-0 rates (especially the outer pixels that contribute to [L1Track](#)); the higher rate [L1Track](#) processing; additional modules, patch panels, and fibres for [L1CTP](#), an additional [MUX](#) module to handle inputs from [L1Track](#) for the [Global Trigger](#), the [RoIE](#) for the [Global Trigger](#); the [RoI](#) distribution; the additional [FELIX-L1Track](#) interface; and finally, additional [AMTPs](#) needed to split the [HTT](#) functionality. In addition, there are other costs associated to the detector systems that are described in the relevant [TDRs](#). Foreseeing the Level-1 functionality in the baseline scenario leads to a very small impact on the overall cost of the system. These considerations are summarised in [Table 14.13](#).

The single-level trigger configuration has the distinct advantage of a simpler commissioning scheme than that of the evolved system. Furthermore, the baseline system will need to cope with trigger rates that are an order of magnitude higher than those in Run 3. However, in commissioning the single-level trigger system, there is no need for disruption to the detector readout while the [HTT](#) is being commissioned. It will be possible to operate the [rHTT](#) passively, sending data to and recording results from the [rHTT](#) without using

the results in the [EF](#) trigger decision<sup>2</sup>. In contrast, commissioning the [L1Track](#) trigger for the evolved scheme requires intervention at the level of the [FELIX](#) for the [ITk](#) outer pixels and at the level of the on-detector [ASICs](#) for the strips. If these actions provoke synchronisation errors, for example, there could be a serious impact on the data-taking, including the possibility of data loss and downtime. Should ATLAS determine it to be necessary or beneficial to the physics programme to move to the evolved architecture at a later stage, the operations team would be integrating well-understood [HTT](#) hardware into the Level-1 trigger, with mature [FELIX](#) and [ITk](#) electronics systems.

Experience from ATLAS operations during Run 1 and Run 2 dictates that reliability, robustness, and stability are most easily achieved when systems are kept as simple as possible.

Table 14.13: *Summary of hooks put into the baseline design in order to accommodate a potential evolution and additional hardware and firmware needed for the evolved system.*

System Component	Hooks for Evolved System	Additional Hardware/Firmware Needed for Evolved System
Level-0 Calo Level-0 Muon	sufficient <a href="#">FPGA</a> resources	minor firmware changes
<a href="#">Global Trigger</a>	extra transceivers	additional <a href="#">MUX</a> modules to receive information from <a href="#">L1Track</a> ; extra <a href="#">GCM</a> modules configured as <a href="#">RoIEs</a> ; extra <a href="#">GCM</a> module for <a href="#">L1CTP</a> interface; additional <a href="#">GEP</a> firmware
Central Trigger	extra optical connectivity	add <a href="#">L1CTP</a> plus additional patch panels and fibres
Readout	none	additional <a href="#">FELIX</a> I/O Cards, servers, and Data Handlers; new <a href="#">FELIX</a> /Data Handler firmware and software; low-latency links to <a href="#">L1Track</a> ;
Dataflow	none	larger bandwidth requirements
Event Filter	none	significant increase in computing power
<a href="#">HTT</a>	<a href="#">rHTT</a> hardware and firmware must meet <a href="#">L1Track</a> latency requirement	separation of regional and global functionality; additional <a href="#">AMTPs</a> ; new firmware

## References

- [14.1] ATLAS Collaboration, *Technical Design Report for the ATLAS Inner Tracker Pixel Detector*, Tech. Rep. CERN-LHCC-2017-021, CERN, Geneva, (*in preparation*).

<sup>2</sup> This is similar to the commissioning strategy used for [FTK](#).

## 14.6 Challenges of the Evolved System

---

- [14.2] *MiniPOD AFBR-814VxyZ, AFBR-824VxyZ, 14 Gbps/Channel Twelve Channel, Parallel Fiber Optics Modules*, Tech. Rep. AV02-4039EN, Avago Technologies, 2013.

## 15 DCS and TDAQ interfaces

The **DCS** [15.1] has the task to permit coherent and safe operation of ATLAS and to serve as a homogeneous interface to all sub-detectors and to the technical infrastructure of the experiment. The **DCS** must be able to bring the detector into any desired operational state, to continuously monitor and archive the operational parameters, to signal any abnormal behaviour to the operator, and to allow manual or automatic actions to be taken. In order to synchronise the state of the detector with the operation of the physics data acquisition system, bi-directional communication between **DCS** and run control must be provided. Finally, the **DCS** has to handle the communication between the ATLAS sub-detectors and other systems that are controlled externally, such as the **LHC** accelerator, the CERN technical services, the ATLAS magnets and the **Detector Safety System (DSS)**.

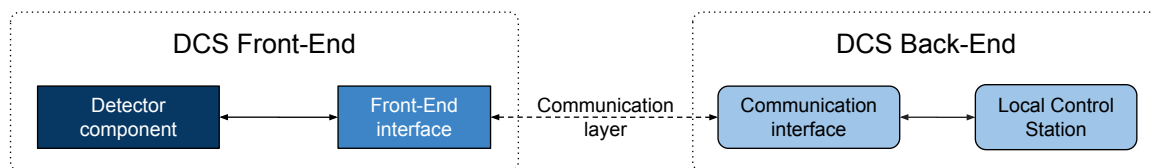


Figure 15.1: *Front-end and back-end components within a DCS system.*

Figure 15.1 shows the architectural components defined for a given system within **DCS** which is divided into Front-End (**FE**) and Back-End (**Back-End (BE)**). The detector hardware is equipped with **FE** interface equipment which can be purpose-built electronics and their associated services such as power supplies or use general purpose I/O components such as the **Embedded Local Monitor Board (ELMB)** [15.2] or **ELMB++** [15.3]. The connection to the **BE** is provided by a communication layer such as Ethernet or **CAN** bus. The **BE** is formed by **Local Control Stations (LCSs)** (**COTS** rack servers) and their communication interfaces (e.g. Ethernet **Network Interface Card (NIC)** or **CAN** bus interface).

The **BE** is organised in three layers (see Fig. 15.2): the **LCSs** for process control of sub-systems, the **Sub-detector Control Stations (SCSs)** for high-level control of a sub-detector allowing stand-alone operation, and the **Global Control Stations (GCSs)** with server applications for common functions and human interfaces in the ATLAS control room for the overall operation. The **BE** consists of control stations, interconnected as a distributed system of **Supervisory Control And Data Acquisition (SCADA)** applications based on **Simatic WinCC Open Architecture (WinCC OA)**. Communication to the **FE** interfaces is performed

## 15.1 Interfaces for On-Detector FE Components

using a middleware layer based on the industry standard [Open Platform Communications Unified Architecture \(OPC UA\)](#).

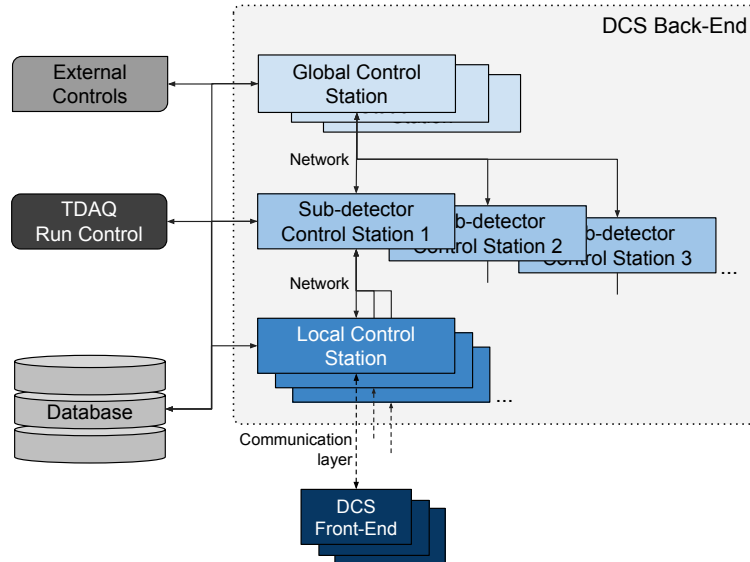


Figure 15.2: Architecture of the *DCS Back-End*.

The full *BE* hierarchy from the operator interface down to the level of individual devices is represented by a distributed *Finite State Machine (FSM)* allowing for standardised operation and error handling in each functional layer. The *SCS* layer also serves as primary interface to the TDAQ run control system while the *GCS* layer integrates with the external control systems (LHC, Cryo, Cooling etc.). All control stations may archive operational parameters to a global database.

## 15.1 Interfaces for On-Detector FE Components

Monitoring and control of on-detector components will follow two different paths with respect to communication and powering. The first path is entirely dedicated to *DCS* functions with independent communication and power lines. This serves as the primary path for important *DCS* data which require a high degree of reliability and availability at all times, including when the detector and its integrated electronics are not fully powered. An example for this path would be an implementation based on the proposed *ELMB++* [15.3].

The secondary path is using the communication infrastructure of the on-detector electronics usually based on the *GBT* or *lpGBT* chipsets which are interfaced with the off-detector back-end via Versatile Links (optical) [15.4] and the *FELIX* system. The controls data in this path may in some cases need to share resources with other types of data such as readout and

trigger data. This path is foreseen for controls data with less critical use cases or if an independent path is impossible to implement, e.g. due to constraints on physical space or available services.

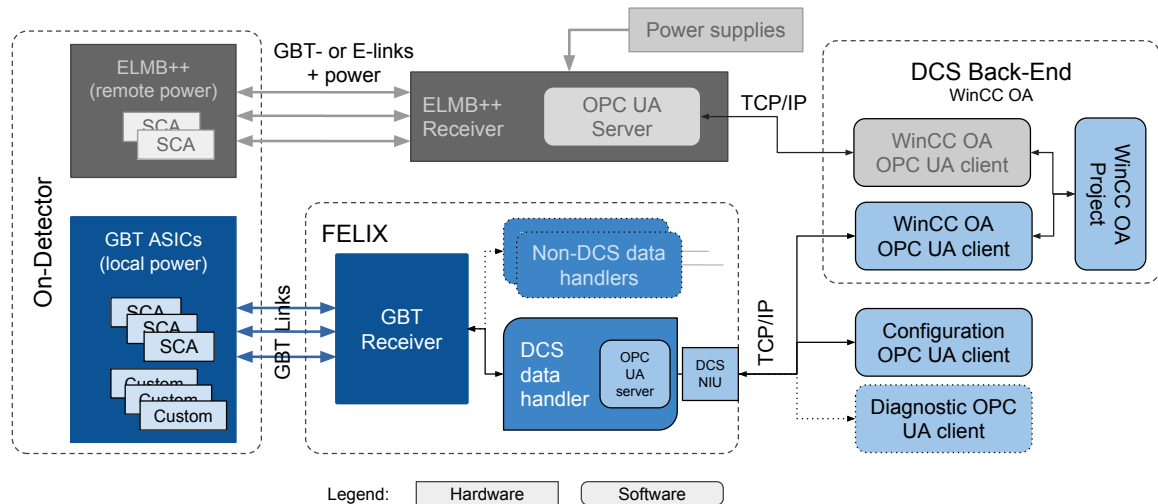


Figure 15.3: On-detector front-end interface schema showing the two paths for controls data and associated powering. The primary path shown is an implementation based on the ELMB++ chain. The secondary path shown is for front-ends such as the *GBT Slow Control Adapter (GBT-SCA)* or custom *ASICs* interfaced with Versatile Links and the *FELIX* system. The handling of DCS data shown separates the data paths in firmware only. Another possibility is to separate the data in software on the *FELIX* host.

Figure 15.3 gives an overview of these two paths for which details are given in the following two sections.

### 15.1.1 Exclusive DCS On-Detector Control Path

Control and supervision of on-detector hardware components in the ATLAS cavern usually covers situations when these components are not powered or the powering has issues which need to be diagnosed. Independent (remote) powering of the *DCS FE* interface hardware is thus vital along with an independent communication path.

In analogy to the current ATLAS solution based on the *ELMB* [15.2] for which the communication is based on *CAN* bus and associated powering is provided by dedicated *CAN* power supplies, the *ELMB++* was proposed [15.3] as the standard solution to implement the primary path for *DCS* systems. Apart from the critical improvement in radiation hardness required by the HL-LHC environmental conditions, the *ELMB++* provides general performance enhancements and additional analog and digital interface standards. The functionality covers:

## 15.1 Interfaces for On-Detector FE Components

- ADC channels with effective conversion rate on the kHz level and 12 bit precision,
- DAC channels (12 bit),
- digital I/O support for SPI, JTAG, I<sup>2</sup>C, GPIO,

and is provided by an *ELMB++ satellite* containing for example **GBT-SCA** ASICs with associated **Electrical chip-to-chip interconnect (e-link)** connectivity and a DC-DC converter for remote powering. Several satellites can be interfaced in a star-point topology directly via **e-links** to a *mobile receiver* if the link distance does not exceed a few meters or to an intermediate *ELMB++ hub*. The hubs feature full **GBT** (with associated **GBT-SCA**) or **lp-GBT** functionality together with a DC-DC converter for remote powering. Towards the back-end, several hubs can be connected via Versatile Links to the mobile receiver. These individual components of the ELMB++ variant are shown in Fig. 15.4 while the topology architecture is illustrated in Fig. 15.5.

### 15.1.2 On-Detector Control via the Readout Path

For the secondary control path where **ASIC** components with **DCS** related functions are interfaced via Versatile Links (e.g. via the **GBT** chipset), the communication on the off-detector side is handled by the **FELIX** system. Two categories of functions should be distinguished: controls data for higher priority control functions which are important for the regular operation of the detector component such as powering control and monitoring and secondly lower priority detector diagnostics data.

If affordable, a dedicated Versatile Link for **DCS** data is foreseen to avoid interference with other data types which could lead to untimely or unsuccessful delivery of controls data. However, in many systems this separation is not feasible, e.g. due to space or power constraints, and the data streams are only possible to be separated within the **FELIX** system.

Within a given Versatile Link, the controls data should be always handled via the **GBT** protocol and associated exclusively to one or several **e-links** and thus controls data can be identified within **FELIX** by the **GBT e-link** identifier. Data for these **e-links** is separated by **FELIX** in firmware and/or software and transferred to an **OPC UA** server which is connected to the **DCS** back-end via a separate physical network interface unit (cf. Fig. 15.3).

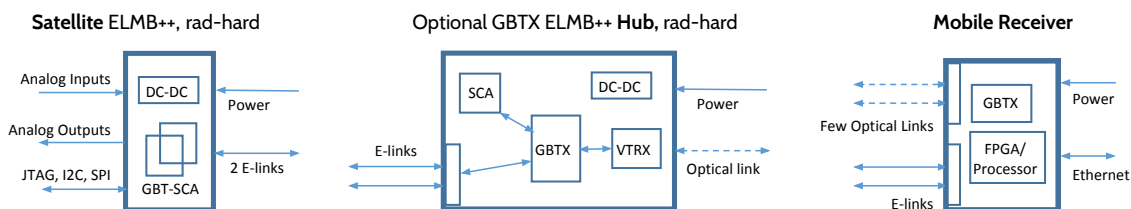


Figure 15.4: *Components of the ELMB++ chain.*



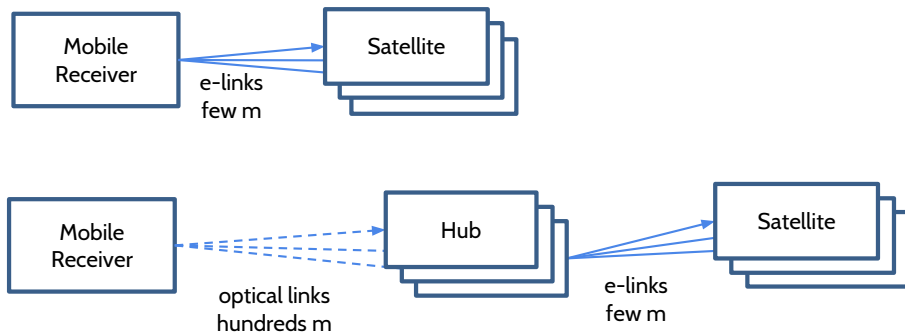


Figure 15.5: Possible topology for the use of the ELMB++ components.

If a separation of controls data is achieved in firmware only, sharing the [PCIe](#) bus and software with the non-[DCS](#) data could be avoided. Further, this approach would allow a closed control loop between the OPC UA server on the [FELIX](#) I/O card and the detector [FE](#) for important [DCS](#) functions. This would minimise downtime by not depending on host software or operating systems of the [FELIX](#) or [DCS BE](#) system hosts. For example in cases where the [FELIX](#) host needs to be rebooted, software processing is completely unavailable or processing suffers from back-pressure of other data types, the [DCS](#) path becomes quickly available again when the [FELIX](#) I/O card is available for data processing.

The lower priority diagnostics data are often merged into the event data streams, either within event data [e-links](#) if [GBT](#) is used or within customised protocols. On the off-detector side, the point of separation of these data from the event data still needs to be established.

## 15.2 Interfaces for Off-Detector Front-End Components

Off-detector [FE](#) components can be divided into the following categories:

- A. Commercial power supplies or similar devices with proprietary communication interface,
- B. Off-detector electronics crates and their commercial or purpose-built boards following the [ATCA](#) or [VME](#) standards,
- C. Other purpose-built devices.

For all device types the standard middleware solution [OPC UA](#) is used to integrate them into the [DCS BE](#). As a general principle, this middleware should be used as close to the actual device as possible in order to profit from a standard communication interface on the physical as well as the logical level. However, this principle cannot be applied equally to the device types A-C and thus different solutions are applied as noted below.

### 15.2.1 Devices with Proprietary Interfaces

Standardisation of proprietary components such as power supplies is provided in the frame of the [Joint COntrols Project \(JCOP\)](#). For each supported hardware component, Ethernet is the communication layer of choice and integration software based on [OPC UA](#) is provided along with framework components for [WinCC OA](#). The equipment manufacturers provide a hardware abstraction layer which is then used inside the [OPC UA](#) server implementation supported by [JCOP](#).

Similarly, for non-standard equipment using proprietary protocols over Ethernet, dedicated [OPC UA](#) servers will be implemented using a standardised server generation framework detailed in Sec. [15.3](#) which is also used for the [OPC UA](#) servers provided by [JCOP](#).

### 15.2.2 Off-detector Electronics based on [ATCA](#) and [VME](#) Standards

Most off-detector electronics will be based on the [ATCA](#) standard together with legacy equipment based on [VME](#). For both standards, [DCS](#) monitoring and control is provided at the level of the crates/shelves and the individual boards.

The proprietary [VME](#) crate interface is covered by a [JCOP OPC UA](#) solution (see also previous section) via Ethernet or [CAN](#) bus.

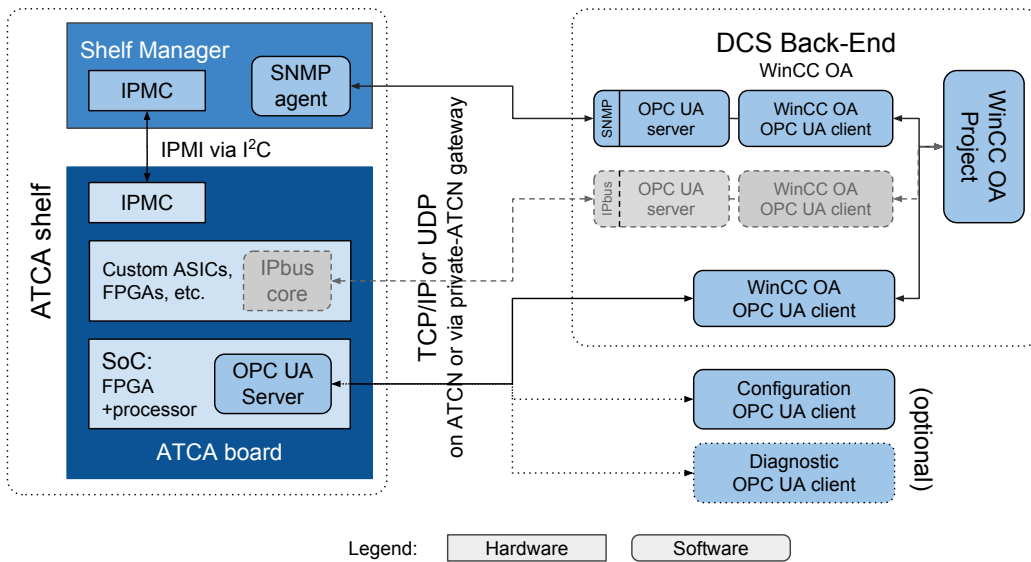


Figure 15.6: Off-detector front-end interface schema for [ATCA](#)-based electronics. Two main paths for [DCS](#) integration are shown: via the [ATCA](#) shelf manager (top) and via a [SoC](#) embedded into the [ATCA](#) board (bottom). The third path in the middle (grey shades and dashed lines) is based on an [IPBus](#) interface used in Phase-I upgrades, which is deprecated for Phase-II designs though.

For [ATCA](#) equipment, [DCS](#) integration is achieved using two complementary control paths as shown in Fig. 15.6. The first path is based on interfacing the CERN standard [ATCA](#) shelf manager via Ethernet. The shelf manager [SNMP](#) interface is then used for an [OPC UA](#) server implementation in the [DCS BE](#) for [WinCC OA](#) integration. In contrast to [VME](#), the [ATCA](#) standard clearly defines the monitoring and control parameters which are available for the shelf. Further, the standard defines the interface for [ATCA](#) boards for which control and monitoring is implemented using communication between an [IPMC](#) on the shelf manager and the board respectively. Thus the standardised board parameters such as health, presence, power, reset and custom sensor data are then also available via the shelf manager [SNMP](#) interface. However, this solution does not scale well above a few hundred parameters and thus cannot be the only [DCS](#) path for [ATCA](#) boards.

The second control path for individual boards is based on embedded processing units such as [SoC](#) components or [FPGA](#) softcore solutions which allow embedding an operating system together with an [OPC UA](#) server directly in the hardware. The [OPC UA](#) server allows a more scalable implementation of monitoring and control functions based on the digital or analog interfaces provided by the associated [FPGA](#) or [SoC](#) firmware.

If an on-board processor element is not available, data exchange via [IPBus](#) is used in some cases as fallback solution for Phase-I upgrades although it is less flexible and requires more complex software on the [DCS BE](#) (an [IPBus OPC UA](#) server). For Phase-II designs this path is considered deprecated and the [SoC](#) approach is preferred.

### 15.2.3 Interfaces for Other Purpose-Built Off-Detector Front-Ends

Any other type of (non-standard) front-end hardware component is integrated using the methods described above. In case a processor element is directly built into the hardware, an embedded [OPC UA](#) server serves as [DCS FE](#) interface connected via Ethernet. Alternatively, an [ELMB++](#) serves as I/O concentrator and allows direct monitoring or control of analog or digital I/O channels.

For [FELIX](#) systems which are used for [DCS](#) functions, monitoring and control of the respective [FELIX](#) hardware and software components themselves are provided directly within the [OPC UA](#) server embedded into [FELIX](#).

## 15.3 Back-End and Middleware

### 15.3.1 Back-End Hardware and Software

The hardware platform for the [BE](#) system is a set of industrial, rack-mounted server machines balancing maximum reliability with low cost, providing redundancy for many components such as power supplies, hard disks and network interfaces. With Ethernet as the

standard communication layer for **FE** integration, the replacement of a **BE** server should incur minimal downtime by relocating the affected **BE** application to a spare server.

The **SCADA** package **WinCC OA** (formerly known as *PVSS*) is the base for the **BE** applications, extended by the **JCOP** framework [15.5]. Since all **FE** systems are interfaced via the network, virtualisation of server operating systems and encapsulation of **BE** applications in software containers could ensure reliability and maintainability of the **BE** software on all levels.

The bi-directional interface between **DCS BE** and TDAQ run control applications will be provided by means of a software implementation with a generic API.

### 15.3.2 Standard Middleware OPC UA

The middleware standard for **DCS** systems is **OPC UA**, an industry standard for machine-to-machine communication in the controls domain, featuring:

- independence from operating systems, availability for a variety of platforms and allowing development in various programming environments. The **OPC UA** reference stack (providing lower layers of **OPC UA** internally to **OPC UA** toolkits) is delivered in standard C. **OPC UA** toolkits are available in C++, Java, .NET, python and others. Moreover, the ubiquitous TCP/IP is the communication technology used by **OPC UA** allowing for full portability of **OPC UA** to any modern network-aware operating system,
- robust data modelling capabilities: thanks to graph-based information model, data can be organised in virtually any fashion, from simple variables (of built-in or custom types) through structures (simple and nested) and trees, up to possibly cyclic graphs representing not only the current state of the system but also its meta-model and associated knowledge [15.6]. In addition, a number of techniques from object-oriented modelling is available, like method invocations, as well as techniques known from relational data modelling like queries and views. Type information is fully exposed to the client(s) and may be accessed as type instances.
- embedding into custom hardware – thanks to limited dependency on standard C and C++, some common cryptography functions and the TCP/IP stack, one can run **OPC UA** servers on any embedded computing element that has enough memory (few MB) and TCP/IP connectivity. A number of successful miniaturised and embedded controllers with **OPC UA** servers are available on the market,
- secure communication: thanks to PKI (Public Key Infrastructure) support in **OPC UA**, one can use state-of-the-art security in **OPC UA** based control infrastructure.

In order to reduce development and maintenance efforts, a framework for **OPC UA** server creation is available – the **Quick OPC UA Server Generation Framework (quasar)** [15.7]. Development starts with creation of a design file, in XML format, describing an object-oriented information model of the target system or device. Using this model, the

framework generates an executable **OPC UA** server application, which exposes the per-design **OPC UA** address space, without the developer writing a single line of code. Furthermore, the framework generates skeleton code into which the developer adds the required target device/system integration logic. This approach allows both developers unfamiliar with the **OPC UA** standard, and advanced **OPC UA** developers, to create servers for the systems they are experts in while greatly reducing design and development effort as compared to developments based purely on COTS **OPC UA** toolkits. Higher level software may further benefit from the explicit device model by using the XML design description as the basis for generating client connectivity configuration and server data representation. Moreover, having the XML design description at hand facilitates automatic generation of validation tools.

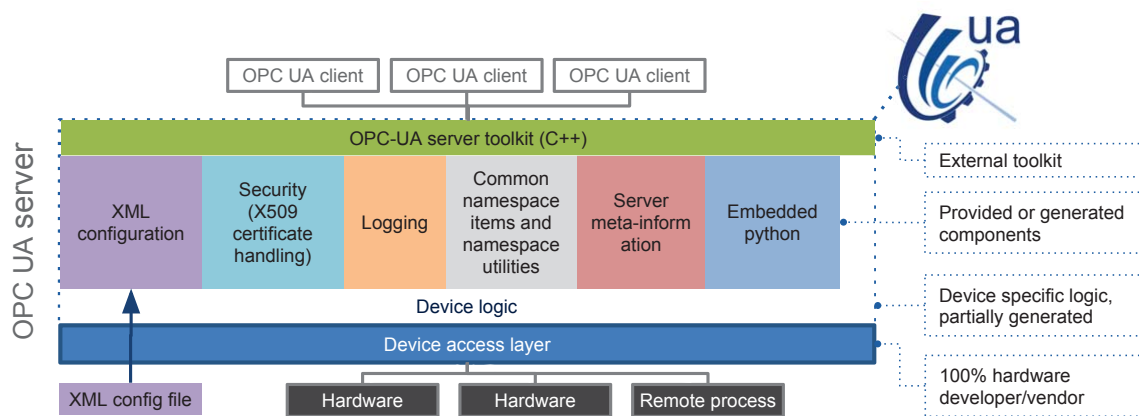


Figure 15.7: *OPC UA server architecture within the `quasar` framework.*

Figure 15.7 gives an overview of the different layers of `quasar` put into context. Controllable devices or systems are accessed using their specific access layer – often provided together with the specific device. The device logic layer functions as interface with the high level layers of `quasar` which comes in several modules covering different functional aspects. The address space module sits in the **OPC UA** end of the server, exposing data towards **OPC UA** clients, and is implemented using a **OPC UA** back-end layer with exchangeable back-end implementations. A configuration module facilitates address space and device instantiation and the definition of their relations. XML is used as the configuration format backed by XML schema definitions. A XML schema to C++ mapping generator (here: `xsd-cxx`) is used to build actual instances from configuration files. An additional subsystem called Calculated Items, operating entirely in the address space, enables creation of new variables which are derived from existing ones using mathematical functions. `quasar` further comes with optional modules such as component based logging, certificate handling, server meta-data, embedded python processing, **WinCC OA** integration tools and SQL/NoSQL archiving with historic data access. A ready-to-use build system based on CMake along with pre-configured tool-chains for several platforms such as `x86_64` or

ARM-based Linux and Microsoft Windows are provided. Finally, an [OPC UA](#) client generation facility called UaoForQuasar is available for building C++ clients for [quasar](#)-based servers.

### 15.3.3 Network

The [ATLAS Control Network \(ATCN\)](#) integration of any [DCS](#) related network interface is implemented based on Ethernet with two main requirements. Firstly, [DCS](#) data should be exchanged between network devices with a minimum of network resources shared with other types of data by using dedicated switches exclusively assigned to [DCS](#) devices. Further, routing of data between [DCS](#) switches should be done with dedicated router hardware components. Secondly, individual [DCS](#) devices should be secured at the network level, e.g. by using [Virtual Local Area Network \(LAN\) \(VLAN\)](#) or private network gateways, allowing only connections to their respective [DCS BE](#) control stations and thus preventing unauthorised access.

## 15.4 DCS of TDAQ Sub-Systems

Introducing the new TDAQ hardware into P1 requires a well configured monitoring and control interface to ensure the system is in an operational state. Integrating these systems into the ATLAS [DCS](#) allows virtual access to the devices installed underground and in the counting rooms. Apart from the clear advantage of monitoring the state of the hardware and its hosting environment, the ability to keep track of the on-board sensors can help to detect a potentially upcoming hardware failure, and allow action to be taken prior to the system triggering a safety system interlock.

The TDAQ [DCS](#) includes all Level-0 Trigger subsystems – i.e. [L0Calo](#) ([eFEX](#), [jFEX](#), [gFEX](#) and [fFEX](#)), [L0Muon](#) (Barrel Sector Logic, Endcap Sector Logic, [MDT](#) Trigger Processor and [NSW](#) Trigger Processor), [MUCTPI](#), Global and [CTP](#) – as well as the [HTT](#) under the [EF](#) subsystem.

The TDAQ [DCS](#) strategy includes monitoring and controlling of the system environment, e.g. the crates in terms of power and cooling, as well as monitoring their critical parameters including: voltages, currents, power-consumption and temperature. Some control options are also included and can be executed with special expert access rights.

All TDAQ front-end subsystems are based on [ATCA](#) technology consisting of shelves and blades. With an [ATCA](#) shelf-manager, information is exchanged between [IPMCs](#) implemented on both the board and the shelf-manager allowing communication between the two over [IPMI](#). This covers basic shelf monitoring and control of e.g. fan speed, crate-slot occupancy, power etc. The [ATCA](#) blade [IPMC](#) carries out the minimal board management functions required by the [ATCA](#) standard, such as payload power control and indication

---

of general board health. The blades contain either [FPGAs](#) or [ASIC](#) chips (or both) for data processing, optical transceivers for data IO, power modules, DC-DC converters, temperature sensors and power-management chips. Most of the systems also contain a [SoC](#) for control and monitoring. Thus integration into the [DCS BE](#) is achieved through two paths, [SoC-to-DCS](#) and [IPMC-to-DCS](#), as described in Fig. 15.6:

- [IPMC-to-DCS](#) interface:  
A dedicated path to [DCS](#) is configured for traffic of [IPMCs](#) to the [DCS BE](#). Either board or shelf information is sent to and from the [ATCA](#) shelf-manager through a [SNMP-to-OPC UA](#) bridge configured on the [LCS](#) machine via [SNMP](#) communication.
- [SoC-to-DCS](#) interface:  
An [OPC UA](#) server constructed on the [SoC](#) using the [quasar](#) tool developed by the central-[DCS](#) team. Monitoring is established by communication between that server and an [OPC UA](#) client configured on the [LCS](#) machine where the information is transferred over TCP IP.

On-board hardware component supervision is done via the monitoring device configured as master of the [I<sup>2</sup>C](#) ([I<sup>2</sup>C](#)) bus of the board. Three possible scenarios are considered: in the first the [SoC](#) is defined as the master, in the second both the [SoC](#) and the on-board [IPMC](#) are defined as masters of two separate [I<sup>2</sup>C](#) networks, and in the rare third case that a [SoC](#) is not embedded, only the [IPMC](#) is defined as master. For the latter case, if additional monitoring of individual board components is needed, an [IPBus](#) interface is required. In all cases, the master retrieves the monitoring information over [I<sup>2</sup>C](#) and transfers it to the [DCS BE](#) via the two paths outlined above. Using a [SoC](#) is considered a more flexible approach in terms of possible extension of parameters to be monitored or controlled. Both the firmware and the software remains under full control of the developers; there are no third-party component dependencies and there are fewer limitations on the possible number of monitored parameters.

On the [DCS BE](#), the configuration of [LCS](#) clients, one per monitoring device is set: an [SNMP-to-OPC UA](#) bridge for communication with the [ATCA](#) shelf-manager and an [OPC UA](#) client for the [SoC](#).

The two paths to the [DCS BE](#) require several separate Ethernet ports for a single [ATCA](#) shelf, one for the shelf manager and several for the aggregated blade interfaces ([SoC](#) or [IPBus](#)), c.f. Fig. 15.6. All network devices for a given subsystem and the corresponding [LCS](#) are connected to at least one [ATCN](#) network switch reserved for the respective subsystem [DCS](#). In total, hundreds of network devices and a few [ATCN](#) network switches are required.

The [DCS](#) provides complete monitoring and control of the blades with human-machine interfaces for the respective hardware. In addition, the [ATCA](#) blades are equipped with power-management features, controlling the power distribution on the board, designed to take immediate action in case of over-current/voltage incidents. A dedicated interlock procedure is foreseen depending on the board functionality.



## References

- [15.1] A. Barriuso Poy, H. Boterenbrood, H. Burckhart, J. Cook, V. Filimonov, S. Franz, O. Gutzwiller, B. Hallgren, V. Khomutnikov, S. Schlenker, and F. Varela, *The Detector Control System of the ATLAS Experiment*, *JINST* **3** (2008) P05006.
- [15.2] B. Hallgren, H. Boterenbrood, H. J. Burckhart, and H. Kvedalen, *The embedded local monitor board (ELMB) in the LHC front- end I/O control system*, proceedings of the 7th Workshop on Electronics for LHC Experiments, Stockholm, Sweden (2001) .
- [15.3] P. Farthouat, *ELMB++ Proposal, Expert meeting on ELMB successor requirements*, 2015. [https://indico.cern.ch/event/535628/contributions/2176605/attachments/1278838/1898759/ELMB\\_Proposal.pdf](https://indico.cern.ch/event/535628/contributions/2176605/attachments/1278838/1898759/ELMB_Proposal.pdf).
- [15.4] *Versatile Link*, (online). <https://espace.cern.ch/project-versatile-link/public/default.aspx>.
- [15.5] O. Holme, M. Gonzalez Berges, P. Golonka, and S. Schmeling, *The JCOP framework*, . [Conf. Proc.C051010,WE2.1(2005)].
- [15.6] W. Mahnke, S. Leitner, and M. Damm, *OPC Unified Architecture*. Springer-Verlag, 2009.
- [15.7] S. Schlenker, D. Abalo Miron, B. Farnham, V. Filimonov, P. Nikiel, and C.-V. Soare, *quasar - A Generic Framework for Rapid Development of OPC UA Servers*, in *Proceedings, 15th International Conference on Accelerator and Large Experimental Physics Control Systems (ICALPCS 2015): Melbourne, Australia, October 17-23, 2015*, p. WEB3O02. 2015. <http://jacow.org/icalpcs2015/papers/web3o02.pdf>.



# 16 System Integration, Installation and Commissioning

In this Chapter initial considerations and planning on installation and commissioning are presented. The plans will mature into a set of documents that will define policies and procedures for each stage of the Installation and Commissioning:

- Documents describing Quality Assurance Policy and Quality Control Procedures, to be adopted during construction by each sub-system. The documents shall be part of the documentation to be presented in the [PRR](#) reviews.
- Documents describing System Integration and Validation Requirements and Specifications for each sub-system of the integration tests at CERN.
- A detailed Installation and Commissioning Plan document shall be prepared before the [PRRs](#) of the sub-systems, see Chapter 19, and a review of the installation plan will be coordinated with the Upgrade Project Office in Technical Coordination.

The chapter is organised as follows:

- Section 16.1 overviews general aspects of the installation and commissioning of the [TDAQ](#) upgrade elements in ATLAS.
- Section 16.4 focuses on validation tests and on the initial system integration of the custom hardware components of the three systems: specifically all the [L0](#) trigger components, the [HTT](#) modules of the [EF](#) system, the [FELIX](#) I/O cards in the [DAQ](#) system. The Section briefly covers QA/QC policies at the production sites, acceptance tests and initial commissioning on the surface in a dedicated setup hosted at a [TDAQ](#) Maintenance Facility, production firmware deployment and system integration tests.
- Section 16.5 briefly describes the installation aspects of the [TDAQ](#) Phase-II hardware in [USA15](#).
- The integration of the [TDAQ](#) components with the ATLAS detector systems in [USA15](#) and the early commissioning with beam is summarised in Section 16.6.

## 16.1 Overview of the TDAQ Phase-II Integration and Installation

The installation and integration of the [TDAQ](#) electronics in ATLAS is the conclusion of several years of development and construction. In general terms, the production of the different components spans over the years 2021-2024. Deliverables will be arriving at CERN at

different times and fully characterised in setups installed on the surface (see Section 16.4.3) before final installation. The sequence of operations for the integration, installation and commissioning of the TDAQ systems can be summarised in the following:

- Single board/unit acceptance tests
- System and/or sub-system integration tests and final validation on the surface
- Installation in ATLAS
- Standalone commissioning
- Detector integration validation and commissioning via dedicated calibration procedures
- Cosmic rays commissioning
- Early beam commissioning
- Operations

### 16.1.1 Location of the TDAQ Phase-II Upgrade Project's Components

Figure 16.1 is a schematic view of the buildings at Point-1 and the access shafts to the ATLAS cavern. The installation of the TDAQ Phase-II upgrade components is in the same locations as the existing Run 2 electronics:

- The counting room and service cavern USA15 houses the electronics racks and the services that needs to be installed close to the detectors. The room is accessible during data-taking operations and hosts all the L0 Trigger sub-systems components, the first elements of the DAQ readout and the HTT hardware.
- SDX1, located on top of the personnel access shaft and used for the personnel and material access control to the underground areas. It also contains the uninterruptible power supplies for the detector services, an electrical sub-station, and a TDAQ room. The TDAQ room is laid out on two floors and houses the Dataflow, the Networking components, and the EF Processing Farm units.

Table 16.1 details the exact location of all the TDAQ components. The motivation for selecting USA15 as baseline location of the HTT sub-system is twofold: (i) compatibility with a possible L0/L1 evolved architecture, where HTT may be used as a Level-1 Track Trigger, and, (ii) the limited space, power, and cooling available in SDX1. Data volume exchange between the EF servers in SDX1 and the HTT electronics in USA15 is relatively limited if compared with the network traffic required to stream the data from the Data Handlers to the Storage Handlers, and, therefore is not a significant concern.

### 16.1.2 Installation and Commissioning Planning

Installation and early commissioning will occur in the years 2025-2026. The installation procedures are relatively simple, as all the electronics is either implemented as ATCA blades or

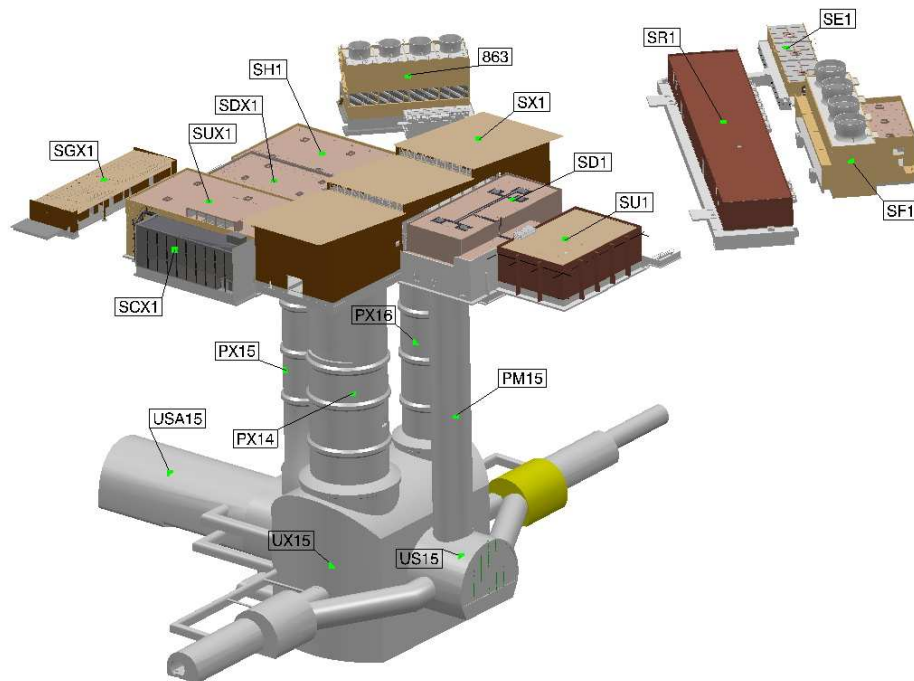


Figure 16.1: Layout of ATLAS Point-1 buildings on the surface and of the access shafts to the ATLAS cavern (from [16.1]). The TDAQ Phase-II upgrade systems are installed in the main counting room and service cavern ([USA15](#)) and in the surface.

as pluggable cards in standard rack-mounted PC server-based units. The installation and commissioning is relatively independent on the activities occurring in the detector cavern in [UX15](#). However, there are a few constraints to schedule installation tasks and work packages, which will require coordination with Technical Coordination, ultimately responsible for any installation task in [USA15](#). Namely:

- **Refurbishing of the counting rooms:** Both [USA15](#) and [SDX1](#) are planned to be renovated: (i) in [USA15](#) interventions to isolate acoustically the area with [ATCA](#) racks from the rest of [USA15](#) will be required by safety regulations; (ii) in addition, cooling, ventilation, and power distribution will also need to be upgraded; (iii) in [SDX1](#) floor support reinforcement will be likely needed as more servers and racks are to be installed.
- **Decommissioning of the legacy electronics racks:** Several of the legacy electronics of Run 2 and Run 3 will be decommissioned and disposed following ATLAS and CERN regulation.
- **Installation of new racks:** The number of crates and of PC server-based units required by all the ATLAS systems to comply with the new trigger and readout require-

## 16.1 Overview of the TDAQ Phase-II Integration and Installation

Table 16.1: Location in ATLAS of the TDAQ systems' and sub-systems' components

System	Sub-system	Component	Location
Level-0 Trigger	L0Calo	eFEX jFEX gFEX fFEX	USA15
	L0Muon	Barrel (RPC) Sector Logic Endcap (TGC) Sector Logic NSW Trigger Processor MDT Trigger Processor	
	MUCTPI	–	
	Global Trigger	–	
	CTP	–	
	TTC	–	
DAQ	Readout	FELIX Data Handler	USA15
	Dataflow	Event Builder Storage Handler Event Aggregator	SDX1
EF	Processor Farm		SDX1
	HTT		USA15

ments will exceed the current available rack space in USA15. ATLAS Technical Coordination is considering the installation of taller racks housing up to 3 ATCA shelves. In addition, deeper PC server racks may have to be installed to enable a wider choice of computing platforms. In fact the current racks in use in Run 2 are not compatible with all commodity-server form-factors.

- **Routing of fibres and cables:** New cable plants for most of the detectors will be required. Removal of the old cables, and the installation and routing of new cables and optical fibres will be a significant activity of Technical Coordination during Long Shutdown 3.
- **Rack Layout Configuration:** Optimisation of the rack layout and of the rack configuration is critical for the L0 Trigger system. TDAQ and Technical Coordination will coordinate the allocation of the ATCA racks to the systems, in particular to minimise contribution of cable routing and interconnects to the L0 trigger latency.

Analysis of the required interventions is on-going by Technical Coordination, and a detailed plan of the operations in USA15 will be available at the beginning of 2018. Figure 16.2 represents a preliminary top-level chart of the tasks to be completed by Technical Coordination before TDAQ can start the installation of the hardware in USA15. It assumes a worst-case scenario, in which the renovation of part of USA15 completes in approximately one year from the end of the Run 3 operations. It would limit the time for installation and commis-

sioning of the TDAQ systems, in particular for the Level-0 Trigger modular electronics in the tall ATCA racks, to 18 months. Technical Coordination is studying the possibility of advance in Long Shutdown 2 some of the tasks listed in Fig. 16.2, which would provide additional schedule float for installation, integration and commissioning.

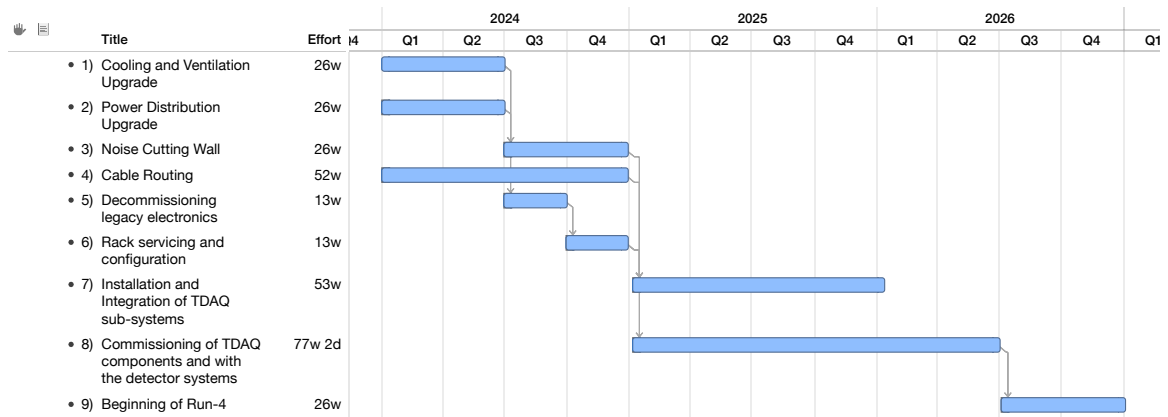


Figure 16.2: Initial planning by Technical Coordination for work packages and tasks related to the installation of racks and modular electronics in USA15. This preliminary chart assumes a worst case scenario where all the tasks will be executed at the beginning of the Long Shutdown 3 shutdown, and no access to the detector systems is allowed for approximately one year.

## 16.2 Laboratories and Facilities

Logistics at CERN will allow the TDAQ Phase-II Upgrade Project to conduct system integration tests in different laboratories and facilities. Two main laboratory spaces are foreseen for integration and validation of hardware components in ATCA format (Section 16.2.1) and for commodity server units (Section 16.2.2).

### 16.2.1 Building 4 surface testing facility

A new laboratory space is being refurbished under the responsibility of ATLAS Technical Coordination (see Figure 16.3). The laboratory is initially meant to be a surface test facility for the integration of the L1Calo Phase-I upgrade components, i.e. FEXs modules, RODs and HUBs. The initial configuration of the laboratory foresees 4 racks for a total of 50kW power distribution and proper water cooling/ventilation.

### 16.2.2 Lab4 Testbed Infrastructure

Figure 16.4 shows the rack layout in the Lab4 facility on Bldg. 4 in the Meyrin site. The laboratory is entirely dedicated to TDAQ developments: currently 16 racks are being equip-

## 16.2 Laboratories and Facilities

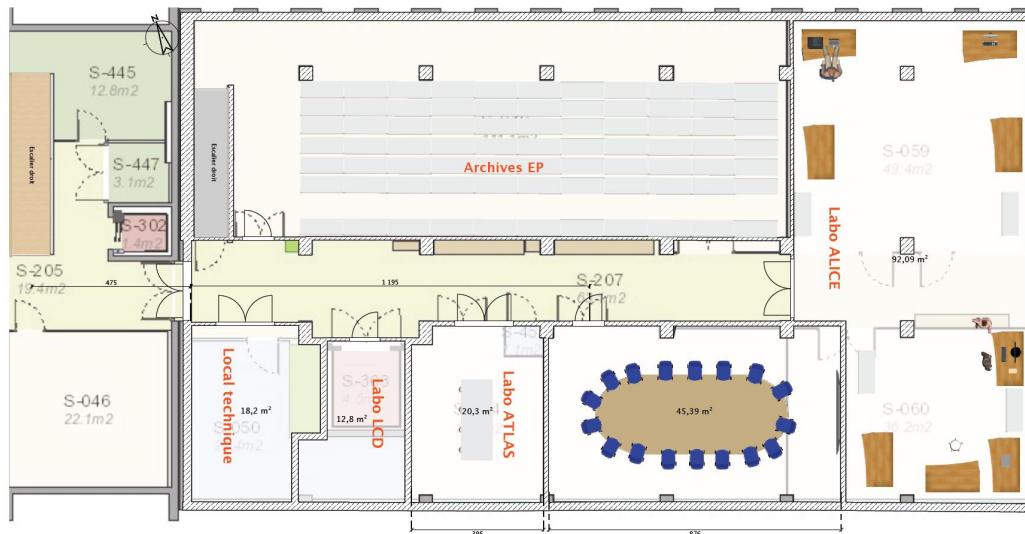


Figure 16.3: Floor view in CERN Bldg. 4 and the 20 m<sup>2</sup> lab space for Phase-I *L1Calo* integration tests

ped for testbeds of *DAQ* and *EF* tests. The facility may be expanded up to 24 racks, and it is being considered as an option for system integration and vertical slice tests of those *TDAQ* Phase-II components based on rack mountable commodity servers.

### 16.2.3 TDAQ Maintenance Facility

The Building 4 laboratory is a potential candidate to host the integration tests of all the *L0* Phase-II sub-systems before installation and commissioning in *USA15*. The lab space should allow to install up to 4-5 racks, each housing 2 *ATCA* shelves.

After completion of the installation in *USA15* the space will be converted to a Slice Test and Maintenance Facility (*TMF*) with minimal required setup for each system and sub-system. The size of each slice will be of a size sufficient to carry all the required sub-system tests, probably shelf-sized. The facility will be maintained alive for the duration of the *HL-LHC* operations to run long term tests, to develop and validate firmware and software before their deployment in *ATLAS* - see Section 16.3 and, ultimately, to troubleshoot hardware modules.

### 16.2.4 Safety

*ATCA* cooling systems, both in the horizontal and vertical cooling configuration, are extremely noisy. CERN HSE group, in collaboration with the EP-ESE group and *ATLAS* Tech-

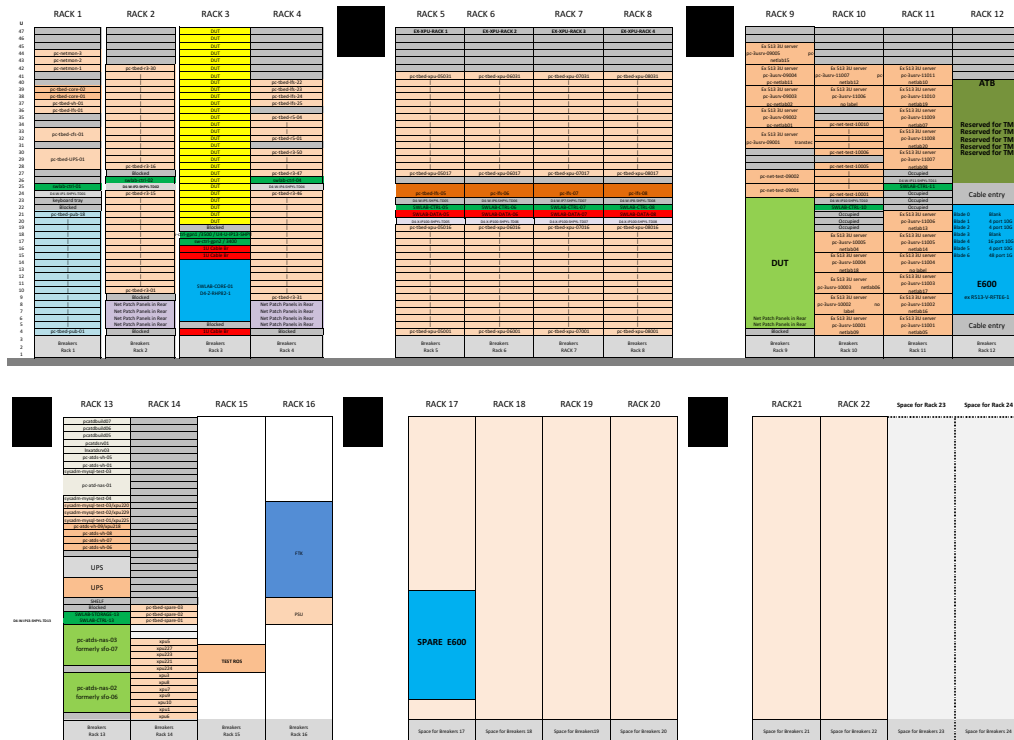


Figure 16.4: Current Lab4 rack configuration for DAQ/EF testbeds [16.2].

nical Coordination has conducted during 2016-2017 measurements in USA15 with racks equipped with two fully loaded shelves [16.3]. Measurements have been made with fans working in two different configurations: 75% of the maximum speed and full speed. Measurements in excess of 85 db(A) were recorded, above the safety level.

At the TMF, worse conditions are expected. Safety precautions become necessary and coordinated with the ATLAS Technical Coordination groups, as similar precautions will be needed also in USA15:

- Isolation walls to reduce the acoustic level noise and the health risks to the users need to be installed
- Access will be regulated and allowed for a limited time only to personnel who has followed proper safety training
- Wearing personal protection equipment will be required
- Yearly medical visits will be mandatory.



### 16.3 Production Firmware Deployment

In most of the [TDAQ](#) Phase-II sub-systems the level of complexity of firmware will increase significantly cf. to what experienced by the [TDAQ](#) for the Run 2 hardware or to what projected for the Phase-I upgrades. Firmware complexity is mainly driven by the need of implementing and integrating in [FPGAs](#) more sophisticated algorithms that reproduce the current offline performance. This is the main mitigation measure against the increased pileup.

Specific project management regulation shall be defined to ensure firmware is properly delivered well in advance for commissioning on the surface and in [USA15](#).

- Firmware shall be considered a CORE deliverable alike hardware components. A zero value will be assigned to firmware deliverable.
- Institutional responsibilities will be accounted for in the [MoU](#) document.
- Tracking of the Firmware developments shall be possible through the definition of milestones throughout the lifetime of the sub-system.
- Firmware design and implementation shall be reviewed in dedicated reviews, to be carried along the main hardware reviews, i.e. Specification, Preliminary ([PDRs](#)), Final ([FDR](#)) and Production Readiness ([PRR](#)) Reviews. The tight coupling between Firmware and Hardware reviews is required to both have a firmware development which allows to do all tests required at the time of the hardware review and to force similar time profiles for effort in hardware and firmware, aimed to reach a high level of overall readiness at the time of commissioning.
- No sub-system component shall be approved for progressing along the hardware design milestones without a full validation of all the firmware functionality required for that specific milestone, see Section [19.2](#).

It is imperative to establish very clear and rigorous procedures on firmware management and revision control. This is clear for example in the [Global Trigger](#) sub-system, because of its complexity in terms of architecture, and in terms of algorithmic firmware integration on both the [MUX](#) and [GEP](#) modules. About 20 firmware working packages have been identified by the level-3 managers, and, responsibilities are shared among several groups in different countries.

- The [TMF](#) infrastructure will be used as validation tool for any firmware block upgrade before it is deployed on the final boards and used during operations on the hardware in [USA15](#).
- For all the Level-0 sub-systems, a Production Firmware Deployment system will be designed and installed in [TMF](#) permanently. As an example, for Global, Production Firmware modules ([PFM](#)), see Section [9.4.6](#), will be designed with same hardware devices as the final [MUX](#) and [GEP](#) blades.



- 
- Each firmware block, developed at a remote site, will need to be commissioned and validated through extensive integration system tests in the production firmware deployment system.
  - Only after validation through the Production Firmware Deployment system, firmware builds integrating the validated blocks will be uploaded on the [USA15](#) modules and final commissioning tests will be conducted in [USA15](#) . The same policy shall adopted on pre-production boards for all the other [TDAQ](#) sub-systems.

The aforementioned guidelines will be adopted as general policy for firmware and software maintenance for the duration of the operations in the [HL-LHC](#) era, as it is currently being done with the [HLT](#) testbeds in Lab4.

## 16.4 Validation and Initial System Integration at CERN of custom Hardware sub-systems

### 16.4.1 Quality Assurance and Quality Control at the production sites

In several sub-systems multiple production sites are being planned. Each site will implement a local test facility to fully characterise their production share. By the time of the sub-systems' [PRR](#) a Quality Assurance Policy and Quality Control Procedures should be established and will be included in the documentation to be reviewed by the [PRR](#) panel.

The main objective of the QA/QC documentation is to ensure that each production site conforms to the agreed policies for what concerns the setup, infrastructure, material handling and tests sequences for the characterisation of the production modules.

QC procedures shall include:

- Visual inspection of the hardware delivered by assembly sub-contractors
- Basic functionality tests, e.g. powering sequence, configuration and monitoring
- Extensive behavioural characterisation of the hardware core and of its input/output interfaces, e.g. BER tests or eye-diagram verification of conformity to ATLAS-wide approved standards like IEEE 802.3bs or similar
- Predetermined test-vectors' validation to demonstrate functionality, or to test fault conditions
- Burn-in stress testing to minimise infant-mortality effects in the early commissioning phase

### 16.4.2 Acceptance validation tests at CERN

Acceptance tests, regulated by the same QA/QC policy, shall be repeated in the [TMF](#) facility once the hardware is delivered at CERN. The objective of these tests is to ensure reproducibility of the results, reliability of the testing procedures executed at the production sites, and of the deliverables' handling (including storage and transport). In addition, validation tests at receiving may need to include also system integration procedures see [Section 16.4.3](#), in particular for sub-systems whose functionality rely on critical communication between modules.

### 16.4.3 Level-0 Trigger Sub-system Integration Tests

Before installation in [USA15](#) each sub-system and each component in a sub-system shall undergo system integration tests. The tests are required to verify and validate communication protocols, interfaces and algorithms that require communication between modules in the system. The general strategy is to define a minimal required setup that reproduce a vertical slice of the sub-system. Such a vertical slice shall be built with production first articles and shall be kept available for the entire duration of the production, installation, commissioning and later for testing further developments of the L0 algorithms. Each production module will be tested through the vertical slice validation with pre-programmed test vector patterns that validate the module's functionality.

#### Phase-I FEXs

Legacy [L0Calo](#) processors ([eFEX](#), [jFEX](#), [gFEX](#)) will undergo at most firmware upgrades. The modules will be fully retested on surface, and system integration validation shall be required for the part of the systems which has interfaces to new LAr or Tile pre-processors and new optical fibres.

#### Forward FEXs

[fFEX](#) production is made only of a few modules in a single [ATCA](#) shelf. System integration on the surface shall be made with output interfaces to a dedicated readout unit and connections to the required Global [MUXs](#) modules.

#### [L0Muon](#) Sector Logic and [MDT](#) Trigger Processors

The vertical slice setup shall be composed by one fully loaded [SL ATCA](#) shelf, one [MDT](#) Trigger Processor shelf and with [MUCTPI](#), [CTP](#), [FELIX](#) first articles or pre-production units.

### Global

Global is a challenging project based on a design that require integration of several complex algorithms in a single [FPGA](#) device. Different stages to validate the system integration are foreseen:

- Firmware will be validated through the [PFM](#) modules (Sections [9.4.6](#), [16.3](#)) in the [TFM](#) facility
- A standalone vertical slice with two fully loaded [ATCA](#) shelves, for [MUXs](#) modules and for [GEPs](#) will be used to validate the time-multiplexing architecture.
- Integrated setups with a minimal sets of [L0Calo](#) processors, and with calorimeter pre-processors will validate the high-speed communication and the reconstruction algorithms of those

### 16.4.4 HTT specific Integration and Initial Commissioning

Integration of the [HTT](#) system before the installation of the production modules in [USA15](#) shall be realised through a vertical slice of the system, corresponding to two tracking units described in Section [13.1](#). The commissioning of the [HTT](#) slice will also require enough [EFPU](#) nodes to feed tracking requests at the expected rate.

The [HTT](#) integration will use:

- a group of [EF](#) nodes requesting tracking for [HL-LHC](#) events from offline Monte-Carlo simulations.
- two [HTTIF](#) modules routing the simulated tracking data to the [HTT](#) processors
- a fully loaded [ATCA](#) shelf with 12 [AMTP](#) blades
- a partially loaded [ATCA](#) shelf with 2 [SSTP](#) blades

The [HTT](#) described here is 2% of the overall system. It will become available with pre-production. An estimate for the required [EF](#) nodes is also 2% of the overall [EF](#). Two racks will be needed to contain the above [HTT](#) and [EF](#) equipment.

Track reconstruction performance and latency shall be among the validation parameters to be measured during the integration tests, before final acceptance and installation in [USA15](#).

### 16.4.5 DAQ Readout Sub-system: FELIX and Data Handlers

System integration tests will be organised for each detector system, as the [FELIX](#) I/O cards may require detector-specific firmware blocks, and data-handlers certainly require detector-specific software. Detector systems are responsible for the organisation of these tests and [TDAQ](#) will collaborate and support to ensure a full validation of their readout.

## 16.5 Installation in USA15

---

In general terms, no large system integration of the [FELIX](#) production I/O cards is planned. Validation tests will be carried at the level of individual readout servers after they will be equipped with the two I/O cards. In general a minimal vertical slice to validate system integration will include:

- one or more detector front-end electronics boards (up to a fully loaded shelf);
- a [FELIX](#) system;
- up to two Data Handler nodes;
- a monitoring node to analyse performance.

An exception is needed for [ITk](#). The detector system will be fully assembled and integrated in the SR1 surface building at Point-1 before being transported and installed in the ATLAS [UX15](#) cavern. The integration and standalone commissioning of the [ITk](#) readout system will be done in parallel, using 1/8, i.e.  $\sim 12.5\%$  of the final required [FELIX](#) cards, in SR1 together with the integration of the [ITk](#) detectors [16.4].

## 16.5 Installation in USA15

The [TDAQ](#) Phase-II installation in [USA15](#) will be preceded by two main tasks: (i) cabling of the fibres from the detector systems to the racks of the trigger processors and of the readout units, (ii) installation and configuration of the [TDAQ](#) racks.

### Cabling

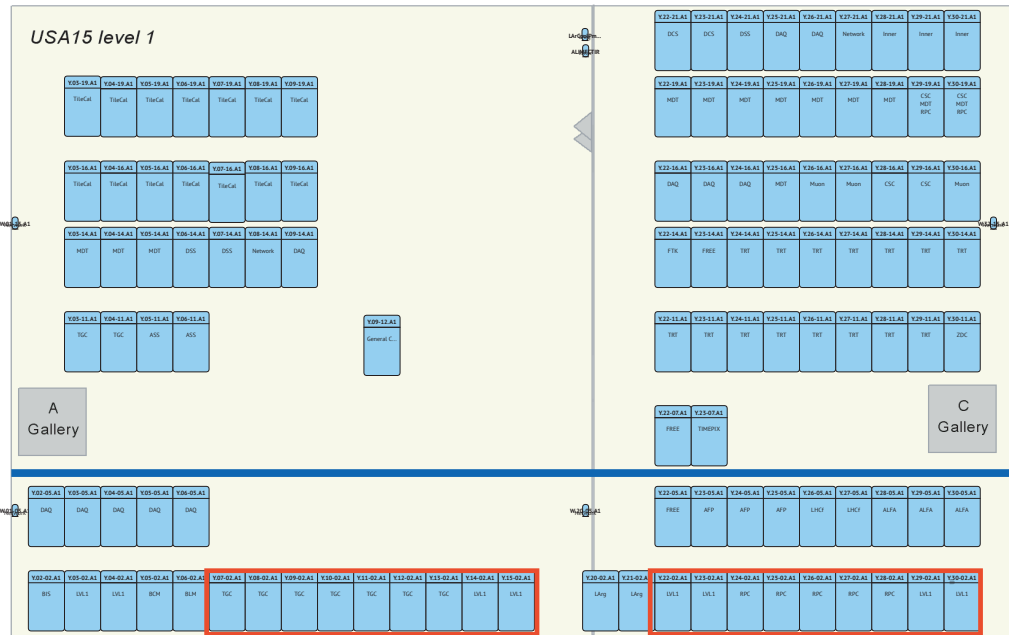
All the cables and optical fibres from [UX15](#) are in the scope of the detector system projects, while all optical fibres and optical plants internal to [TDAQ](#) are within this scope. Technical Coordination personnel will lay down and route all the cables up to the final destination point.

A cabling database will record each cables, the source and receiving terminal points, and its final length after connectorisation. Measurements of cable lengths, i.e. propagation delays will be conducted in coordination between the detector systems, Technical Coordination and [TDAQ](#) for all the cases where latency is relevant.

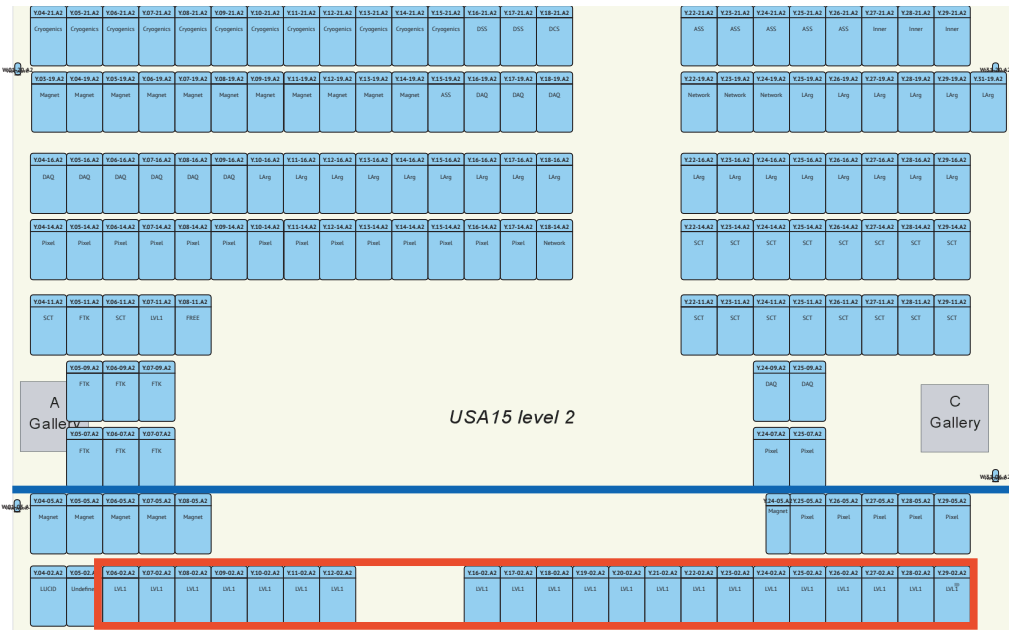
### USA15 Rack Layout

[USA15](#) racks are distributed in several rows over the two floors, as shown in Fig. 16.5, extracted from the Rack Wizard tool on Glance [16.5].

The front rows, shown in both Figs 16.5a, 16.5b below the thick blue line that represent the noise isolation wall, house a total 67 racks in the two floors of [USA15](#). Most of these



(a) Level-1 (top) floor of USA15



(b) Level-2 (bottom) floor of USA15

Figure 16.5: Current layout and rack configuration of the USA15 floors. Encircled in red the possible locations of the Level-0 Trigger electronics. In blue the noise-cutting wall to isolate acoustically the USA15 area with the ATCA racks

racks will be available in Phase-II for allocating the L0 Trigger sub-systems. As the readout electronics of many detectors will be decommissioned during Long Shutdown 3, most of the racks in the back rows will be reused for installing detector FELIX and Data Handler nodes.

The exact configuration of the racks is under discussion with Technical Coordination and all the ATLAS detector systems and it will be fully specified in the Installation and Commissioning Plan document.

### TDAQ sub-systems' configuration in USA15

Table 16.2 summarises the numbers of ATCA racks allocated in USA15 to all the TDAQ Trigger sub-systems. A total of 65 racks are required: 9 for the L0 Trigger system, 27 for the DAQ Readout nodes, and 28 racks for the HTT sub-system. A first estimate of the maximum power consumption per rack and the total power consumption of each TDAQ sub-system racks is given in the last four columns. The table indicates a range of power consumption: a current best estimate (CBE), based on an approximate bottom-up analysis of the components on a board, and a maximum possible value (MPV) that takes into account possible uncertainties on those estimates and margins assigned. The total power consumptions of those 61 racks is estimated to be between approximately 670 and 820 kW.

Table 16.2: Summary table with number of shelves, server units and racks in USA15 for the TDAQ system. First estimates of maximum power consumption per rack and the total of power consumption of the TDAQ sub-systems. Current Best Estimates (CBE) and Maximum Possible Values (MPV), which take into account known uncertainties and margins to add

Sub-system	No. of ATCA shelves or commodity servers	No. of racks	Max. Power per rack [kW]		Total Power/system /system [kW]	
			CBE	MPV	CBE	MPV
Level-0 Trigger System	25	9				
LOCalo	4		13	16	18	25
LOMuon	15		13	15	56	68
Global Trigger	4		13	16	24	29
CTP,MUCTPI	2		5	6	5	6
HTT	56	28	11	16	314	392
Readout	812	27	11	13	245	298
<b>TOTAL</b>		<b>65</b>			<b>662</b>	<b>818</b>

Such a layout assumes each Level-0 Trigger System rack to be configured with three (3) ATCA shelves, as shown in Figure 16.6, and a shelf with 12 user-blades, a shelf manager and, in some cases, an extra-switch. The HTT sub-system uses standard racks with two (2) ATCA shelves.

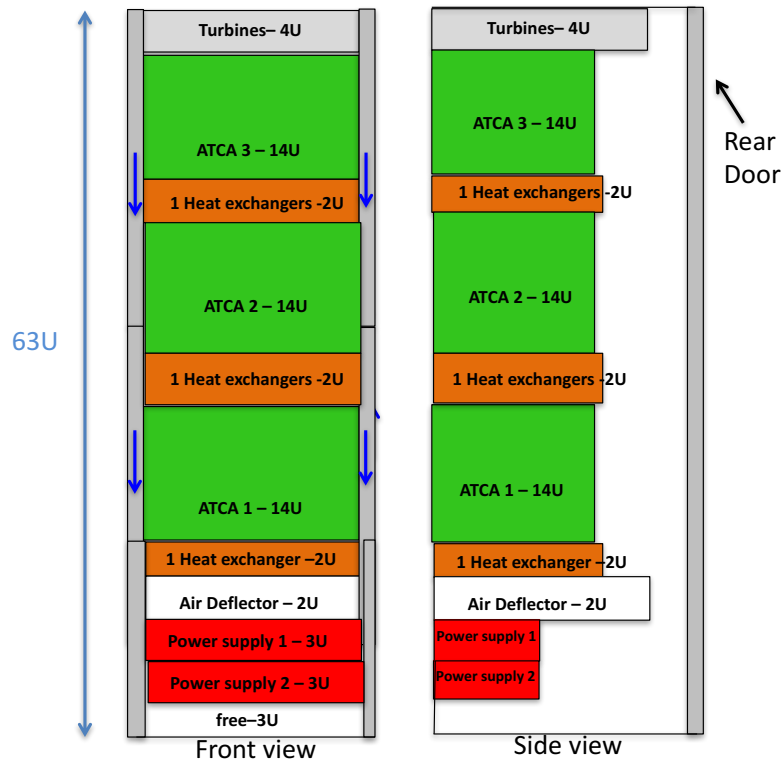


Figure 16.6: Rack layout with three ATCA shelves and three heat exchangers.

**LOMuon sub-system configuration** A possible arrangement of racks and configuration is shown, as example, in Tables 16.3 and 16.4 for the LOMuon sub-system.

Table 16.3: Example of Racks' configuration in the LOMuon sub-system. TP stands for Trigger Processor.

	SL Shelves	MDT TP Shelves	NSW TP Shelves	Rack Power	
				CBE [kW]	MPV [kW]
<b>LOMuon Endcap</b>				34	41
Rack #1	2	0	1	11	13
Rack #2	0	2 Large	1	13	15
Rack #3	2	1 Small	0	10	13
<b>LOMuon Barrel</b>				22	27
Rack #1	2 Large	1 Small	0	10	13
Rack #2	1 Small	2 Large	0	12	14
<b>TOTAL</b>				<b>56</b>	<b>68</b>

Table 16.4: Possible configuration of the ATCA shelves in the L0Muon sub-system

	No. of blades or shelves	Power/Unit CBE [kW]	Unc. [%]	Margin [%]	Power[kW] /Unit MPV [kW]
<b>L0Muon Endcap/TGC Shelf</b>		3.5			4.3
SL	12	0.20	10	10	0.24
Shelf Manager	1	0.15	10	10	0.36
Fan Tray	1	1.00	10	10	1.21
<b>L0Muon Endcap/NSW TP Shelf</b>		3.9			4.7
NSW TP	8	0.35	10	10	0.42
Shelf Manager	1	0.15	10	10	0.18
Fan Tray	1	1.00	10	10	1.21
<b>L0Muon Barrel/RPC Large Shelf</b>		3.5			4.3
SL	12	0.20	10	10	0.24
Shelf Manager	1	0.15	10	10	0.18
Fan Tray	1	1.00	10	10	1.21
<b>L0Muon Barrel/RPC Small Shelf</b>		2.7			3.3
SL	8	0.20	10	10	0.24
Shelf Manager	1	0.15	10	10	0.18
Fan Tray	1	1.00	10	10	1.21
<b>L0Muon /MDT Large Shelf</b>		4.4			5.3
MDT TP	12	0.20	10	10	0.24
MDT TP RTM	12	0.20	10	10	0.24
Shelf Manager	1	0.15	10	10	0.36
Fan Tray	1	1.00	10	10	1.21
<b>L0Muon /MDT Small Shelf</b>		3.3			4.0
MDT TP	8	0.20	10	10	0.24
MDT TP RTM	8	0.20	10	10	0.24
Shelf Manager	1	0.15	10	10	0.36
Fan Tray	1	1.00	10	10	1.21



---

## 16.6 Integration with the Detector Systems and Early Commissioning

The main commissioning phase begins when all infrastructure in [USA15](#) is installed and connected (including [ATCA](#) shelves, blades, fibre plant, [TTC](#) and readout fibres). Integration tests executed in the [TMF](#) laboratory will be repeated to assess the correct functions of the individual hardware components. Joint commissioning tests with the detector systems or among different sub-systems will be conducted to validate the correct interfaces and proper mapping between modules.

A common plan for schedule of the detectors' installation, of the commissioning of their electronics, and on their needs of early availability of the relevant [TDAQ](#) components will need to be defined early 2018, through dedicated meetings with Upgrade and Technical Coordination.

### Detector system calibrations

As part of the initial commissioning, the detectors will be organised in independent partitions and calibration scans will be executed on each partition. Calibration runs will obviously allow [TDAQ](#) and the ATLAS detectors to commission all the components of Readout and Dataflow sub-systems associated to a particular detector partition.

They will also allow commissioning of [L0](#) Trigger elements. This is particularly true for the calibrations of the [LAr](#) and [Tile](#) calorimeters. Precision calibration loops allow determining the energy scale at the calorimeter's cell-level, and, consequently, of the [L0Calo](#) – i.e. of the [FEXs](#) – modules and of the [Global Trigger](#) processors.

### Early commissioning with cosmics and with beams

Final validation of the systems in [USA15](#) will use cosmic rays triggers and high-rate stress tests to probe operational regions otherwise untested with calibrations. In general, cosmics data are sufficient to measure individual signals and overall system timings to the bunch-crossing levels. However, final tuning and verification of the trigger and readout chain will require [LHC](#) beam data, for example to achieve the ultimate timing performance.

Finally, once the commissioning strategies aforementioned and the parameters of the systems are well established and tuned, early data enable the study of the system performance and of the reconstruction algorithms deployed in the trigger processors. Rates and efficiencies of trigger object candidates can be studied in detail for example by tag and probe methods, and comparing online vs. offline object reconstructions.

## 16.7 Installation and Commissioning Plans of DAQ/EF components in SDX1

As outlined in Section 16.1.1 SDX1 is the surface building near the personnel access shaft. It hosts the underground access infrastructure, the TDAQ room and the Uninterruptible Power Supply (UPS) system. Technical Coordination is planning a renovation of the infrastructure to provide additional cooling capabilities to the ATLAS detector systems in UX15 and USA15, detaching the circuits dedicated to the ventilation of the surface building and the rack cooling in the TDAQ room, from the existing chilled and mixed water production stations, providing new cooling capability also to the racks that hosts the EFPU farms and the DAQ storage nodes.

The TDAQ room is laid out on two floors, equipped with approximately 100 standard racks for server applications, as shown in Figure 16.7:

A typical configuration of the SDX1 rack is shown in Figure 16.8. The front and rear views of the racks in the figure show 40 motherboards installed and the corresponding network interconnections.

### Installation and Integration of the EFPU and Dataflow nodes

The installation sequence and all the steps to integrate PC servers in SDX1 follow the well defined procedures used for the Run 1 and Run 2 operations, for example during the rolling replacement of the HLT PC servers.

For the processing nodes the procedure takes into account that the upgraded EFPU integrates to pre-existing nodes. It is summarised in the following steps:

- Units installed and cabled in the SDX1 racks
- A sys-admin burn-in tests of 24 hours
- A 24 hour burn-in with DAQ common software running
- A vertical slice tests simulating dataflow between a few racks commissions the nodes and the associated network
- Vertical slices tests gradually scaling up to the full system

Storage nodes in the Dataflow sub-system undergo similar procedures, but as storage will be completely replaced with new devices, a configuration step is required:

- Units installed and cabled in the SDX1 racks
- A sys-admin burn-in tests of 24 hours
- Configuration of the Storage handlers
- Vertical slice tests together with the EFPU nodes, scaling gradually to the full system.

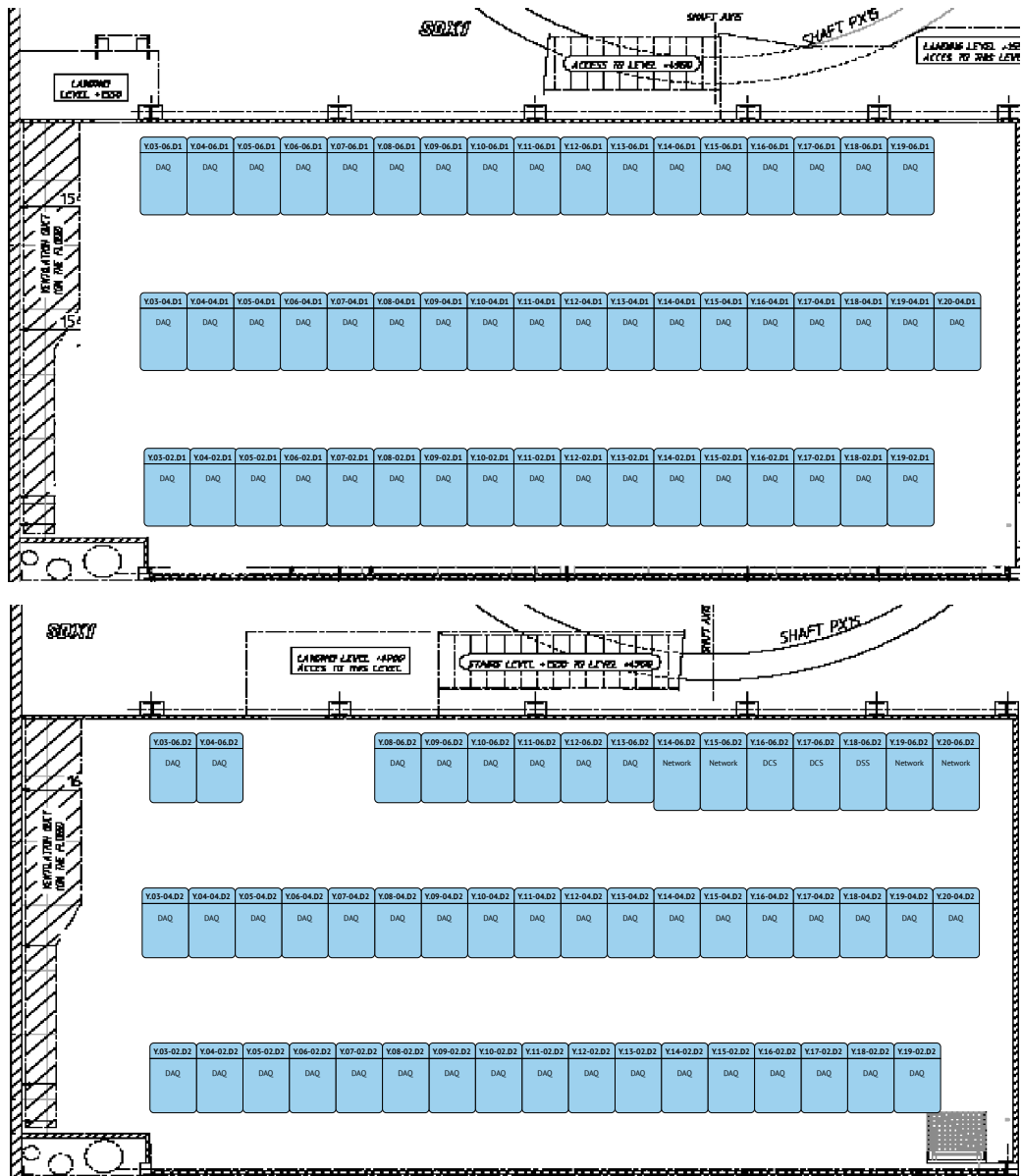


Figure 16.7: Current layout and rack configuration of the SDX1 Level-1 (top) and Level-2 (bottom) floors



Figure 16.8: Photographs of a typical event processing rack configuration in [SDX1](#). The left photograph shows the front side of the servers, while right one the rear network interconnections to the rack-level switches.

## Software Integration

Integration of the software and installation or upgrades of the operating systems of all the TDAQ nodes follow also a standard procedure:

- Development and pre-commissioning will be conducted in dedicated test nodes in the Lab4 facility
- On the SDX1 nodes Technical run periods are foreseen:
  - A. Technical runs will be conducted every 6-8 weeks for approximately one week.
  - B. During a technical run all the DAQ and EF resources will be available only for TDAQ commissioning, i.e. there will be no involvement of detector systems, and all the nodes will be run in emulation mode.
  - C. Software commissioning will be run on small partition initially to test the data flow.
  - D. Tests will be repeated scaling up the dimensions of a partition gradually to the full system.

## References

- [16.1] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST 3 (2008) S08003.
- [16.2] TDAQ TestBed, (online).  
<https://twiki.cern.ch/twiki/bin/view/Atlas/Lab4Testbed>.
- [16.3] C. Bortolin, *ATCA Cooling Report*, Tech. Rep. ATU-GE-ER-0015 v.1 / EDMS Doc. No 1609794. <https://edms.cern.ch/document/1609794/1>.
- [16.4] ATLAS Collaboration, *Technical Design Report for the ATLAS Inner Tracker Strip Detector*, Tech. Rep. CERN-LHCC-2017-005, CERN, Geneva, Apr, 2017.  
<https://cds.cern.ch/record/2257755>.
- [16.5] ATLS RackWizard, (online).  
<https://glance.cern.ch/atlas/rackwizard/menu>.



## **Part III**

### **Project Management and Organisation**





# 17 Project Management and Organisation

This chapter focuses on the [TDAQ Phase-II Upgrade Project \(UPR\)](#) organisational structure and summarises the basic project governance framework in which the project management plan is executed. The [TDAQ Phase-II upgrades](#) are developed and executed by a group of ATLAS Institutions that commit formally through a [MoU](#) document to design, construct, install and commission new components in the ATLAS Trigger and [DAQ](#) system. As such, the upgrade activities are embedded in the [TDAQ](#) system organisation, and ultimately integrated in its operation model. However, to control and administer the activities and the tasks in an effective and timely manner, the Phase-II upgrades are operated through a dedicated organisation, the [TDAQ Phase-II UPR](#), with a relatively autonomous governance and management structure, well defined responsibilities and reporting lines to the [TDAQ](#) system and, ultimately, to ATLAS.

The following sections describe the relationship between the [TDAQ UPR](#), [TDAQ](#), and ATLAS management entities in detail. Section [17.1](#) is an overview of the ATLAS organisation as it relates to the management of the upgrade activities. The structure and operation of [TDAQ](#) is summarised in Section [17.2](#). Finally, Section [17.3](#) details the structure, the governance and the management of the [UPR](#) at the top three levels of the organisation.

## 17.1 Overview of the ATLAS Upgrade Organisation

The overall organisation, the management and the integration of the ATLAS Phase-II upgrade projects is completely defined in an internal report, the ATLAS Upgrade Organisation document [[17.1](#)], released in December 2016. This section summarises briefly the essential points needed to understand the relationship of the [UPR](#) with ATLAS management and with the ATLAS detector system with which the [UPR](#) project is integrated.

For reference, a simplified view of the ATLAS Upgrade organisation is provided in Figure [17.1](#), extracted from [[17.2](#)] and updated to reflect the new ATLAS management organisation in place since March 2017.

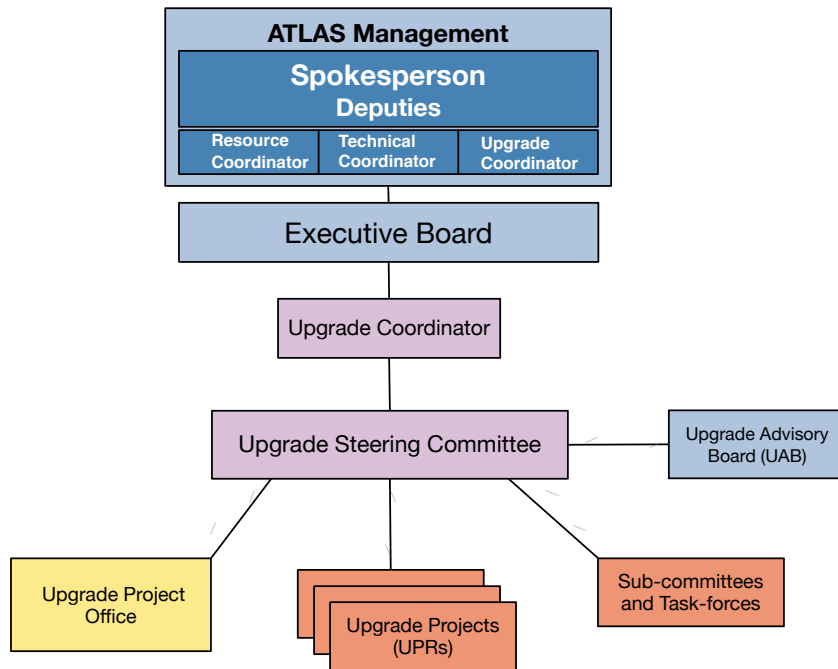


Figure 17.1: A schematic view of the ATLAS Upgrade organisation.

### 17.1.1 ATLAS management

The overall execution of ATLAS is the responsibility of the ATLAS Management, led by the Spokesperson, two Deputy Spokespersons, the Technical Coordinator, the Resource Coordinator, and the Upgrade Coordinator [17.3]. The Spokesperson and her/his Deputies have the overall responsibility to globally manage all the aspects of the ATLAS experiment. The Spokesperson represents ATLAS with respect to CERN, funding agencies and other outside bodies.

### 17.1.2 Executive Board

The Executive Board (EB) is the highest-level body for directing the execution of the ATLAS project and for direct communication between the ATLAS management and the systems. It is chaired by the Spokesperson with the Technical Coordinator as deputy chair. The EB includes: the ATLAS management, the System Project Leaders (Inner Detector, Liquid Argon and Tile Calorimeters, Muon, TDAQ, ITk Project, Forward Detectors), the Activity Coordinators (Run, Trigger, Data Preparation, Computing, Physics), and the Publication Committee Chair. Ex-officio members are the ATLAS Collaboration Board (CB) Chair, and

the **CB** Deputy-Chair. In addition, the Spokesperson appoints additional members-at-large to ensure an overall balance and competence in the **EB**.

### 17.1.3 Upgrade Coordinator

The overall steering and monitoring of the upgrade activities are delegated to the Upgrade Steering Committee (**USC**), which is a sub-committee of the **EB**, with an extended membership. The **USC** is chaired by the Upgrade Coordinator (**UC**), who is elected by the Collaboration Board (**CB**) with a renewable two-year appointment.

The **UC** follows each phase of the **UPRs**, from their initial R&D phase, through construction and into the system integration phase. The **UC** also appoints committees to assist with major decisions that cut across upgrade projects, or that have global impacts on the performance of any given upgrade project. The Spokesperson and the **UC** work closely together in preparing upgrade-related interactions with the CERN **LHCC** and **UCG** review committees.

### 17.1.4 Upgrade Steering Committee

The **UC** chairs the Upgrade Steering Committee (**USC**), the committee which is formally responsible for the oversight of the entire Upgrade programme. The **USC** operates in close collaboration with the **TC** and the **TC PO**, which links to the technical coordination and operation of the entire experiment.

The **USC** consists of: the **UC** as **USC** Chair, and a Deputy Chair designated by the **TC**, ATLAS management, the Upgrade Project Leaders (**UPLs**), the conveners of the Sub-committees, the Upgrade Physics Conveners, the Upgrade Software Coordinators, the Electronics Coordinator, the Review Office Chair, the Run Coordinator, the Computing Coordinator, the Physics Coordinator, and the Project Leaders of the main ATLAS systems.

The involvement of the System Project Leaders within the main decision-making body should ensure appropriate interaction with any corresponding **UPRs** affecting that System.

### 17.1.5 Upgrade Advisory Board

An Upgrade Advisory Board (**UAB**) is formed of ATLAS members who act as upgrade representatives of the communities supported by each Funding Agency (**FA**) contributing to ATLAS. The **UAB** is chaired by the **RC** and discusses primarily the availability of resources, and planning across the full spectrum of upgrades, including financial and workforce planning. The **UAB** member for a given **FA** may be either the National Contact Physicist (**NCP**) or a separate ATLAS collaborator representing the community supported by the **FA**.

### 17.1.6 Interactions with External Committees and Organisations

Formal reporting of ATLAS status and planning to the CERN Directorate, to the [LHCC](#) and to the [RRB](#) is among the responsibilities of the ATLAS Spokesperson. In the case of discussions with the [LHCC](#) related to upgrades, the Spokesperson may delegate or share the responsibility with the [UC](#). Contacts with individual ATLAS [FAs](#) are primarily the responsibility of the [NCPs](#) and/or handled by the [UAB](#) member for that [FA](#). However, the Spokesperson shall handle directly discussions with [FAs](#) and CERN Management whenever required, with the presence of the [NCP](#) and/or [UAB](#) member if possible.

### 17.1.7 Upgrade Projects

The upgrades of ATLAS are formally organised through Upgrade Projects. An Upgrade Project ([UPR](#)) comprises in general the plans and the activities of one of the existing ATLAS systems listed in Section 17.1.2. An [UPR](#) may be structured, managed and operated by an autonomous organisation, but its activities and plans should be developed in consultation and agreement with the parent system's management.

#### [UPR approval steps](#)

[UPRs](#) are approved in several stages. When a project has reached the level of maturity agreed by the parent system, with the [UC](#) and the [USC](#), the project passes first through initial stages of approvals as shown in Fig. 17.2: (i) an Initial Design Review ([IDR](#)) verifies that the technical requirements have been identified and that the initial design meets those requirements; (ii) a kick-off meeting is held by the Institute Board ([IB](#)) of the parent system to collect expression of interests from any institutes in ATLAS that wish to participate; (iii) based on the [IDR](#) committee's report and a recommendation by the [USC](#) members, the [UC](#) requests approval of the project by the [EB](#); (iv) the [EB](#) approval decision is endorsed by the ATLAS Collaboration Board ([CB](#)).

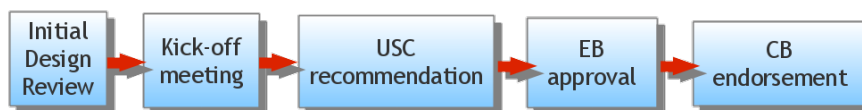


Figure 17.2: *Internal approval process of an Upgrade Project (UPR)*

#### [UPR Organisation and relationship to the parent system](#)

With the exception of [ITk](#), each [UPR](#) is assisted by the parent system's Institute Board ([IB](#)). The [IB](#) monitors the activities of the [UPR](#), in line with the [UPR's](#) [MoU](#). The [IB](#) takes decisions

---

in matters of resources and policy, and endorses technical decisions pertaining to the [UPR](#) recommended by the [UPR](#).

### **Upgrade Project Leader**

The [UPR](#) activities are coordinated by an Upgrade Project Leader ([UPL](#)), initially appointed by the System Project Leader and endorsed by the System [IB](#) until the approval of the project by the CERN [RB](#). After approval the [UPL](#) is elected by the System [IB](#) through the standard ATLAS election rules. The [UPL](#) reports on progress and achievements to the [USC](#), and to the parent System's management body.

### **UPR Management**

The [UPL](#) may propose a new management structure in consultation with the parent System's Project Leader, the [UC](#) and the System's [IB](#) Chair. Modifications to the [UPR](#) management structure shall be agreed with the [UC](#), and formally endorsed by the System's [IB](#).

### **Completion of the [UPR](#)**

After completion of the integration of the [UPR](#) into ATLAS, the [UPR](#) organisation dissolves and the parent system takes over the exploitation and maintenance of the delivered detector elements.

## **17.2 Organisation and management of the TDAQ systems and the TDAQ Upgrades**

The [TDAQ](#) system covers the design, development, construction, installation, commissioning and operation of all hardware and software components that perform the selection and acquisition of events from the ATLAS detector. The system organisation and its management structure is described extensively in Ref. [17.4] and the essential points are summarised in this Section. The following sections describe the continuing roles of the [TDAQ](#) Institute Board (17.2.1), the [TDAQ](#) Management Team (17.2.2), the [TDAQ](#) Steering Group (17.2.3), and how resources are controlled. In Section 17.2.5 the relationship of the Upgrade Project to parent system [TDAQ](#) system is described. The upgrade organisation and management plan is presented in Section 17.3.

The organigram of the ATLAS [TDAQ](#) system and its relationship to ATLAS management are shown in Figure 17.3.

## 17.2 Organisation and management of the TDAQ systems and the TDAQ Upgrades

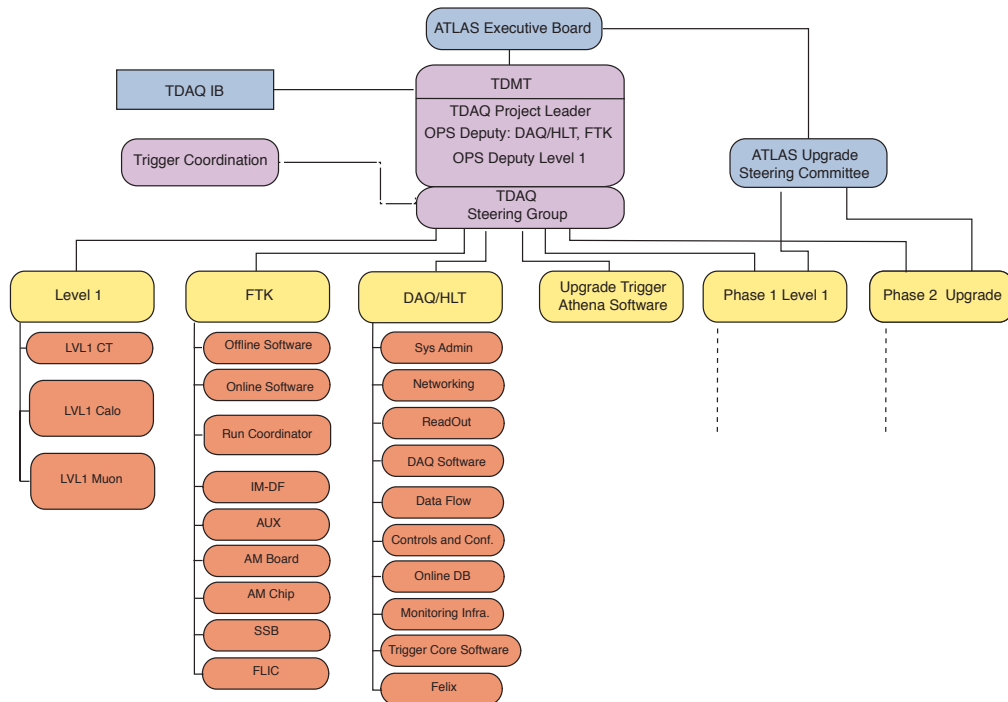


Figure 17.3: Organisation and management structure of the ATLAS TDAQ system and relationships to ATLAS.

### 17.2.1 TDAQ Institutional Board

The TDAQ Institute Board (TDIB) is the policy and decision making body of the ATLAS TDAQ system. It is subordinate to the ATLAS Collaboration Board. The TDIB takes decisions on sharing of resources and responsibilities. Major issues affecting the overall performance of ATLAS must be brought forward to the collaboration as a whole for decision in the ATLAS Collaboration Board. The institutions are the source of money and effort [17.5], and therefore all major questions involving sharing of responsibilities and contribution of resources have to be agreed upon by the TDIB. In major matters concerning resources the Institute Board shall consult the ATLAS Resources Coordinator.

Members of the TDIB are:

- The TDIB Chair
- A representative per institution contributing to the TDAQ system operations
- The members of the regular TDAQ Steering Group (ex-officio) - See Section 17.2.3
- The Chair(s) of sub-committee initiated by TDIB (ex-officio)
- The ATLAS Management (ex-officio)

## 17.2.2 TDAQ Management Team

The TDAQ Management Team (TDMT) is responsible to the TDIB for the execution of the TDAQ system. The TDMT is formed of the TDAQ Project Leader and her/his deputies. The Project Leader is elected by the TDIB, after nomination of candidates and due consultation within the TDAQ community, for a term of two years, and he/she can be re-elected with a 2/3 majority [17.5]. The Project Leader will appoint at least two deputies to share part of her/his responsibilities, and concentrate in complementary areas of TDAQ. These appointments will be approved by the TDIB. As indicated in Figure 17.3, at the time of drafting this TDR, the Project Leader has two appointed deputies, one for L1 Trigger operations, the other for the operation of the DAQ, HLT and FTK activities.

## 17.2.3 TDAQ Steering Group

The TDAQ Steering Group (TDSG) encompasses all the activities of the TDAQ system, and is composed of the people leading those efforts. The TDSG is the main body where executive decisions are taken on technical matters and it proposes recommendations to the TDIB on matters of sharing resources and responsibilities. The TDSG is chaired by the TDAQ Project Leader or by a delegated member of the TDMT. The frequency of TDSG meetings shall be sufficiently high to ensure the effective leadership of TDAQ system execution. The TDMT shall routinely consult the TDSG whenever significant issues arise.

The following itemised list defines TDSG membership:

- The members of the TDMT
- The leaders of TDAQ activities
- The TDIB Chair (ex-officio)
- The Chairs of sub-committees initiated by the TDIB
- The members at large chosen to provide balance and competence in the TDSG
- The ATLAS Run coordinators (ex-officio)
- The ATLAS Trigger coordinators<sup>1</sup> (ex-officio)
- The ATLAS Software & Computing coordinators (ex-officio)
- ATLAS Management (ex-officio)

## 17.2.4 TDAQ Resources

The TDIB Chair and the TDMT are responsible for resources matters, particularly the use of TDAQ funds and the sharing of financial responsibilities. The TDAQ funds include funds for maintenance and operation, including replacements, and non-CORE funds for supporting the ATLAS system. The use of past, current and foreseen TDAQ funds, as well

<sup>1</sup> In ATLAS, the Trigger Activity Area has primary responsibility for the Trigger Menu that is used during operations and the algorithmic software that is run in the HLT.

as the development of proposals regarding sharing of financial responsibilities need to be brought to the TDAQ Institute Board for approval. The TDIB Chair and the TDMT can decide together to appoint a person to assist in their task on resources matters. The appointed person needs to be endorsed by the TDIB. As most TDAQ M&O funds are provided by the full ATLAS Collaboration as Category A M&O, the ATLAS Resources Coordinator is consulted in all related matters.

### 17.2.5 TDAQ system and TDAQ Upgrade Projects

The TDAQ upgrade tasks will be executed by one more more ATLAS Institutions that commit to constructing, installing and commissioning new hardware and software components. These Institutions are, to a large extent, already members of the TDAQ. However, in compliance with ATLAS rules, any other ATLAS Institutions are eligible to join the UPR project activities and assume responsibilities within. Non-TDAQ Institutions are not required to formally join the system, but are invited to attend the regular TDIB meetings, to participate in the discussions and in the decision making process related to any upgrade activity, including for example the election of the Upgrade Project Leader (UPL).

The oversight and coordination of the ATLAS upgrade programme is detailed in Ref. [17.1] and outlined in Section 17.1. TDAQ upgrade R&D and project tasks form an integral part of the TDAQ system planning and are embedded within the TDAQ system. However, different strategies for the Phase-I and Phase-II upgrades have been employed.

The major construction work in the TDAQ Phase-I upgrade is related to the Level-1 Trigger, and falls under the responsibility of the Level-1 Upgrade Coordinator, appointed by the TDAQ Project Leader as described in the Phase-I TDR [17.6]. This person reports to the ATLAS USC together with the TDAQ Project Leader, and also to the TDAQ Steering Group. The TDAQ Project Leader directly reports to the TDIB on the status of the Phase-I upgrades.

The Phase-II TDAQ upgrade project has an elected Upgrade Project Leader (UPL), a dedicated internal organisation and an autonomous management structure for the duration of the project lifetime. Phase-II resources management is handled directly by the UPL and will be formalised through an independent MoU document. The UPL reports directly to the TDIB on the technical progress in the project and on any key resources aspect. The UPL will consult regularly with the TDAQ Project Leader and keep him or her apprised of any upgrade issues that impact operations, the Phase-I upgrade or which are relevant for the ATLAS Executive Board. An extended TDSG, which includes the regular TDSG members and the UPR management, is the main body where executive decisions are taken on technical matters, and is chaired by the UPL as documented in Section 17.3. After completion, the UPR organisation dissolves and the M&O of the delivered system, hardware, firmware and software is integrated into the overall TDAQ M&O activities. The ATLAS Management shall verify the institutional commitments of the non-TDAQ Institutions that have



---

joined the [UPR](#) after its completion, for example to the long-term M&O programme during operations in Run 4 and beyond.

## 17.3 TDAQ UPR Organisation and Upgrade Project Management Plan

The current (interim) management structure of the [TDAQ HL-LHC UPR](#) is shown in Fig. 17.4. This structure is responsible for directing and managing the [UPR](#) activities starting from the ATLAS approval of the [UPR](#).

The [UPR](#) is managed through the [TDAQ](#) Upgrade Project Leader ([UPL](#)) and her/his Management Team ([UPMT](#)); this team is responsible for the execution and the control of the tasks and of the resources required to ensure completion of the project activities. The executive decisions are made through the extended [TDSG](#) ([eTDSG](#)) group, which facilitates communications between the [UPR](#) management and the Level-3 managers, and ensures full integration of the [UPR](#) in the [TDAQ](#) system.

The [UPR](#) comprises three systems at Level-2 of the organisation structure: the Level-0 ([L0](#)) trigger, the Data Acquisition ([DAQ](#)), and the Event Filter ([EF](#)) & Performance systems. The organisational structure reflects the main [TDAQ](#) operation activities and the project's system-wide deliverables. Each system is organised at Level-3 of the management structure in sub-systems, which are functionally and organisationally independent sub-units corresponding to specific deliverables and tasks.

### 17.3.1 UPR Management Team

The [UPMT](#) is responsible for executing and directing all the [UPR](#) activities and coordination areas until all deliverables of the project have completed construction, installation and initial commissioning in ATLAS. Individual items will be turned over to the main [TDAQ](#) system as they are completed. Upon project completion all deliverables, the [UPR](#) is dissolved into the main [TDAQ](#) system operations.

**Composition:** The [UPL](#) is assisted upon appointment by four Level-2 manager/coordinators, who form together the [UPMT](#):

- Upgrade Project Leader [UPL](#) (chair)
- [UPR](#) Level-0 Trigger Coordinator (Deputy [UPL](#))
- [UPR](#) Data Acquisition Coordinator (Deputy [UPL](#))
- [UPR](#) Event Filter & Performance Coordinator (Deputy [UPL](#))
- [UPR](#) Project Office Chair

17.3 TDAQ UPR Organisation and Upgrade Project Management Plan

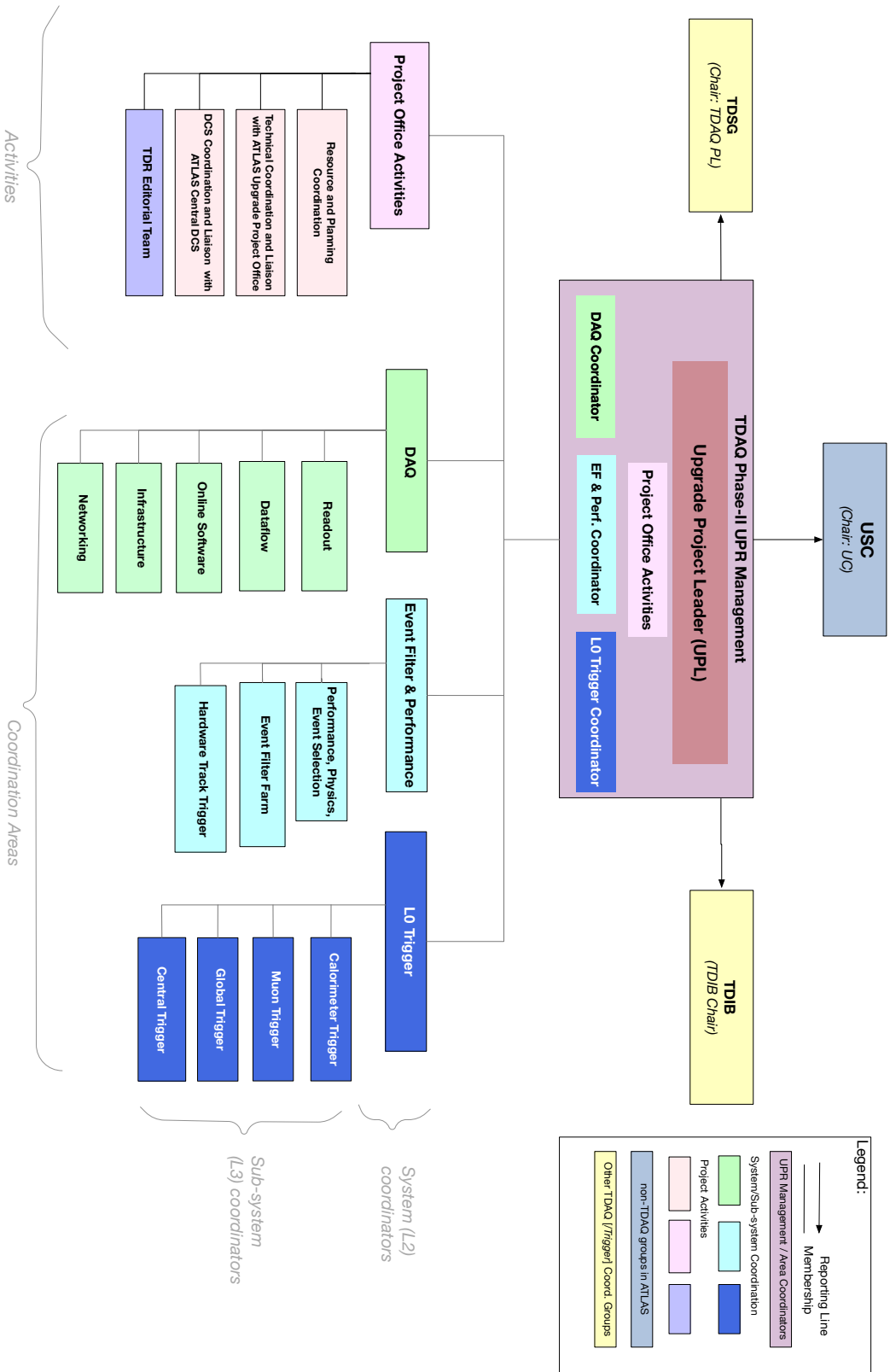


Figure 17.4: Organisation and management structure of the TDAQ HL-LHC UPR until CERN approval of the TDR, expected by April 2018. The detailed structure is subject to (minor) changes as the scope of the project finalises in the early phases of the TDR preparation.

**Mandate and Responsibilities:** The **UPMT** is the central element of the **UPR** organisation and is responsible for managing the overall project. It monitors the activities and the execution of all the project elements and of all coordination areas, and resolves all the integration aspects across the sub-projects. The responsibilities of the **UPMT** span all phases of the project, from R&D through detailed design, production, system assembly and integration, installation and commissioning.

### 17.3.2 Upgrade Project Leader (UPL)

**Term:** The **UPL** is elected by the **TDIB** for a renewable two-year term. The elections are regulated by the ATLAS and **TDAQ** system policy.

**Interim UPL (iUPL):** A first interim appointment by the **TDAQ** Project Leader was endorsed by the **TDIB** in September 2016 for the duration of the approval process of the **UPR**, i.e. until the **TDR** approval by the CERN **RB**. The expected completion of the term is April 2018.

**Mandate and Responsibilities:** Ultimately responsible for the **UPR**, and manages, in consultation with the **eTDSG**, its execution at CERN and with collaborating institutions to ensure that the project is completed within approved cost, schedule, and technical scope.

**Reporting Lines** There are three well defined reporting lines of the **UPR** management to **TDAQ** and ATLAS management:

- The **UPL** represents the **TDAQ** Phase-II **UPR** in the **TDSG** where he/she reports on the status of the project. The **UPL** and the **TDAQ** system's Project Leader will have frequent meetings to discuss planning, and technical status and progress.
- The **UPL** shall report to the **TDIB** on the status of the project, seek approval as needed on specific decisions proposed after a discussion within the **eTDSG**, and on any resources and financial matter of the project.
- The **UPL** shall also report regularly to ATLAS management, and, in particular, to the ATLAS **UC** through the **USC** and to the ATLAS **RC** on technical and resources-related matters, respectively, as outlined in Section 17.1.4.

### 17.3.3 Deputy Upgrade Project Leaders

One or more **DUPLs** are appointed by the **UPL** to assist in all matters relating to Project's activities and resources, including the planning, procurement, disposition of institutional

resources, progress reports on the executed activities, liaison with Institutional Representatives, fabrication and QA of TDAQ components, and their timely delivery for installation in ATLAS during LS3.

The DUPLs may be delegated, on a case by case basis, to represent the project to the ATLAS management or in CERN-wide review processes.

**Interim Mandate and Term:** In January 2017, the iUPL appointed the UPR L0 Trigger Coordinator, the DAQ Coordinator and the EF & Performance Coordinator as DUPLs for the whole duration of the interim mandate of the iUPL.

### 17.3.4 Extended TDAQ Steering Group

The Extended TDAQ Steering Group (eTDSG) is an extension of the TDSG (see Section 17.2.3); it is chaired by the UPL or by a delegated member of the UPMT. The eTDSG is the main body where executive decisions are taken on technical matters and it proposes recommendations to the TDIB on matters of sharing resources and responsibilities.

The following are members of the eTDSG:

- UPR Level-3 Coordinators/Managers
- TDAQ Upgrade Project Office chair
- TDAQ Upgrade Project Office Activities' Coordinators (see Section 17.3.8)
- TDIB Chair (ex-officio)
- Members of the TDMT (ex-officio)
- Members of the regular TDSG
- ATLAS Upgrade Coordinator (ex-officio)
- Technical Coordination's and Upgrade Project Office representative (ex-officio)

The UPL shall routinely consult the eTDSG whenever significant issues arise, and organise regular meetings whose frequency shall be sufficiently high to ensure an effective leadership of the UPR execution.

### 17.3.5 UPR Level-0 Trigger Coordinator

The UPR L0 Trigger Coordinator is directly responsible for the definition and the management of all the Trigger deliverables, as identified in the UPR PBS and WBS and related software activities.

He/she will oversee the R&D programme, design, fabrication, assembly and construction, installation and commissioning of the hardware L0 Calorimeter Trigger (forward FEX), the L0 Muon Trigger (RPC and TGC Sector Logic, MDT and NSW Trigger Processors) , the

Global Trigger (Aggregating stages, Global Processing Units), and the Central Trigger systems (CTP units, MUCTPI, TTC network distribution and interfaces).

He/she is also expected to manage a UPR Level-0 Trigger Coordination body as specified in Section 17.3.9.

### 17.3.6 UPR DAQ Coordinator

The UPR DAQ Coordinator is directly responsible for the definition and the management of all the DAQ deliverables and associated software activities, as identified in the UPR PBS and WBS specified in Sections 18.4 and 19.3, respectively.

The DAQ Coordination Group will work closely with the existing DAQ/HLT Coordination Group. The two coordination groups will have some common members. This choice is justified by several factors: (i) the limited size of the groups, (ii) the synergies with the existing activities in TDAQ operations and the tasks defined within and for the TDAQ Phase-I upgrades, and, most importantly, (iii) the fact that the DAQ/EF architecture and resources will evolve smoothly from the current Run 2 system through the Phase-I and Phase-II upgrades and the lifetime of the HL-LHC.

### 17.3.7 UPR EF & Performance Coordinator

The EF & Performance Coordinator is directly responsible for the definition and the management of all the EF deliverables, as identified in the UPR PBS and WBS, specified in Sections 18.4 and 19.3 respectively.

He/she will oversee as applicable the R&D programme, design, fabrication, assembly and construction, specification, procurement, installation and commissioning of the Event Filter Processing Units (EFPUs), possibly including Accelerator Processing Units, and of the Hardware-based Tracking for the Trigger (HTT). He/she will also oversee the design and development of the EF software, and of the Physics performance and Event Selection activity

The EF & Performance Coordination Group will work closely with the existing DAQ/HLT Coordination Group and Trigger Coordination Group. These coordination groups may have some members in common. This choice is justified by several factors: (i) the limited size of the groups, (ii) the synergies with the existing M&O activities and TDAQ Phase-I upgrades, and (iii) the fact that the DAQ/EF architecture and resources will evolve smoothly from the current Run 2 system through the Phase-I and Phase-II upgrades and the lifetime of the HL-LHC.

### 17.3.8 TDAQ Upgrade Project Office

The **TDAQ** Upgrade Project Office assists the **UPMT** with specific technical aspects of the project, to ensure a smooth integration with the **TDAQ** system and with ATLAS.

#### UPR Technical Coordinator

The principal responsibility of the **UPR** Technical Coordinator (**UPR-TC**) is to act as Technical Engineering Manager and Architect, focusing on global technical aspects of the **UPR**. The **UPR-TC** will assist the **UPL** chairing dedicated meetings of the **eTDSG** as the need arises. He/she will represent the **UPR** in the ATLAS **UPO** and, in general, with ATLAS **TC** and in discussion with other ATLAS detector systems. He/she will coordinate with the ATLAS Review Office for ATLAS-wide reviews.

#### UPR Project Resource Coordinator and Risk Manager

The principal responsibility is to assist the **UPL** on the Project Management aspects of the **UPR**, maintaining and monitoring the **CORE** costs, schedule and risk register. He/she will assist the **UPL** chairing dedicated meetings of the **eTDSG** or dedicated **eTDSG** sessions as the need arises. He/she will, in consultation with the **TDIB**, survey and monitor that the available person-power at each institute responsible for any **CORE** deliverables matches what is required.

#### UPR DCS Coordinator

The principal responsibility of the **UPR DCS** Coordinator is to guarantee the consistency of the Control, Configuration and Monitoring functions of the different hardware deliverables, and consistency with ATLAS central **DCS** strategies and plans.

### 17.3.9 Organisation of the TDAQ UPR Systems and Sub-systems

The Level-2 system (Level-0 Trigger, **DAQ** and **EF**) coordinators, and the Level-3 sub-system coordinators oversee the design, construction, installation, and commissioning of their sub-systems. Each system and sub-system is organised according to an individual structure proposed by the coordinators in consultation with the **UPL**. Figures 17.5, 17.6 and 17.7 show the organisation charts for the three systems in the **TDAQ UPR**.

Each Level-2 system has a coordination body chaired by the System Coordinator, and where each Level-3 sub-system is represented by its coordinator. In addition, the **DCS** and the Technical Coordinators are ex-officio members of each System coordination group to ensure

coherency across the whole **TDAQ UPR** of the implementation of the **DCS** functions in the hardware modular electronics, and to guarantee consistency with ATLAS guidelines regarding installation and integration in both **USA15** and **SDX1**.

Each sub-system is divided into specific activity areas, or tasks, that are mapped onto Level-4 and Level-5 items of the **WBS**. Coordinator Areas may have one or more conveners.

### Organisation of the Level-0 Trigger System

The Level-0 Trigger system is organised in four sub-systems and several coordination areas, as shown in the organigram in Figure 17.5.

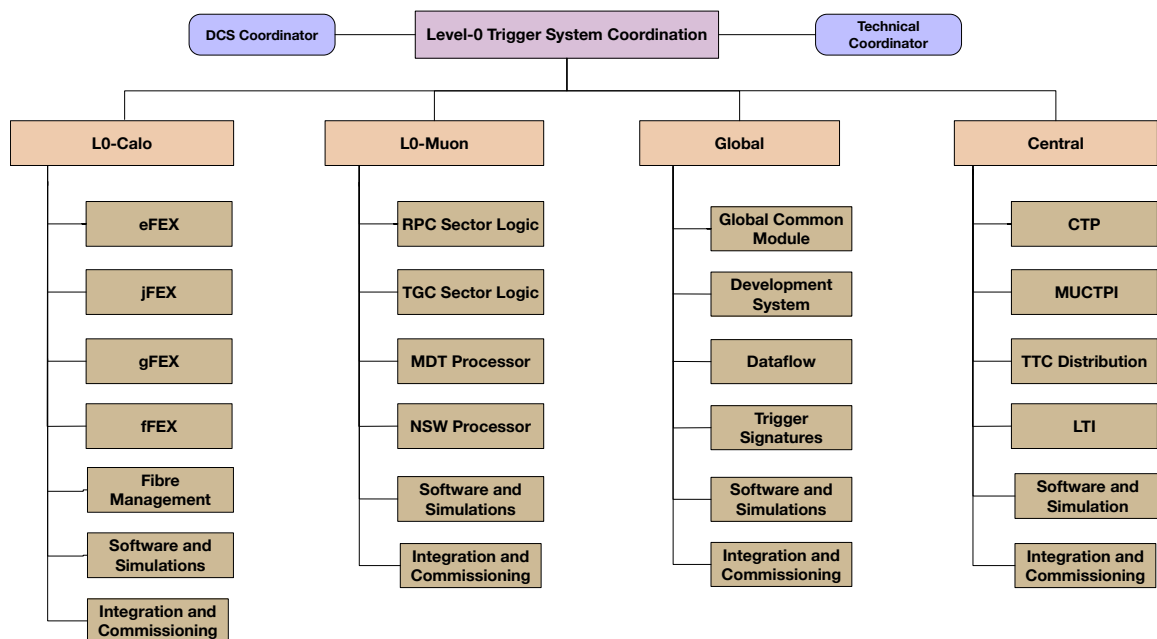


Figure 17.5: Organisation of the sub-systems in the Level-0 Trigger system

**Level-0 Calorimeter Trigger** The existing L1Calo group coordinates and organises meetings to discuss topics related to either Operations, Phase-I upgrades, or Phase-II upgrades of this sub-system.

**Level-0 Muon Trigger** Regular meetings are organised by the Level-3 coordinators among the conveners of each activity area, the experts, as well as the colleagues from the Muon Phase-II Upgrade Project to discuss issues related to the muon trigger that are relevant for both upgrade projects.

**Global Trigger** Two Level-3 coordinators manage the activity of the sub-system, together coordinating the hardware developments, the algorithmic firmware and software, and the core framework firmware. The sub-system also manages integration meetings with other sub-systems within the [TDAQ UPR](#) organisation and with other ATLAS upgrade projects on a regular basis.

**Central Trigger** The Central Trigger group manages and coordinates the activities for the [CTP](#), the [MUCTPI](#), the [LTI](#) and the [TTC](#) distribution. The group, also in charge of the existing central trigger system and its Phase-I upgrade, has tight connections with the CERN micro-electronics group, which, for example, has a key role in [TTC](#) developments.

### Organisation of the Data Acquisition system

The [DAQ](#) system comprises the Readout, Dataflow, Network and Online software sub-systems. The activities are managed by the associated Level-3 coordinators who work closely together with the existing [DAQ/HLT](#) coordination group, which is responsible for the Operations and Phase-I upgrades.

**Readout** The Level-3 coordinators manage the activities of the sub-system, including custom hardware development and commodity hardware procurement, firmware development, software development, and the integration and testing with detector systems. The Phase-I [FELIX](#) and swROD group will contribute to the R&D needed and integrate into the group.

**Dataflow** The Level-3 coordinator manages the activities of the sub-system and works with the existing Dataflow group for using the current infrastructure to perform evaluation and testing.

**Network** The Level-3 coordinator manages the activities of the sub-system and works with the existing network group for R&D activities, such as technology watching and technology evaluation.

**Online Software** The Level-3 coordinator manages the activities of the sub-system and works with the existing Configuration and Control group, and Monitoring group, for the long term development and the evolution strategy. Most of the online software elements will be maintained and upgraded adiabatically in the next several years by the [DAQ/HLT](#) coordination group. However, some of the core software infrastructure, may be significantly redesigned to benefit from the availability of new technologies. The chart in [Figure 17.6](#) shows only those specific Phase-II upgrade activities.



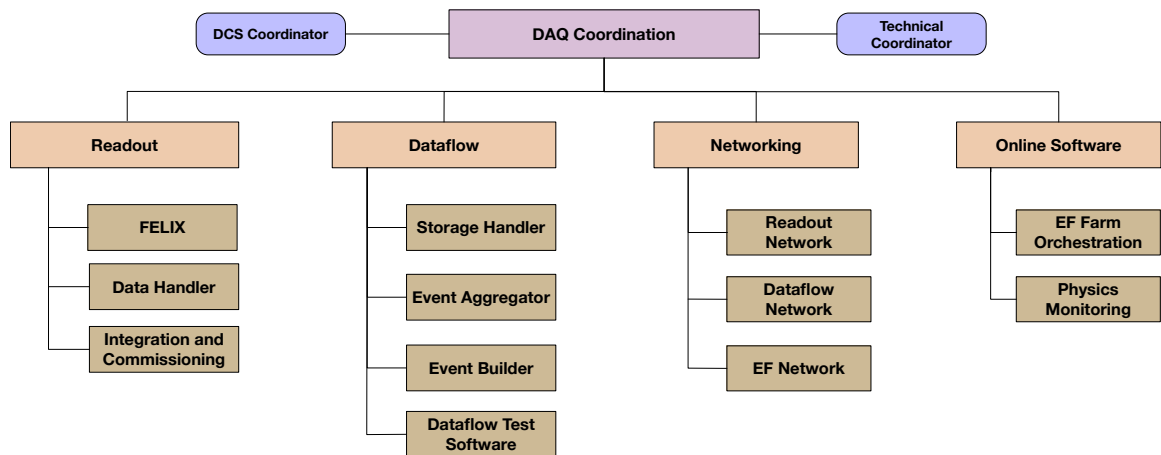


Figure 17.6: Organisation of the sub-systems in the *DAQ* system

### Organisation of the Event Filter & Performance System

The *EF* & Performance system comprises the *HTT* and *EF* sub-systems, and the Physics, Performance and Event Selection (*PPES*) activity. The Level-3 coordinators manage these three areas.

**Event Filter Software** Developments of the *EF* Tracking, Calorimeter, Muon and Core Software are organised within the relevant sub-groups of the existing Trigger activity area and its coordination group. Software developments and performance studies for conditions in Run 2, Run 3 and Run 4 are discussed in those forums. The Level-2 and Level-3 coordinators are active members of the Trigger groups, and shall be responsible (i) to guarantee that the developments of the trigger upgrade software are consistent with the *UPR* schedule, and (ii) to facilitate the contacts required with other upgrade subsystems.

**Event Filter processors** The computing requirements of the *EF*, calculated using a model described in Section 12.4, rely on performance information from all of the sub-groups. They will be periodically updated to check and if necessary revise the original estimates and ultimately as input to the procurement process.

Topics related to the *EFPU*s are discussed within the existing *DAQ/HLT* coordination group.

**Hardware Tracking for Trigger** The *HTT* sub-system has an established group with regular meetings where both hardware development, simulations and studies of the regional tracking and the global tracking scan are reported.

**Physics, Performance and Event Selection** The following activities are coordinated:

- Physics requirements capture
- Coordination of simulation of the TDAQ upgrades
- Performance studies of the overall upgrade and its subsystems
- Design of event selection strategies, including trigger menus and rate estimates

The group works closely with subsystems of the Level-0 Trigger and with the HTT and EF sub-systems to provide physics-motivated requirements and give performance feedback on their designs. These activities will continue to be organised periodically post-TDR, as required either by the UPR or external requests and initiatives such as a forthcoming CERN Yellow Report on HL-LHC. The group works also closely with members of the Trigger Activity area and its Coordination group for what concerns the strategies of the trigger menus in Run 4. It is expected that definite plans for the Run 4 trigger menus will come within the purview of the Trigger Coordination Group toward the end of Run 3.

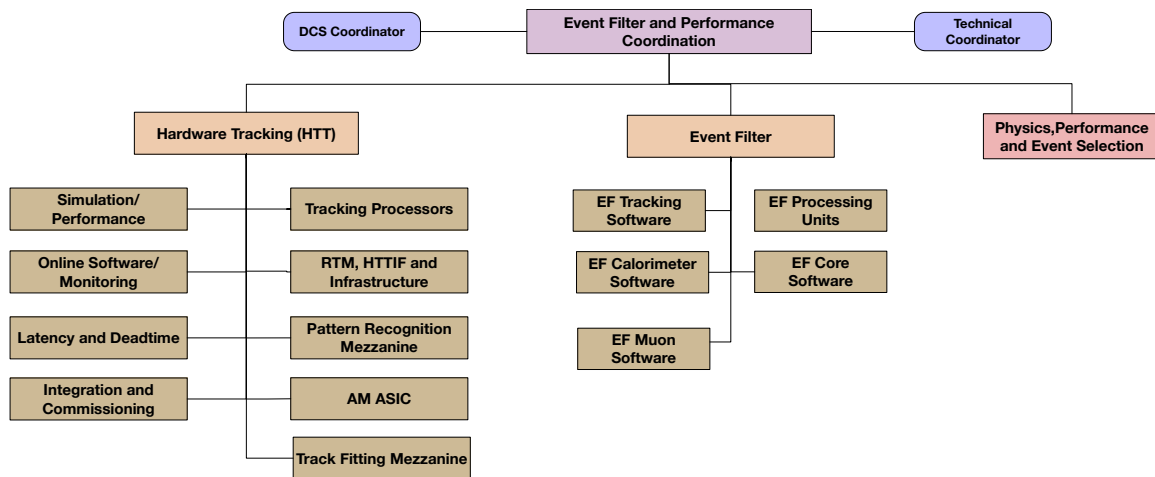


Figure 17.7: Organisation of the sub-systems in the EF & Performance system

## References

- [17.1] ATLAS Collaboration, *ATLAS Upgrade Organisation*, Tech. Rep. ATU-ORG-MG-0009 v.1 / EDMS Doc. No. 1815123, CERN, Geneva, Jun, 2017.  
<https://edms.cern.ch/document/1815123/1>.
- [17.2] ATLAS Collaboration, *ATLAS Phase-II Upgrade Scoping Document*, Tech. Rep. CERN-LHCC-2015-020. LHCC-G-166, CERN, Geneva, Sep, 2015.  
<https://cds.cern.ch/record/2055248>.

- [17.3] ATLAS Collaboration, *ATLAS Organisation*, Tech. Rep. ATLAS Internal Note GEN-NO-009, CERN, Geneva, Dec, 2007.  
[https://atlas.web.cern.ch/Atlas/private/ATLAS\\_CB/CB\\_Approved\\_Documents/A59\\_ATLAS\\_Organisation%5B1.2%5D.pdf](https://atlas.web.cern.ch/Atlas/private/ATLAS_CB/CB_Approved_Documents/A59_ATLAS_Organisation%5B1.2%5D.pdf).
- [17.4] The TDAQ Group Collaboration, *Constitution of the ATLAS Trigger/DAQ Institute Board*, <https://twiki.cern.ch/twiki/bin/view/Atlas/AtlasTdaqIbConstitution>.
- [17.5] ATLAS Collaboration, *ATLAS System Organisation*, Tech. Rep. ATLAS Internal Note GEN-NO-015, CERN, Geneva, Nov, 2006.  
[https://atlas.web.cern.ch/Atlas/private/ATLAS\\_CB/CB\\_Approved\\_Documents/A56\\_ATLAS\\_System\\_Organisation%5B2.1%5D.pdf](https://atlas.web.cern.ch/Atlas/private/ATLAS_CB/CB_Approved_Documents/A56_ATLAS_System_Organisation%5B2.1%5D.pdf).
- [17.6] ATLAS Collaboration, *Technical Design Report for the Phase-I Upgrade of the ATLAS TDAQ System*, Tech. Rep. CERN-LHCC-2013-018. ATLAS-TDR-023, CERN, Geneva, Sep, 2013. <https://cds.cern.ch/record/1602235>.



# 18 Project Cost Estimates

In this chapter the cost estimates of the [TDAQ](#) upgrades are presented. The project is organised in a set of deliverables that are summarised in the Product Breakdown Structure table in Section [18.4](#). The chapter is logically organised in two parts. A first general part introduces: (i) the Cost Management Plan in Section [18.1](#) summarising the functions and the processes adopted by the project to estimate the costs, define budgets, and to manage and monitor the costs during the [UPR](#)'s lifetime; and (ii) the cost policy and methodology, described in Sections [18.2](#) and [18.3](#), used for the [TDAQ](#) deliverables. The second part of the chapter is a collection of tables and figures presenting the cost estimate of each of the systems of the [UPR](#) (Section [18.5](#)) and the expected spending profiles (Section [18.6](#)).

## 18.1 Overview of the Cost Management Plan

The Cost Management Plan of the [TDAQ UPR](#) is based on the following fundamental functions/processes that are used to manage costs and budgets within the project.

**Cost Estimation** Cost estimation is the iterative process of developing, assembling and predicting, by successive approximations, the financial resources needed to complete the production of the project's deliverables. The cost estimating process prepares the cost tables and spending profiles, which are reviewed by the [UCG](#) group and are the baseline for the preparation of the [TDAQ](#) Memorandum of Understanding ([MoU](#)) document.

**Cost Budgeting** Cost budgeting is the set of accounting functions that establishes budgets and define the processes and the procedures for measuring, analysing and reviewing the forecasted and the actual costs. Budgeting is executed after the approval by the [RB](#) of the project and, once the [MoU](#) process has been completed.

**Cost Control** Cost control is the process concerning the collection, the analysis, the monitoring, the reporting and the managing of the costs on a regular basis. The [UPR](#) Resource Coordinator ([UPR-RC](#)), who reports directly the [UPR](#)'s Project Office chair, and, ultimately to the [UPL](#), will communicate and interact with the Institute's leaders to ensure, for example, that cost overruns do not endanger schedule or quality. The [UPR-RC](#) is also responsible

## 18.2 CORE Costing Policy

---

for controlling the **CORE** (see Section 18.2) cost estimates during the initial phases of the project, until a **MoU** document has been approved. Once an **MoU** document is in force, the **UPR** management will collect reports, for comparison with the **CORE** cost profile and help the institutes take corrective actions to achieve minimum variances. Changes to the **CORE** cost baseline profile will be documented along with an updated forecast for project completion.

**Internal Scrutiny and reporting to TDAQ** The **UPR** Resource Coordinator (**UPR-RC**) is responsible, within the **UPR** Project Office, for executing these functions during the lifetime of the **UPR**. Within the **UPR** Project Office, the **UPR-RC** reports on the Cost/Scheduling status of the project, which will be reviewed on a quarterly basis by the **UPL**, who ultimately approves the report within the project.

The **UPL** reports to the **TDIB** on a yearly basis on the status of costs and spending of the project. The **TDIB** formally approves the budget and the spending profile of the project within **TDAQ**. In addition, the **TDIB** chair may appoint a special Scrutiny Group within **TDAQ** an annual independent internal review of the financial status of the **UPR**.

**Reporting to ATLAS** Frequent interactions with the **ATLAS** management are used to monitor the status of the project. The **TDAQ UPR** report will be used by the **ATLAS** Resources Coordinator in his/her interaction with the **CERN RRB** and with the **RRB's** Scrutiny Group.

## 18.2 CORE Costing Policy

At the time of the original construction, the **LHCC** established a **COSting REview (CORE)** committee charged with reviewing the experiment's costs. A costing policy was established to define a metric that assigns cost values to each deliverable. Since then **CORE** equivalent costs are used to estimate and evaluate the effective costs of upgrade projects.

**CORE** cost does not represent the entire cost of the **TDAQ UPR**. There are non-**CORE** costs that are required to complete successful production and delivery of the components of the **TDAQ** project. It is the responsibility of each Institution participating in the **UPR** to obtain financial support from its Funding Agency for both **CORE** and non-**CORE** expenditure on the deliverables where the institution is responsible.

The policy for the evaluation of the **TDAQ UPR CORE** costs is based on a few key elements:

- 
- All production units that need to be installed in either [USA15](#) or [SDX1](#), and/or are required for the operations of the [TDAQ](#) system during [HL-LHC](#) shall be considered as *ATLAS deliverables* and their cost estimates shall be included in the [CORE](#) value of the project.
  - Specialised infrastructure, tooling and equipment required to maintain production hardware, firmware and software of the [TDAQ](#) system during operations at the [HL-LHC](#) are also *deliverables* and shall be included in the [CORE](#) costs of the project.
  - Evaluation of the [CORE](#) costs of a deliverable shall include, but not be limited to:
    - Cost value of the materials and components that define a hardware deliverable.
    - Non-recurring engineering ([NRE](#)) costs associated for example with [ASIC](#) production, [PCB](#) manufacturing or equivalent.
    - Allowances for production yields in both manufacturing, assembly, and QA/QC testing.
  - Pre-production items, manufactured and assembled with final tooling and labour to demonstrate production readiness, are to be considered [CORE](#) items and part of production.
  - Generic infrastructure costs, such as prototype and development test stands costs, shall not be part of [CORE](#).
  - Services and labour costs for manufacturing and QA/QC procedures performed by industrial subcontractors shall be included in the [CORE](#) estimates.
  - Labour and infrastructure at a production site of a participating Institute shall not be accounted in [CORE](#).
  - In general spares are not to be considered to be [CORE](#) items. As they represent failures occurring during operations, in general spares are to be considered part of the Maintenance and Operation (M&O) programme. Exception is made for those deliverables whose failure would cause a significant impact to the operations of ATLAS. In this case, spare units shall be included in the production and accounted for in the [CORE](#) estimates. These units shall be available at CERN for a possible emergency replacement in [USA15](#) for any failing module.
  - For applications where spares are accounted in [CORE](#), the fraction of spare units to be manufactured in addition to the regular production shall not exceed 5% of the total quantity, except in cases where a very limited number of units are produced.
  - No contingency shall be included in the [CORE](#) value of any deliverable. Institutions responsible for a deliverable are also responsible for the financial coverage and the additional costs that a deliverable might incur beyond what is specified in its [CORE](#) value.

### 18.3 Costing Methodology

A bottom-up methodology has been adopted for the estimate of the [CORE](#) values of the elements of the [TDAQ UPR](#). Detailed checking and scrutiny, firstly within [TDAQ](#), and, in

a second phase, by the Upgrade Project Office (UPO), has been performed on the estimates for each individual element of the UPR's systems and sub-systems.

- The quantities for each element of the upgrade have been calculated based on the architecture and design presented in this Technical Design Report.
- Each quantity is assigned an uncertainty reflecting possible changes of configurations, such as changes in interfaces because of the design evolution of a detector, new upgrade elements in the ATLAS upgrade scope, etc.
- Yield accounts for failure and loss during all the phases of production up to and including the installation of deliverables in the USA15 and SDX1 areas. A preliminary model of the production process is defined for a given deliverable, and a yield factor is estimated for each phase of construction based on past experience with projects of similar complexity and/or equivalent deliverables.
- Based on the policy given in Section 18.2, for mission-critical deliverables of the Level-0 Trigger system, some spare units are accounted in the CORE estimates. The fraction of spares to be considered in the production depends on the total production quantity for each sub-system, and shall be reported case by case in the Basis of Estimates (BOEs) documents associated to the UCG review that is scheduled in Q1 2018.
- Cost estimates for each item are quoted in the currency expected for the final procurements. Exchange rates to Swiss Francs (CHF) have been calculated as the average over the period of September 1st to November 30th, 2016, following the ATLAS convention used first in the ITk-strip TDR [18.1].

Table 18.1: Exchange rates used for the cost estimates. They represent the 3-month average over the period of September 1st to November 30th 2016, following the convention of the ITk-strip TDR [18.1].

Currency	Swiss Francs [CHF]
Euro [EUR]	1.085
British Pound [GBP]	1.246
U.S. Dollar [USD]	0.986
Japanese Yen [100 JPY]	0.942
Israeli Shekel [ILS]	0.2588
Chinese Yuan [CNY]	0.1461

- In general, values reported in this TDR reflect the estimates of a deliverable as if it were to be manufactured today. No de-rating factors are applied for procurements that will be executed in 3-5 years unless explicitly stated.
- An exception is made for the costs of the CPU-based commodity servers of the Event Filter system, as detailed in Section 18.5.3, where a performance improvement model is assumed as a function of time at a fixed cost.



- Similarly, assumptions are made on performance as function of purchase date when estimating costs of commercial off-the-shelf (COTS) integrated circuits (e.g. FPGAs), or high-speed serial computer expansion bus standards and related devices (e.g. speed of PCIe controllers).

### 18.3.1 Estimate Uncertainties

CORE costing is based on the Current Best Estimate (CBE) concept, i.e. on the information available at the time of the estimate, and it is naturally associated with an uncertainty. Several factors determine the level of cost uncertainty of an item: (i) the maturity of the technical development and design of a particular item, (ii) recent experience in other construction projects of items with similar technical complexity, (iii) the availability of vendor quotes through tendering processes or standard catalogues for commercial-off-the-shelf (COTS) components, (iv) understanding the procurement processes by the Institutions responsible for the deliverable item. To describe the level of uncertainty of an estimate, a quality factor ranging from 1 to 5 is used. The quality factors and the criteria are shown in Table 18.2: QF1 has the highest certainty and is based on a vendor quote for the final item or a catalogue price; QF5 has the lowest amount of certainty and is based on a rough estimate without any detailed design. For assembled items, i.e. comprising of several sub-assemblies or components with different quality factors, a cost-weighted average quality factor is calculated.

Table 18.2: *Quality Factor (QF) definitions used to estimate uncertainties on the CORE values of the project elements. QF1 has the lowest uncertainty, while QF5 has the highest one.*

Factor	Definition of the criteria based upon
QF1	(i) Items for which there is a recent (1 year max.) quote or catalogue price, based on a nearly completed design and for which there is more than one potential vendor (ii) Items that are a copy or are almost identical to an existing design for which there is a recent catalogue price or quote and for which there is more than on potential vendor.
QF2	Items that just fall short of satisfying the QF1 criteria: (i) Items that have only one potential vendor. (ii) Estimates based on a detailed, but not completed, design. (iii) Items adapted from an existing design with minor modifications. (iv) Items having quotes >1 year old, but deemed still to be sufficiently reliable based on experience.
QF3	(i) Items with quotes > 2 years old.

continued ...

## 18.4 UPR Project Breakdown Structure (PBS)

Factor	Definition of the criteria based upon
	(ii) Items whose cost estimates are based on a conceptual design or adapted from an existing design with extensive modifications. (iii) Items whose costs are estimated using physicist or engineering experience regardless of the maturity of the design.
QF4	Items that have unproven fabrication yields or for which there are unique issues, e.g. a special-order item and/or a single preferred vendor
QF5	Items that are still in a conceptual stage with no detailed specifications or design

## 18.4 UPR Project Breakdown Structure (PBS)

The **UPR** comprises a set of hardware, firmware and software deliverables that are organised in a hierarchically structured multi-levelled tree, the Product Breakdown Structure (**PBS**). The **PBS** mirrors the functional partition of the project into three (3) systems, twelve (12) sub-systems, and thirty-nine (39) components. The overall **TDAQ UPR** corresponds to the first level of the **PBS**, the three systems, i.e. Level-0, **DAQ** and **EF**, correspond to the Level-2 items in the **PBS** structure, while Level-3 items corresponds to sub-systems, e.g. the Level-0 Muon Trigger sub-system. In general, each Level-4 item in the **PBS** corresponds to one or more hardware deliverables. Firmware and software tasks related to the configuration, monitoring, control and operation of the hardware are included in the same **PBS** item.

Table 18.3 shows the **PBS**, expanded to the Level-4 of the tree structure, with a short dictionary entry for each item. A full **PBS** dictionary [18.2] down to Level-5/Level-6 of the structure is maintained in the “Project Management” section of the Electronics Data Management System (**EDMS**) folders dedicated to the **TDAQ UPR**, and is made available to the **LHCC** and **UCG** review panels through the **UCG** confidential material package.

The **PBS** seeds a Work Breakdown Structure (**WBS**), which defines the tasks and the processes to be executed for the design, prototyping, construction and the installation/commissioning phases of the **UPR**, as described in Chapter 19.

Table 18.3: *The TDAQ UPR Product Breakdown Structure (PBS).*

PBS Code	Item	Description
1	<b>TDAQ</b>	
1.1	<b>Level-0 Trigger</b>	Lowest-level, hardware-based Trigger system
1.1.1	<b>CTP</b> sub-system	Central Trigger Processor that forms the overall trigger decision and distributes timing information ( <b>TTC</b> ) to all sub-detectors

continued ...

<b>PBS Code</b>	<b>Item</b>	<b>Description</b>
1.1.1.1	CTPMI module	Interface of the CTP to LHC timing signals. It makes available timing signals on the ATCA back-plane for distribution to other CTP modules
1.1.1.2	CTPIN module	ATCA blade that accepts electrical and optical inputs for non-latency-critical signals to the CTP
1.1.1.3	CTPCORE module Processor	ATCA blade, implementing the core functions of the Central Trigger and is the source of signals to the TTC system
1.1.1.4	Infrastructure	ATCA shelf, patch panels and optical fibres
1.1.2	MUCTPI sub-system	Interface that combines and processes muon trigger information from the Barrel and Endcap Muon Trigger systems, and interfaces to the Global Trigger and to the CTP
1.1.2.1	MUCTPI module	ATCA blade handling one quarter of the muon trigger geometric coverage
1.1.3	TTC system	Trigger Timing and Control infrastructure
1.1.3.1	LTI	Local Trigger Interface module: ATCA blade that is an integral part of TTC and provides an interface between the CTP sub-system and FELIX or sub-detector-specific electronics through a Passive Optical Network (PON).
1.1.3.2	Infrastructure	ATCA shelves for LTIs, Passive Optical Network (PON) trees elements (splitters and fibres)
1.1.4	Calorimeter Trigger	This system defines Level-0 Trigger Objects comprising electron/gamma, tau, jet, large-R jet objects and global variable like $E_T^{\text{miss}}$ , and provide this information to Global trigger.
1.1.4.1	High-eta feature extraction	ATCA blade-based system, which defines electron/gamma, tau and jet Level-0 candidates in the Forward region
1.1.4.2	Fibre management	Optical fibre plant to handle Phase-2 interfaces to the calorimeters and to the Global trigger
1.1.5	Muon Trigger	System that defines Muon Level-0 Trigger Objects and provides this information to the MUCTPI.
1.1.5.1	TGC SL	TGC Sector Logic Board: an ATCA blade that defines Muon candidates for a Muon Endcap trigger sector and provides a seed for the MDT Trigger Processor; it subsequently combines the MDT trigger result with the RPC/TGC one and provides the information to the MUCTPI.

continued ...

## 18.4 UPR Project Breakdown Structure (PBS)

<b>PBS Code</b>	<b>Item</b>	<b>Description</b>
1.1.5.2	<a href="#">RPC SL</a>	<a href="#">RPC</a> Sector Logic Board: an <a href="#">ATCA</a> blade that defines Muon candidates for a Muon Barrel trigger sector and provides a seed for the <a href="#">MDT</a> Trigger Processor; it subsequently combines the <a href="#">MDT</a> trigger result with the <a href="#">RPC/TGC</a> one and provides the information to the <a href="#">MUCTPI</a> .
1.1.5.3	<a href="#">MDT</a> Processor Main Board	<a href="#">ATCA</a> blade that processes <a href="#">MDT</a> hits in regions identified by Barrel or Endcap <a href="#">SL</a> Board and provides Muon candidates transverse momentum and coordinates measurements through precision tracking to <a href="#">MUCTPI</a> .
1.1.5.4	<a href="#">MDT</a> Rear Transition module	Auxiliary module to an <a href="#">MDT</a> Processor Main Board
1.1.5.5	<a href="#">MDT</a> Sector processor mezzanine	Auxiliary module to an <a href="#">MDT</a> Processor main board; it performs the segment finding part of the tracking algorithm.
1.1.5.6	<a href="#">NSW</a> Trigger carrier board	<a href="#">ATCA</a> blade that is the main carrier board providing all services to the system.
1.1.5.7	<a href="#">NSW</a> Trigger Mezzanine	Auxiliary module to the <a href="#">NSW</a> Trigger Carrier Board, it contains all the MultiGigaBit interfaces and implements the Muon identification using the New Small Wheel detector data. It provides muon candidates to the Endcap Sector Logic.
1.1.5.8	Infrastructure	<a href="#">ATCA</a> shelves for all Muon Trigger Processors
1.1.6	Global Trigger sub-system	The Global Trigger receives and processes trigger information from the calorimeter and muon triggers. It also uses full-granularity calorimeter information to refine calorimeter objects and performs topology-based multi-object selections.
1.1.6.1	Common Module	An <a href="#">ATCA</a> blade, it is the basic hardware platform used for all Global trigger functional elements.
1.1.6.2	<a href="#">PFM</a>	Production Firmware deployment Module: <a href="#">ATCA</a> blade used as hardware platform for all algorithm and infrastructure developments
1.1.6.3	Infrastructure	<a href="#">ATCA</a> shelves and fibre plant needed to interface Global trigger elements
1.2	<b>DAQ</b>	The <a href="#">DAQ</a> system transports and stores data accepted by the Level-0 trigger for Event Filter processing.

continued ...

<b>PBS Code</b>	<b>Item</b>	<b>Description</b>
1.2.1	Detector Readout	The Detector Readout subsystem receives event data from detector front-end links and facilitates detector specific processing.
1.2.1.1	<a href="#">FELIX</a>	The Front-End LInk eXchange ( <a href="#">FELIX</a> ) routes detector data from custom serial links to a commodity network.
1.2.1.2	Data Handler	The Data Handler receives data from <a href="#">FELIX</a> via a commodity multi-gigabit network and hosts detector-specific processing.
1.2.2	Dataflow	The Dataflow subsystem buffers, transports, aggregates and compresses event data.
1.2.2.1	Storage Handler	The Storage Handler buffers data received from the Detector Readout subsystem through the Event Builder to decouple the Readout and Event Filter.
1.2.2.2	Event Aggregator	The Event Aggregator receives events accepted by the Event Filter, performs compression, and then prepares output files for transfer to permanent storage
1.2.3	Network	The network for all <a href="#">TDAQ</a> communications and data transport
1.2.3.1	<a href="#">FELIX</a> network	The network connecting the <a href="#">FELIX</a> servers and the Data Handlers with additional requirements for <a href="#">DCS</a> , control, and monitoring infrastructure inter-connection
1.2.3.2	Routers and fibres	The routers and fibres to be used for the data network and control network
1.2.3.3	Top-of-Racks (ToR) switches and fibres	The switch and fibres in an Event Filter Farm rack for connecting to the core router
1.2.4	Infrastructure	Infrastructure to facilitate the operation of all <a href="#">DAQ</a> PC units in <a href="#">SDX1</a> (racks and coolers for electronics in <a href="#">USA15</a> are under the responsibility of Technical Coordination.
1.2.4.1	Server Rack	Racks to host the PC servers and network switches
1.2.4.2	Rack cooler	Cooler to provide cooling in the server rack
1.3	<b>Event Filter</b>	The Event Filter ( <a href="#">EF</a> ) reduces the L0 trigger rate by two orders of magnitude through reconstruction and selection of events with commodity processors assisted by hardware tracking processors
1.3.1	Tracking Hardware ( <a href="#">HTT</a> )	The co-processor that will perform track reconstruction for the Event Filter.

continued ...

## 18.5 UPR CORE Costing Tables

PBS Code	Item	Description
1.3.1.1	AM Tracking Processor (AMTP)	The ATCA main card that will perform first-stage or second-stage processing
1.3.1.2	AMTP Rear transition Module (RTM)	The Rear Transition Module provides optical connections for the external interfaces of the TP
1.3.1.3	PRM	Pattern Recognition Mezzanine: The card where AM-based pattern recognition and the first stage track fitting will be performed
1.3.1.4	AM ASIC	The Associative Memory ASIC that will find the initial track candidates through pattern matching.
1.3.1.5	TFM	Track fitting Mezzanine: The card where first-stage track candidates will be extrapolated to second-stage layers, where track fits using all layers will then be performed.
1.3.1.6	Infrastructure	The ATCA infrastructure dedicated to HTT that will be common within TDAQ.
1.3.1.7	Hardware Tracking Interface	The interface between the EF processors and the HTT units
1.3.2	Processing Units	The commodity compute servers and relative infrastructure.
1.3.2.1	Servers	The commodity compute servers, which will perform the EF event processing.

## 18.5 UPR CORE Costing Tables

Table 18.4 presents the CORE cost estimates of the three systems constituting the TDAQ upgrade project.<sup>1</sup>

Table 18.4: CORE Cost summary of the TDAQ UPR project for each system.

PBS Code	System	CORE Cost [kCHF]
<b>1</b>	<b>TDAQ Total</b>	<b>44,880</b>
1.1	Level-0 Trigger	10,393
1.2	DAQ	13,640
1.3	Event Filter	20,847

<sup>1</sup> Commodity items in the current system that will not be changed for the upgrade and that are not listed in the tables in this section will be retained for Phase-II and are subject to the rolling M&O replacement policy.

### 18.5.1 Level-0 Trigger

The **CORE** costing of the Level-0 Trigger system is summarised in Table 18.5 and shown as a pie chart in Figure 18.1. The estimates reflect the design detailed in Chapters 7- 10. The **L0Muon** and **Global Trigger** sub-systems are the main cost drivers within the Level-0 Trigger system because of the large number of **ATCA** blades and of the large processing **FPGAs** required. The highest uncertainty in the cost of the sub-system components is due to the uncertainty in the estimate of the processing resources needed, and consequent **FPGA** choices. QF3 is in general assigned to all the trigger processors. QF4 is used for the **Global Trigger GEP** and **MUX** modules, since new generation **FPGAs** that are expected to be available on a 2-3 year timescale should provide substantially more processing power at fixed costs. The **UCG** confidential material includes Basis of Estimates' documents that detail and justify the values shown in Table 18.5 down to the Level-5 of the **PBS**.

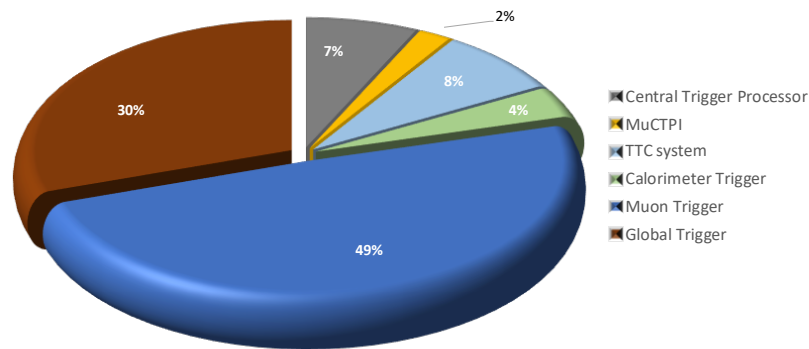


Figure 18.1: **CORE** cost distribution of the Level-0 Trigger System.

Table 18.5: **CORE** Cost Estimate of the Level-0 Trigger system in the **TDAQ UPR**. **CORE** values of the system's components represent Current Best Estimates.

<b>PBS Code</b>	<b>Item</b>	<b>Cost [kCHF]</b>	<b>Cost Quality</b>
<b>1.1</b>	<b>Level-0 Trigger System</b>	<b>10,393</b>	
<b>1.1.1</b>	<b>Central Trigger Processor</b>	<b>726</b>	
1.1.1.1	CTPMI	58	QF3
1.1.1.2	CTPIN	116	QF3
1.1.1.3	CTPCORE	501	QF3
1.1.1.4	Infrastructure	51	QF1
<b>1.1.2</b>	<b>MUCTPI System</b>	<b>213</b>	

continued ...

<b>PBS Code</b>	<b>Item</b>	<b>Cost (CBE) [kCHF]</b>	<b>Cost Quality</b>
1.1.2.1	<b>MUCTPI</b> Modules	209	QF2
1.1.2.2	Infrastructure	4	QF1
<b>1.1.3</b>	<b>TTC system</b>	<b>817</b>	
1.1.3.1	<b>LTI</b>	643	QF3
1.1.3.2	Infrastructure	174	QF1
<b>1.1.4</b>	<b>Calorimeter Trigger</b>	<b>389</b>	
1.1.4.1	High-eta feature extraction	357	QF3
1.1.4.2	Fibre management and Infrastructure	32	QF2
<b>1.1.5</b>	<b>Muon Trigger</b>	<b>5,111</b>	
1.1.5.1	<b>TGC SL</b> Board	1,479	QF3
1.1.5.2	<b>RPC SL</b> Board	1,024	QF3
1.1.5.3	<b>MDT</b> Processor Main Board	939	QF3
1.1.5.4	<b>MDT RTM</b>	118	QF3
1.1.5.5	<b>MDT</b> Sector Processor Mezzanine	636	QF3
1.1.5.6	<b>NSW</b> Trigger Carrier Board	127	QF3
1.1.5.7	<b>NSW</b> Trigger Mezzanine	607	QF3
1.1.5.8	Infrastructure	181	QF3
<b>1.1.6</b>	<b>Global Trigger</b>	<b>3,135</b>	
1.1.6.1	Common Module	2,422	QF3
1.1.6.4	<b>PFM</b>	493	QF3
1.1.6.5	Infrastructure	220	QF2

### 18.5.2 Data Acquisition System

The **CORE** cost estimate of the **DAQ** system is shown in Table 18.6 and in Figure 18.2. The estimates reflect the design detailed in Chapter 11. The detector readout costs are driven essentially by the number of links required to read out the the ATLAS detector systems at the specified trigger rate of 1 MHz and for an event size including 200 additional events from pileup. Dataflow and network costs are essentially determined by the ATLAS event size expected at the **HL-LHC** (see Section 11.5). The **UCG** confidential material includes Basis of Estimates' documents that detail and justify the values shown in Table 18.6 down to the Level-5 of the **PBS**.

### 18.5.3 Event Filter

The **CORE** cost estimate of the **EF** system is shown in Table 18.7 and in Figure 18.3. The estimates reflect the technical design presented in Chapters 12 and 13 for the **EFPU** servers and for the custom-hardware tracking co-processors, respectively.



Table 18.6: *CORE* Cost Estimate of the *DAQ* system in the *TDAQ UPR*. *CORE* values of the system's components represent current best estimates.

<b>PBS Code</b>	<b>Item</b>	<b>Cost [kCHF]</b>	<b>Cost Quality</b>
<b>1.2</b>	<b>DAQ</b>	<b>13,640</b>	
<b>1.2.1</b>	<b>Detector Readout</b>	<b>6,186</b>	
1.2.1.1	FELIX	4,262	QF3
1.2.1.2	Data Handler	1,924	QF3
<b>1.2.2</b>	<b>Dataflow</b>	<b>4,034</b>	
1.2.2.1	Storage Handler	4,006	QF3
1.2.2.2	Event Aggregator	28	QF3
<b>1.2.3</b>	<b>Network</b>	<b>3,315</b>	
1.2.3.1	FELIX network	916	QF2
1.2.3.2	Routers and fibres	2,109	QF3
1.2.3.3	ToR switches and fibres	290	QF3
<b>1.2.4</b>	<b>Infrastructure</b>	<b>105</b>	
1.2.4.1	Server Rack	45	QF3
1.2.4.2	Rack cooler	60	QF3

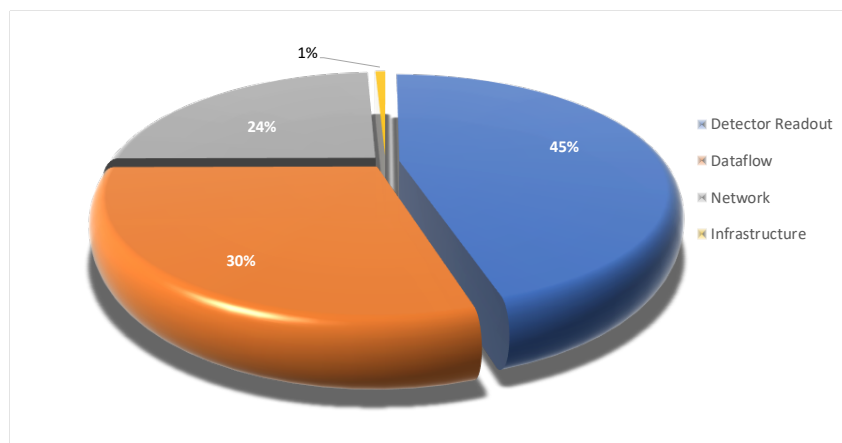


Figure 18.2: *CORE* cost distribution of the *DAQ* System.

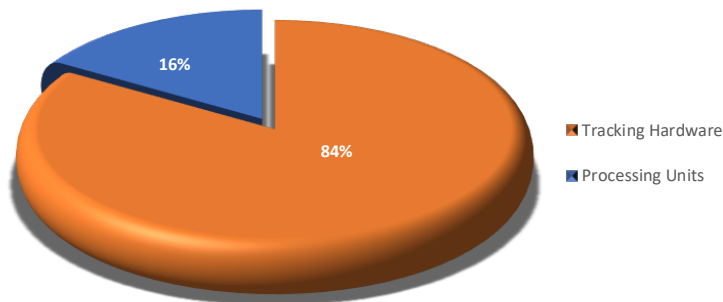


Figure 18.3: CORE cost distribution of the EF System.

Table 18.7: CORE Cost Estimate of the EF system in the TDAQ UPR. CORE values of the system’s components represent current best estimates.

PBS Code	Item	Cost [kCHF]	Cost Quality
<b>1.3</b>	<b>Event Filter</b>	<b>20,847</b>	
<b>1.3.1</b>	<b>Tracking Hardware</b>	<b>17,448</b>	
1.3.1.1	Tracking Processor (TP)	6,451	QF3
1.3.1.2	TP Rear transition Module (RTM)	695	QF2
1.3.1.3	Pattern Recognition Mezzanine (PRM)	4,678	QF3
1.3.1.4	AM ASIC	3,293	QF4
1.3.1.5	Track fitting Mezzanine (TFM)	1,261	QF3
1.3.1.6	Infrastructure	658	QF2
1.3.1.7	Hardware Tracking Interface	411	QF2
<b>1.3.2</b>	<b>Processing Units</b>	<b>3,399.0</b>	
1.3.2.1	Servers	3,399.0	QF3

The UCG confidential material includes Basis of Estimates’ and additional documentation that detail and justify the values shown in Table 18.7 down to the Level-5 of the PBS. The remaining sections just briefly outline a few considerations on costs for the different components of the EF sub-systems.

### Tracking Hardware

The dimensions of the system described in Chapter 13 determine the quantities in the first part of Table 18.7. The FPGAs of the 672 TP main boards and in the 1152 PRM mezzanines are the cost drivers of the HTT sub-system. The cost uncertainties are determined by the

---

choice of the [FPGAs](#) in the [TP](#) modules and the [PRM](#) and [TFM](#) mezzanines, where a quality factor QF4 is assigned. The [ASIC](#) costs are driven by the [NRE](#) costs, which account for two separate mask-sets for pre-production and production, and include a 70% yield production. An overall quality factor QF3 is assigned.

### Event Filter Processing Units

The [CORE](#) cost estimates of the [EFPU](#) sub-system are based on the technical considerations detailed in Chapter 12. A total of  $\sim 4.5$  MHS06 have been estimated to be required for the [EF](#) system at the ultimate performance configuration of the [HL-LHC](#). Two factors contribute to the [CORE](#) cost estimates: (i) the CPU cost extrapolations, together with the projected hardware configuration of the commodity servers in 2026; and (ii) the continuation of the M&O rolling replacement foreseen to continue throughout the lifetime of the experiment.

**CPU cost extrapolations:** The latest CERN IT extrapolations indicate computing costs between 1.4 and 3.3 CHF/HS06 in 2026, as shown in Figure 18.4. The extrapolations are based on past purchases, assuming an exponential decrease. Similar extrapolations, based on the ATLAS [TDAQ](#) purchases, provide a compatible figure of roughly 2 CHF/HS06. This type of extrapolations have shown a large variability. For example the 2013 CERN IT analysis indicated a compute cost of 1.1 CHF/HS06 in 2023, significantly below the current figures. The variability probably reflects the market-driven nature of the compute technology sector. Considering the CERN IT and ATLAS [TDAQ](#) results, as well as their evolution in the past few years, a value of 2 CHF/HS06 was chosen as estimated compute cost in 2026.

**M&O Rolling-replacement:** The [TDAQ](#) M&O policy to partly replace computing resources at the end of the 5-year lifetime of a server defines a rolling-replacement procedure where approximately a third of the [EF](#) farm is replaced every 2-3 years. In the event of a long-shutdown of the [LHC](#) complex, the rolling-replacement is foreseen for the first year when the [LHC](#) operations are resumed. To achieve the projected 4.5 MHS06 required for the operations at the ultimate performance configuration of [HL-LHC](#) 1133 servers should be installed by the end of the [LS3](#) shutdown in addition to the rolling replacements, which is the cost reported in Table 18.7.

## 18.6 Costing Profiles

Figure 18.5 shows the spending profile projected in the calendar years 2018-2026. The stacked columns represent the sum of the three systems the upgrade project is comprised of: the Level-0 trigger (blue), [DAQ](#) (green) and the [EF](#) (yellow) systems.

## 18.6 Costing Profiles

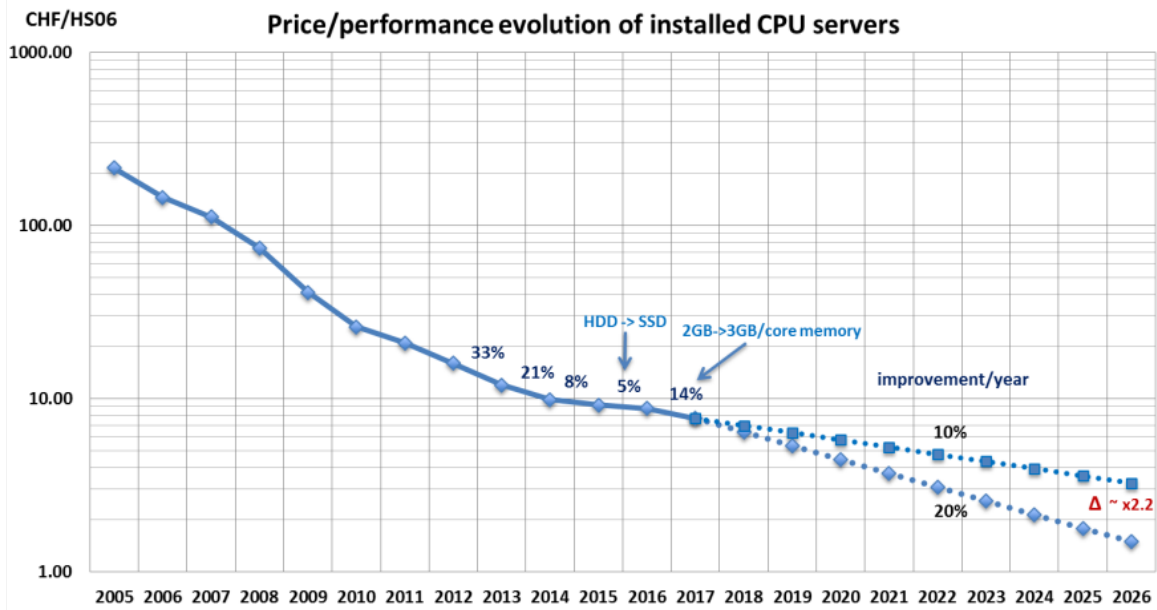


Figure 18.4: Preliminary extrapolation of CPU costs based on CERN procurements [18.3]. Two extrapolation curves are given: a pessimistic scenario with a 10% annual reduction, and a scenario considered realistic with a 20% annual reduction of the CPU cost per HS06 unit. An overall factor 2.2 difference between the two scenarios is calculated at the end of the LS3 shutdown.

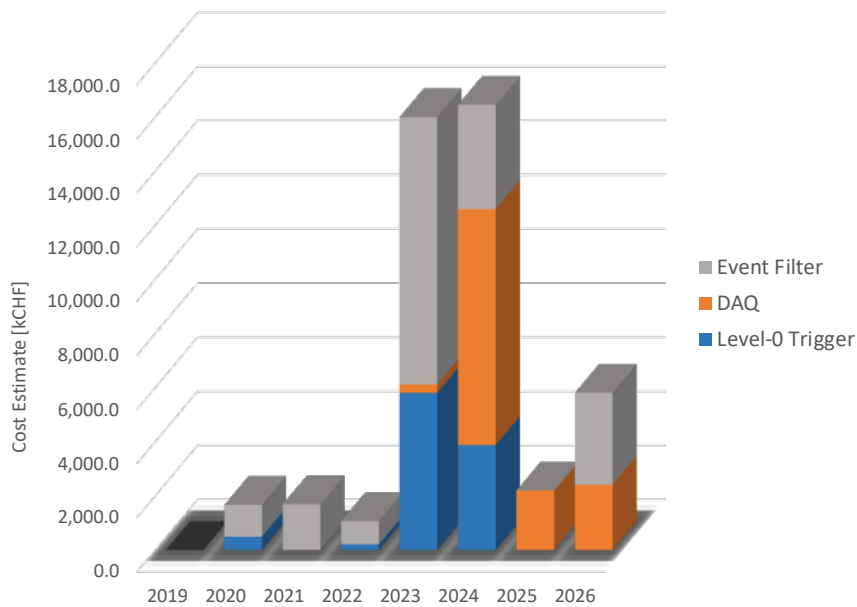


Figure 18.5: Estimated CORE spending profile of TDAQ upgrade project. The stacked histogram shows the contributions from the PBS Level-2 systems: the Level-0 trigger, DAQ and EF.

The expected spending time profiles, expanded at Level-3 of the PBS (sub-systems) for each of the systems, are shown in Figures 18.6, 18.7, and 18.8, respectively. The profiles reflect the schedule discussed in Chapter 19.

### 18.6.1 Spending profile of the Level-0 Trigger System

The Level-0 Trigger sub-systems components are essentially designed as ATCA modular electronics, and their production spans the years 2022-2024.

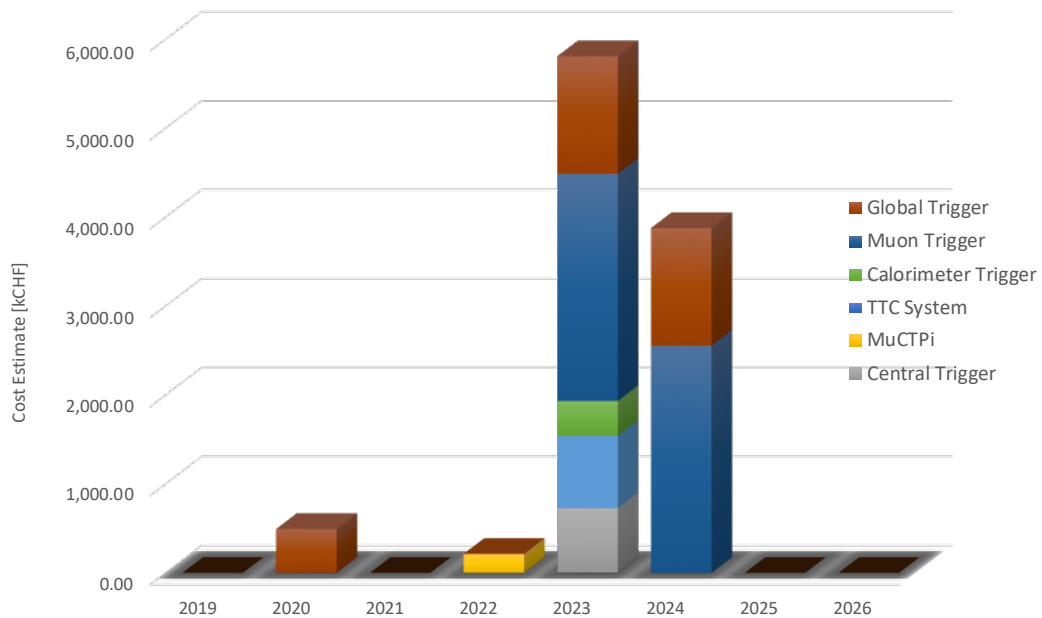


Figure 18.6: Estimated CORE spending profile for the Level-0 trigger system. The stacked histogram shows the contributions from the PBS Level-3 items.

### 18.6.2 Spending profile of the DAQ System

The cost of the DAQ upgrade is spread between 2024 and 2026. The plan includes the support for the surface tests of the ITk detectors during assembly through pre-production units. Some DAQ elements, in particular for what concerns networking, are COTS components purchased toward the end of LS3 and the beginning of the HL-LHC operations.

## 18.6 Costing Profiles

---

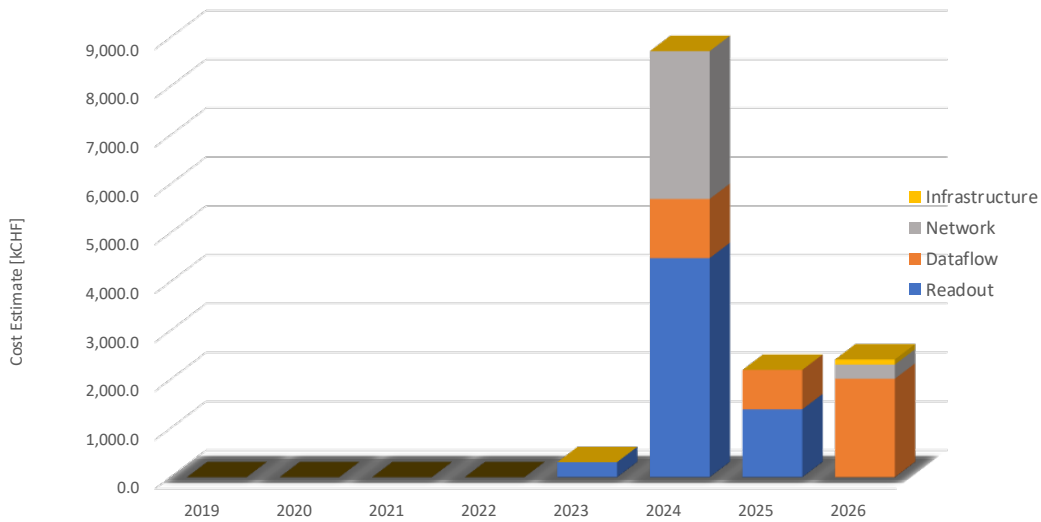


Figure 18.7: Estimated *CORE* spending profile for the *DAQ* system. The stacked histogram shows the contributions from the *PBS* Level-3 items.

### 18.6.3 Spending profile of the EF Trigger System

The *HTT* sub-system is based on *ATCA* boards and contains a challenging and technically complex *ASIC*. The production peaks in the years 2020-2024, with pre-production *ASICs* already produced as early as 2020. Procurement of the *EFPU* servers is carried out conveniently as late as possible, consistently also with the schedule for the M&O rolling replacement programme (see Section 18.5.3).

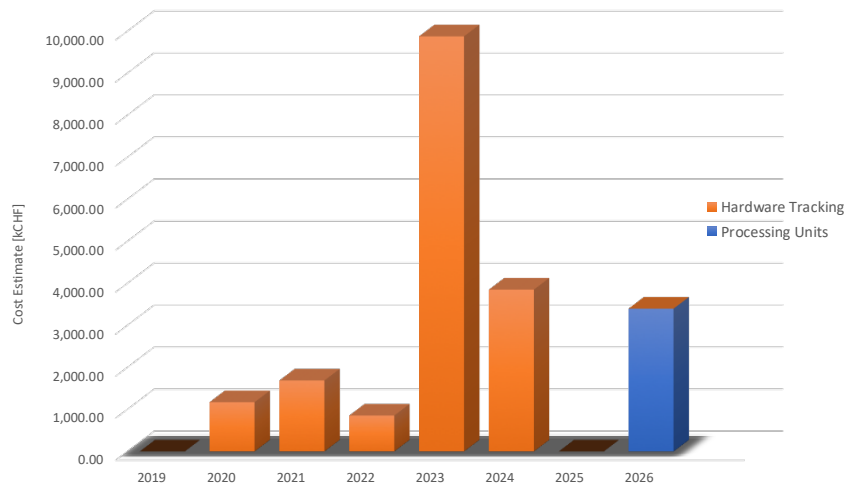


Figure 18.8: Estimated *CORE* spending profile for the *EF* system. The stacked histogram shows the contributions from the *PBS* Level-3 items.

## References

- [18.1] ATLAS Collaboration, *Technical Design Report for the ATLAS Inner Tracker Strip Detector*, Tech. Rep. CERN-LHCC-2017-005, CERN, Geneva, Apr, 2017. <https://cds.cern.ch/record/2257755>.
- [18.2] The TDAQ Group, *TDAQ Phase-II PBS dictionary*, Tech. Rep. ATL2-D-MC-0012, EDMS Id 1817664 v.2, CERN, Geneva, Oct, 2017. <https://edms.cern.ch/document/1817664/2>.
- [18.3] H. Meinhard, *Technology Evolution*, Tech. Rep. CHEP 2016, CERN, Geneva, Oct, 2016. <https://indico.cern.ch/event/555063/contributions/2285842/attachments/1350817/2039143/2016-10-08-WLCGWorkshop-TechnologyEvolution.pdf>.





# 19 Planning and Schedule

This chapter focuses on the planning and an initial high-level schedule of the UPR that summarises the information documented in the Schedule Management Plan (SMP) [19.1], which is part of the *Confidential Material* package to be reviewed by the UCG in early 2018, and available on the EDMS area of the project [19.2].

Detailed bottom-up plans of the activities of each sub-system [19.2] are used as the basis for the overall project plan, schedule and milestones. The plan described in this TDR and in the SMP builds on the experience of the TDAQ Phase-I upgrade project [19.3].

The schedule includes all the tasks required until the completion of the project with installation and commissioning finished, and all their interdependencies. The plans will evolve as the project moves forward, including the addition of more detailed breakdown of tasks as they approach. This will enhance the ability of the UPR management to track progress, detect delays as early as possible, and take pre-emptive actions.

The chapter is organised as follows: Section 19.1 summarises the principal concepts and guidelines extracted from the project's SMP; Section 19.2 describes the production plan and the reviewing process that controls and monitors the technical progress of the project for a successful and timely delivery to ATLAS. Finally, Section 19.3 summarises the initial plans of the project's systems and sub-systems, describing their top-level production schedule, and presenting in full detail one reference example, the L0Muon Trigger Barrel and Endcap SL.

## 19.1 Overview of the Schedule Management Plan

The complete plan to monitor, control, and revise, if necessary, the schedule of the project is documented in a separated document [19.1]. The basic elements of the Plan are:

- The systems' and sub-systems' coordinators are responsible for identifying all the activities and the tasks required for the design, construction, installation and commissioning of a given deliverable. The tasks and their duration are documented in a structured hierarchical work breakdown document (WBS) which goes down to Level-5 or Level-6.

## 19.1 Overview of the Schedule Management Plan

---

- The coordinators shall agree with the project's **UPR-RC** on the consistency and completeness of the **WBS**'s tasks required to deliver successfully to ATLAS the system/sub-systems' elements.
- The coordinators prepare the schedule for each Level-2 and Level-3 deliverable. The schedules are approved by the **UPR-RC** through an internal scrutiny process, that is chaired by the **UPR-RC**. The **UPR-RC** reports the conclusions of the internal scrutiny to the **eTDSG**, requesting the report's endorsement by the **UPL** after a discussion and the unanimous agreement of the **eTDSG** members.
- The **UPR-RC** assists the **UPL** to develop a comprehensive schedule, in accordance with the general schedule developed by the ATLAS **USC**, to seek the necessary review process to baseline the schedule, to oversee the progress, and to take necessary corrective actions to ensure that the project remains on schedule.
- The schedule of each Level-2 and Level-3 deliverable shall include all the formal review steps outlined in Section 19.2, which are necessary tools for the **UPR** management to monitor and track that the design and/or production make adequate progress towards their completion.
- After each Level-2 and Level-3 deliverable's approval, it is the responsibility of the system and sub-system's coordinators to implement, execute, and track the progress of their project against the baselined schedule for their respective deliverables. They report regularly on the progress to the **eTDSG**.
- In the event of significant accumulated schedule variance, or in the event of technical issues emerging during the production phase, the **UPR-RC** organises a Production Assessment Review (**PAR**) at a production site to identify the root-causes of the technical issues and to propose corrective actions. The **UPR-RC** reports to the **UPL** who approves formally the proposed corrected actions coming from the **PAR**, following a discussion and agreement by the **eTDSG**.
- The **UPR-RC** may request the **UPL**'s approval for a revision-controlled change of the baselined schedule in the event of a significant variance. The **UPL** approval is conditional to the majority agreement by the **eTDSG** members.
- The **UPL** reports regularly to the **TDIB** on the state of progress compared to the baselined schedule.
- The **UPL** informs immediately the **TDIB** and the ATLAS **USC**, in the event of becoming aware of a critical schedule variance.
- Upon request by the ATLAS Management, and, specifically by the **UC**, the **UPL** reports to the ATLAS **USC** on progress with respect to the baselined schedule. ATLAS Management through the **UPO** may also require periodic formal reviews of the status of the **UPR**, along the same lines as the **ASSO** reviews during the original construction phase.

---

## 19.2 Design, Production Plan and Hardware/Firmware Integration

The development and production process of each sub-system, which is broken down into tasks in the [WBS](#), is mapped onto a few generic phases during the project lifetime. At the end of each phase a formal review establishes whether the phase has concluded successfully: for example, if it is compliant to the specifications, and satisfies all the quality and reliability requirements. If so, it is approved to progress to the next stage of the development/production cycle. The review may be an internal process within the [TDAQ](#) organisation, or an agreement with ATLAS Management to hold a formal ATLAS event organised either by the ATLAS Electronics Coordinator, in his/her function of chair of the ATLAS Electronics Review Office, or by a designated software responsible for software matters.



Figure 19.1: *Sequence of activities and approval reviewing processes at different stages of the development and of the production cycle of any project's deliverable.*

Each coordinator shall define and include in the system/sub-system planning all the milestones and phases (from initial developments to the final installation), which are represented graphically in Figure 19.1 and described in the paragraphs that follow.

The statement is valid not only for the hardware deliverables, but for firmware and software components as well, as they shall be considered deliverables in the same way as their hardware counterparts.

The complexity and resources of single devices have grown by more than two orders of magnitude since the the original ATLAS construction, even with respect to the large digital ASIC designs of some sub-systems. Recognising the significance of this, firmware activities shall be integrated into the UPR planning in a more structured and detailed fashion than has been done in the past. In particular, the firmware development and the hardware design shall be planned and scheduled coherently from the Requirements capture up to the Commissioning of each sub-system. The need for this approach is particularly true for the TDAQ UPR where most of the sub-systems contain technically challenging PCBs that host multiple high-performance FPGAs, in which complex algorithms are implemented which must fit the resource limits of the selected device.

There are four types of design reviews for each sub-system component:

- the Specification Review (SR),
- the Preliminary Design Review (PDR),
- the Final Design Review (FDR),
- the Production Readiness Review (PRR).

These reviews are done in agreement with ATLAS Management and are organised by the ATLAS Electronics Coordinator, in his/her function of chair of the ATLAS Electronics Review Office. Each of these reviews is described in a dedicated paragraph below.

For TDAQ components that contain a large effort on software development software reviews will be carried out at the time of SR, PDR, FDR, PRR. The scope of these reviews will be defined in a dedicated document.

A number of Internal TDAQ Reviews are also foreseen at various stages, for example at the end of the requirement capture, or just before installation of a subsystem. The scope of these reviews varies and they are briefly described in a single dedicated paragraph below.

### 19.2.1 Specification Review

The first step of the component development is the definition of a set of specifications and its proposed architecture. A Specification Document is prepared, based on a set of requirements, described in a “User Requirement Document” (URD) previously reviewed during an Internal TDAQ Review. The SR reviews the Specification Document that describes the required functionality and performance of the device, its interfaces to other devices, and reliability. The specified interfaces must be cross-checked for consistency with the corresponding components specifications. The approval of the SR starts the formal design phase of a prototype that captures all or a large fraction of the functions specified in the Specification Document. At the end of the design cycle a PDR is organised.

## 19.2.2 Preliminary Design Review

The **PDR** establishes whether the design meets all the aspects of the requirements and specifications, including all interfaces to other components. The prototype may not implement all functionality required in final system. If so, the prototype will need its own specifications, distinct from that reviewed previously in the **SR**. To avoid ambiguity, in this case, the **PDR** will address the prototype spec and any differences between this and the **SR** specification.

The review process is based upon the following documentation:

- The **URD** and the Specification Document to be used as benchmarks for the prototype design evaluation.
- Block diagrams - both at the board level and at the level of single **FPGAs**.
- Full Schematics-set of prototype **PCB** designs.
- **PCB** layout and **PCB** stack-up information when deemed necessary by the complexity of the component.
- High-level and/or SPICE-level simulations of the main functions.
- Preliminary estimate of the power consumption of the board appropriately dressed and configured.
- Preliminary thermo-mechanical analysis through a Finite Element Analysis program.
- Design and Verification of all the input/output interfaces.

**PDRs** should be held prior to the submission of any prototype design for fabrication or assembly.

The **PDR** shall also include in its scope the initial review of the **FPGA** firmware, held as a separate review if judged so by the Electronics Coordinator. The details of the scope of the firmware review at the time of the **PDR** will be described in a separate document.

## 19.2.3 Final Design Review

The **FDR** is the last step of the design phase of a component deliverable. The purpose of the review is to assess whether the design is mature enough for pre-production. All requirements and specifications are reviewed. A successful performance measurement of the prototype(s) that have been built allows the project to finalise the design in preparation for small scale pre-production, which will be approved upon a successful **FDR** review.

The review's scope and the documentation to be prepared for the hardware deliverables are similar to the **PDR**, emphasising the design modifications made since the design of the last prototype iteration (see Section 19.2.2).

The **FDR** shall also include in its scope a review of the **FPGA** firmware, held as a separate review if judged so by the Electronics Coordinator. The details of the scope of the firmware

review at the time of the [FDR](#) will be described in a separate document, but as a general guideline, the following aspects related to the firmware implementation will be reviewed:

- FPGA resource utilisation.
- Firmware framework description
- External interface tests (such as interface link-speed tests, performance measurements of individual interfaces for serialiser/deserialise latency at a given baud rate and a defined data protocol, etc)
- Single board internal slice test (for the realtime path if applicable, otherwise for the readout path, with at least one-channel input, a single processing algorithm, and at least one-channel output).
- Power consumption based on real implementation of the most power-hungry functionalities

### 19.2.4 Production Readiness Review

The [PRR](#) shall address quality management (QA/QC), version control (hardware and firmware) and any outstanding items from earlier reviews.

During the Production Readiness Review, the following items shall be considered:

- full functionality test with single board, including readout and monitoring path;
- system slice-test, where this project is part of a bigger test stand (many-channel in, many-channel out if applies, or a single processing algorithm, out of many, running on all input channels);
- test hardware and software needed for production testing;
- production-related issues (plan, monitoring of important parameters like yield);
- production test methodology (dedicated firmware, dedicated test bench, paths used as for example: [JTAG](#) chain coverage and the necessary functional tests where [JTAG](#) does not cover parts of the design for example);
- QA/QC plan.

At the level of single-board test, full system test or system integration test, a description of the [DAQ](#) software necessary for configuration, test and monitoring, together with a description of its development and major milestones should be presented. Power estimates need to be re-analysed during the review. Test coverage of all test benches should be evaluated, if they have changed in terms of input data and output checks, or if some new ones have been added.

The [PRR](#) shall also include in its scope a review of the [FPGA](#) firmware, held as a separate review if judged so by the Electronics Coordinator. The details of the scope of the firmware review at the time of the [PRR](#) will be described in a separate document,

The [PRR](#) is the last ATLAS-wide review chaired by the ATLAS Electronics Coordinator. Additional Internal TDAQ reviews, described below, are held to monitor the component development up to Installation and Commissioning.

### 19.2.5 Internal TDAQ Reviews

The reviews described below are an internal process of [TDAQ](#) and will be organised within the context of the [eTDSG](#). Internal TDAQ reviews are not always mandatory for deliverables, but are carried as needed.

**Requirement Capture and Internal Requirement Review.** The definition of a set of requirements is the first step of a design. A "User Requirement Document" ([URD](#)) that contains a list of short, well-defined statements shall be prepared. The [URD](#) contains:

- A brief introduction with a description of the [UPR](#) system and sub-systems to which the hardware deliverable belongs, or where firmware will be deployed;
- A list of the required functionalities;
- Physics Performance requirements and more generally Performance requirements for all functionalities;
- A list of all required interfaces to integrate the project into the [UPR](#) sub-system;
- A list of requirements that the project puts on the rest of the [UPR](#) systems and, more in general, on [TDAQ](#).

A high-level simulation model of the component may be available at this stage to define these requirements. The [URD](#) is the document used for the definition of the specifications and will be part of the documentation required for the [SR](#). This internal review is mandatory for all deliverables.

**Internal Follow-up Reports and Status/Progress Design Review.** During the construction of the prototypes and/or the evaluation of the Key Performance Indicators, the sub-system coordinator will report regularly to the [eTDSG](#) on the status of the prototype design, construction and, eventually, the evaluation of its performance. Once the performance evaluation is completed, a Progress Design Review may be organised to verify the compliance of the prototype design to the specifications. This review is only organised upon request of the [UPR](#) management and it might be an essential tool for [UPL](#) and his/her management team to verify whether the sub-system is on schedule, or to minimise any schedule and cost impacts deriving from potential design issues at an early stage. In the event that this review fails, a second prototyping cycle may be required through a Follow-up [PDR](#).



**Production Advancement Review.** During the series production, it may be useful to schedule advancement reviews to monitor the progress of production through to its completion. Such a review may help to ensure that any non-conformities or schedule problems arising during production are addressed promptly. This internal review is not mandatory for all deliverables, but is carried as needed.

**Production completion and Installation Readiness Review.** Once the production of a hardware component of the UPR has completed, and all the hardware deliverables have been transported to CERN, its integration with the final, i.e. “production”, firmware shall be done at the TMF facility and reviewed before the final deployment of the configured hardware in the USA15 counting room.

The following items shall be included in the scope of the “Installation Readiness Review”:

- Full system test in the TDAQ TMF surface facility.
- Full realtime path test with all baseline algorithms implemented.
- System integration tests with detectors electronics.
- Monitoring tasks and other TDAQ software (databases, online).

**Firmware Review at the time of Installation.** A close-out review of a component firmware will be held at the end of the Production phase, to assess the suitability of the firmware for long-term support, covering documentation, version management and regression testing.

Other firmware reviews will be held after installation, during commissioning and during operation, to deploy new trigger algorithms for example, but they are outside the scope of this TDR and are not further described.

## 19.3 Production Schedule and Milestones

Figure 19.2 shows the top-level schedule of the UPR down to the Level-3 of the PBS/WBS.

In order to quantify the criticality of items for the schedule, the planned dates for completion of production for the primary deliverables have been compared to the required dates for their installation. The schedule float is defined as the number of working days available between the end of production and the required date of installation for each deliverable, after accounting for the time required for surface commissioning and testing. Sufficient float is available to accommodate unforeseen delays in all aspects of the project.

A full set of Gantt charts, for each Level-3 item in the WBS, is available in the Planning section of the UPR’s EDMS structure [19.2], and will be included in the “Confidential material”



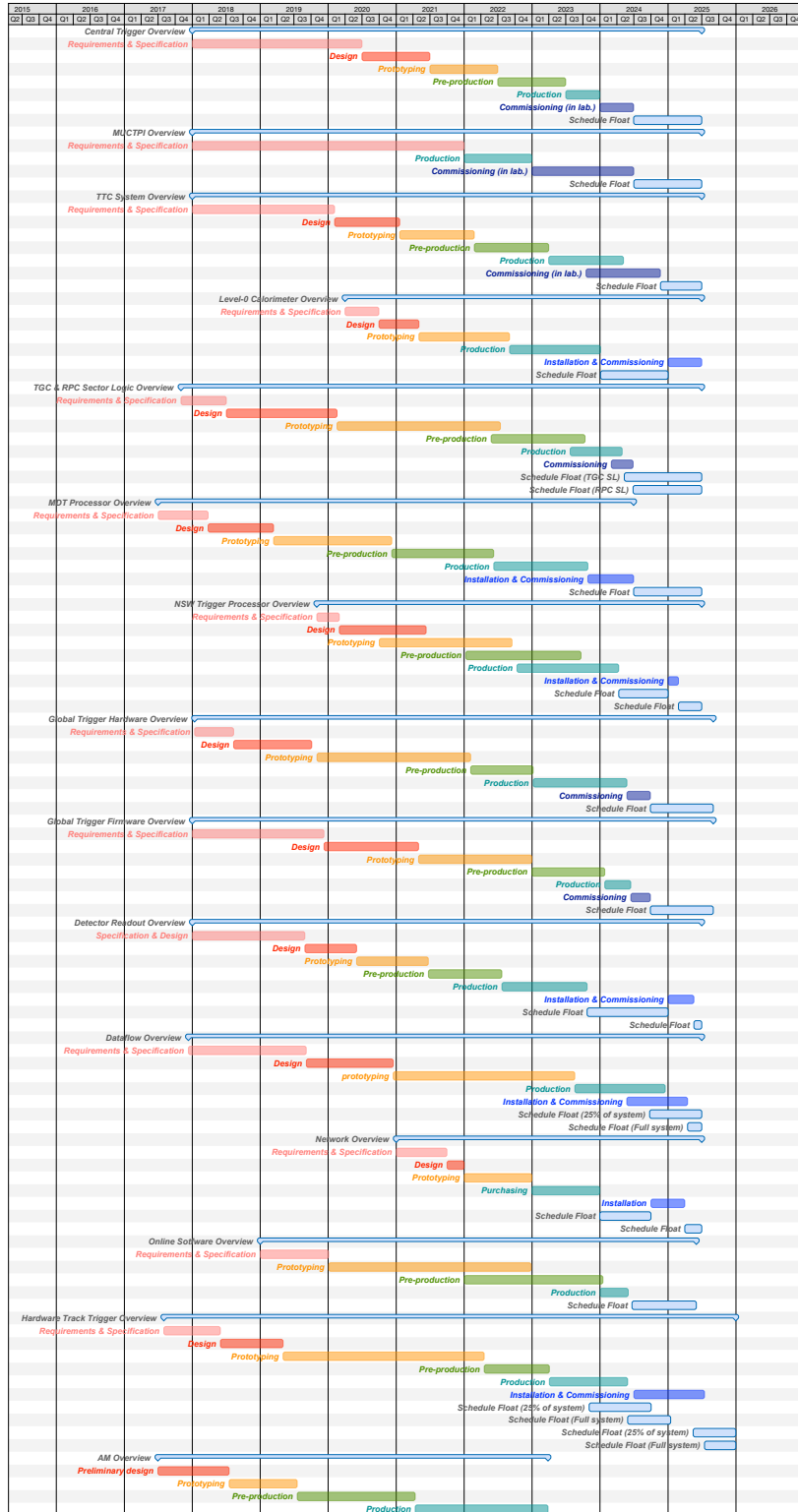


Figure 19.2: High-level schedule for the UPR deliverables at the Level-3 of the PBS. The schedule includes design, prototyping, pre-production, production, installation and commissioning.

### 19.3 Production Schedule and Milestones

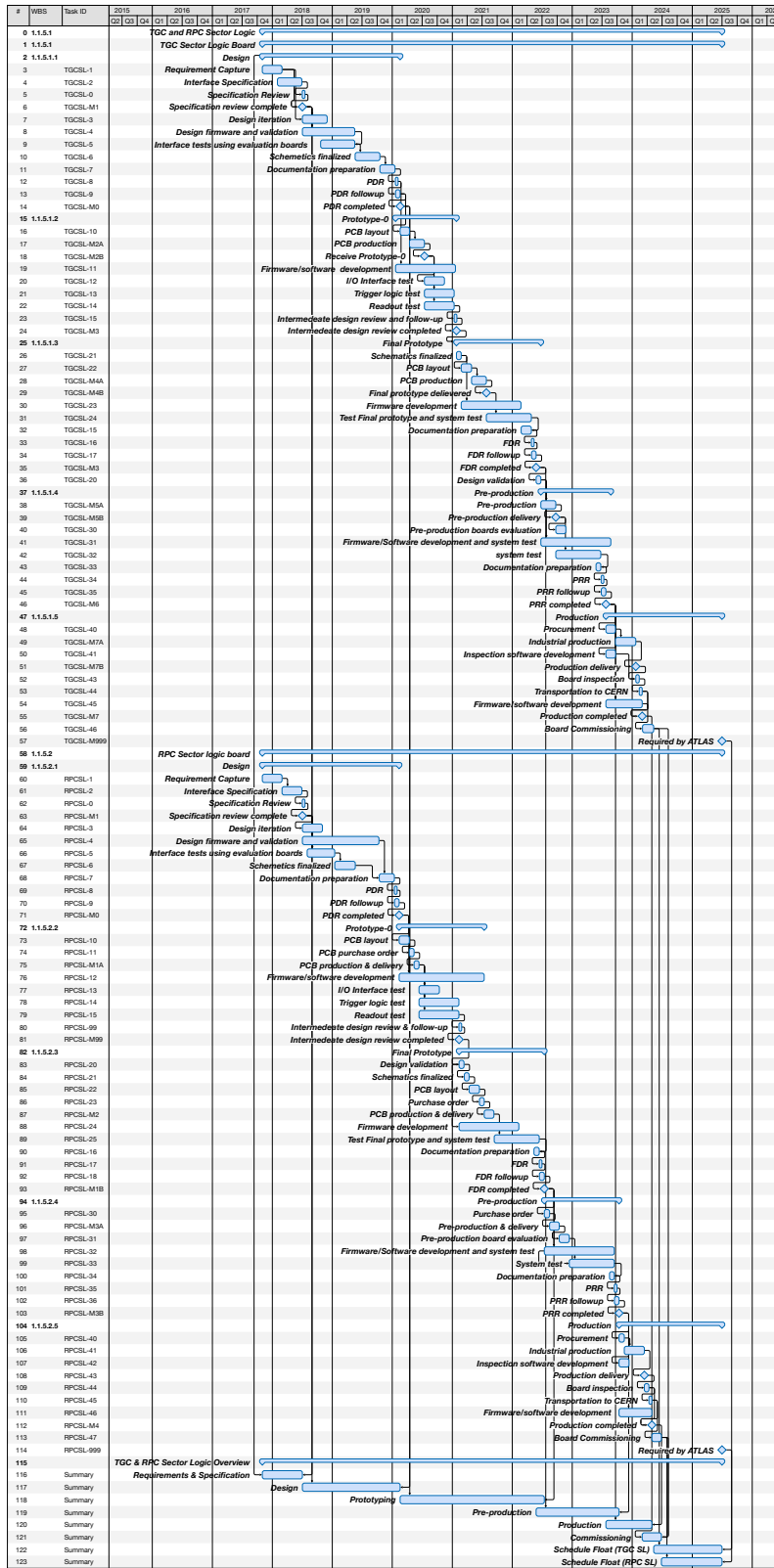


Figure 19.3: Example of expanded Gantt chart for the Level-0 Muon Trigger Barrel and Endcap SL.

for the UCG review in Q1-2018. For illustrative purposes only the fully expanded planning of the Level-0 Muon Trigger (TGC and RPC) SL is shown in Figure 19.3.

### 19.3.1 Level-0 Trigger System's Milestones

Tables 19.1- 19.4 summarise the principal milestones for the Level-0 Trigger sub-systems during the construction.

Table 19.1: *Principal milestones for the Level-0 Central Trigger sub-system that includes the following Level-3 deliverables: CTP, MUCTPI, and TTC.*

WBS	Descr.	Milestone	Date
1.1.1	CTP	Specification Review	30.06.2020
		Preliminary Design Review	30.06.2021
		Final Design Review	30.06.2022
		Firmware Design Review	30.06.2023
		Production Readiness Review	30.06.2023
1.1.2	MUCTPI	Specification Review	31.12.2021
		Firmware Design Review	30.06.2023
1.1.3	TTC	Specification Review	30.03.2020
		Preliminary Design Review	31.12.2020
		Final Design Review	03.01.2022
		Firmware Design Review	30.06.2022
		Production Readiness Review	03.01.2023

Table 19.2: *Principal milestones for the Level-0 Calorimeter Trigger sub-system.*

WBS	Descr.	Milestone	Date
1.1.4	Level-0 Calorimeter	Specification Review	30.09.2020
		Preliminary Design Review	03.05.2021
		Final Design Review	01.09.2022
		Production Readiness Review	01.02.2023
		Firmware Design Review (Phase-I FEXs)	01.01.2024

### 19.3 Production Schedule and Milestones

Table 19.3: *Principal milestones for the Level-0 Muon Trigger sub-system (WBS: 1.1.5).*

WBS	Descr.	Milestone	Date
1.1.5.1	<b>TGC Sector Logic Board</b>	Specification Review	02.07.2018
		Preliminary Design Review	17.02.2020
		Final Design Review	25.05.2022
		Production Readiness Review	24.07.2023
1.1.5.2	<b>RPC Sector Logic Board</b>	Specification Review	02.07.2018
		Preliminary Design Review	11.02.2020
		Final Design Review	14.07.2022
		Production Readiness Review	12.10.2023
1.1.5.3	<b>MDT Processor Main Board</b>	Specification Review	02.07.2018
		Preliminary Design Review	28.02.2020
		Final Design Review	20.07.2021
		Production Readiness Review	22.12.2022
1.1.5.4	<b>MDT Rear Transition Module</b>	Specification Review	02.07.2018
		Preliminary Design Review	28.02.2020
		Final Design Review	22.07.2021
		Production Readiness Review	19.12.2022
1.1.5.5	<b>MDT Sector Processor Mezzanine</b>	Specification Review	02.07.2018
		Preliminary Design Review	28.02.2020
		Final Design Review	20.07.2021
		Production Readiness Review	16.12.2022
1.1.5.6	<b>NSW Carrier</b>	Specification Review	01.03.2020
		Preliminary Design Review	02.10.2020
		Final Design Review	10.01.2022
		Production Readiness Review	11.10.2022
1.1.5.7	<b>NSW Mezzanine Card</b>	Specification Review	01.03.2020
		Preliminary Design Review	11.06.2021
		Final Design Review	15.09.2022
		Production Readiness Review	21.09.2023

Table 19.4: *Principal milestones for the Level-0 Global Trigger sub-system (WBS: 1.1.6).*

WBS	Descr.	Milestone	Date
1.1.6.1	<b>GCM</b>	Specification Review	05.10.2019
		Preliminary Design Review	05.10.2019
		Final Design Review	05.02.2022
		Production Readiness Review	07.01.2023
1.1.6.2	<b>PFM</b>	Specification Review	25.04.2018
		Preliminary Design Review	13.04.2019
		Final Design Review	04.01.2020
		Production Readiness Review	06.06.2020
1.1.6.3	<b>Data Aggregator (Firmware)</b>	Specification Review	10.12.2019
		Preliminary Design Review	23.01.2021
		Final Design Review	10.12.2022
		Production Readiness Review	08.12.2023
1.1.6.4	<b>Trigger Framework (Firmware)</b>	Specification Review	10.12.2019
		Preliminary Design Review	09.01.2021
		Final Design Review	05.11.2022
		Production Readiness Review	14.07.2023
1.1.6.5	<b>Trigger Signature (Firmware)</b>	Specification Review	10.12.2019
		Preliminary Design Review	01.05.2021
		Final Design Review	31.12.2022
		Production Readiness Review	26.01.2024

### 19.3.2 DAQ System's Milestones

Table 19.5 summarises the principal milestones for the DAQ sub-systems during the construction.

### 19.3.3 EF System's Milestones

Tables 19.6 and 19.7 summarise the principal milestones for the two sub-systems in the EF system during the construction.

### 19.3 Production Schedule and Milestones

Table 19.5: *Principal milestones for the sub-systems of DAQ.*

WBS	Descr.	Milestone	Date
1.2.1	<b>Detector Readout</b>	Specification Review	28.08.2019
		Preliminary Design Review	01.06.2020
		Final Design Review	22.06.2021
		Production Readiness Review	22.07.2022
1.2.2	<b>Dataflow</b>	Specification Review	05.09.2019
		Preliminary Design Review	15.12.2020
		Final Design Review	18.08.2023
		Production Readiness Review	
1.2.3	<b>Network</b>	Specification Review	30.09.2021
		Preliminary Design Review	31.12.2021
		Final Design Review	31.12.2022
1.2.5	<b>Online software</b>		
1.2.5.1	<b>EF Farm Orchestration</b>	Selection of the cluster orchestrator system	03.01.2020
		Prototype System Review	10.12.2021
		Pre-production System Review	12.09.2022
		Production system fully operational	31.05.2024
1.2.5.2	<b>Physics Monitoring</b>	Requirement Document Review	13.10.2021
		Prototype Test Completion and Review	27.12.2022
		Final Design Review	06.04.2023
		Production Deployment Review	17.01.2024

Table 19.6: *Principal milestones for the HTT sub-system (WBS: 1.3.1).*

WBS	Descr.	Milestone	Date	
1.3.1	<b>HTT</b>	System-wide Specification Review	01.06.2018	
		Preliminary Design Review	30.05.2019	
1.3.1.1	<b>TP</b>	Final Design Review	20.04.2022	
		Production Readiness Review	29.03.2023	
1.3.1.2	<b>TP RTM</b>	Final Design Review	02.11.2021	
		Production Readiness Review	29.03.2023	
1.3.1.3	<b>PRM</b>	Final Design Review	20.11.2021	
		Production Readiness Review	20.03.2023	
1.3.1.4	<b>AM ASIC</b>	AM08 (Prototype)		
			Preliminary Design Review	20.07.2018
		AM09pre (Preproduction)		
			Final Design Review	17.07.2019
	AM09 (Production)			
		Production Readiness Review	14.08.2020	
1.3.1.5	<b>TFM</b>	Final Design Review	28.01.2022	
		Production Readiness Review	20.03.2023	

Table 19.7: *Principal milestones for the EFPU sub-system (WBS 1.3.2).*

WBS	Descr.	Milestone	Date
1.3.2.1	Servers	Technology decision	01.04.2026
1.3.2.2	EF Compute Requirements	Updated EF farm compute requirements	10.04.2023
		Final EF farm compute requirement for purchase	28.05.2025
1.3.2.3	Accelerator evaluation	PDR: Design of EF farm using accelerators	10.05.2024
		FDR: Conclusion on use of accelerators for rolling replacement procurement	13.01.2025
1.3.3	Software	EF reconstruction software initial prototyping complete	01.10.2019
		Software Design Reviews (all complete)	02.08.2024
		Software ready for Run 4 commissioning	26.02.2025
		Software ready for MC26a	26.02.2025

## References

- [19.1] The TDAQ Group, *TDAQ Phase-II Upgrade Schedule Management Plan*, Tech. Rep. ATL2-D-MG-0003, EDMS Id 1874385 v.1, CERN, Geneva, Nov, 2017. <https://edms.cern.ch/document/1874385/1>.
- [19.2] The TDAQ Group, *TDAQ Phase-II Upgrade Project Management Plan Documents*, Tech. Rep. CERN-0000184878, CERN, Geneva, Nov, 2017. <https://edms.cern.ch/project/CERN-0000184878>.
- [19.3] ATLAS Collaboration, *Technical Design Report for the Phase-I Upgrade of the ATLAS TDAQ System*, Tech. Rep. CERN-LHCC-2013-018. ATLAS-TDR-023, CERN, Geneva, Sep, 2013. <https://cds.cern.ch/record/1602235>.





## 20 Resources Requirements and Institutional Responsibilities

This chapter documents the tasks and the associated required human resources for each system and sub-system, estimated bottom-up based on input from the institutes. In addition, the [UPL](#) and the [TDIB](#) chair are surveying the aspirations of the Institutes, and investigating with their Representatives whether the required resources are available locally among those Institutes that have expressed interest in the different systems and sub-systems. Many Institutes have already expressed their intention of contributing to the [TDAQ UPR](#). This is an iterative process that will ultimately complete with an agreement between ATLAS and the Institutes, and by signing a “Memorandum of Understanding” ([MoU](#)) document for the [TDAQ Phase-II Upgrade Project \(UPR\)](#) construction after the approval of the [UPR](#) by the CERN Resource Review Board ([RRB](#)). The [MoU](#) will be signed by ATLAS and the Funding Agencies of the participating Institutes. Any formal commitment is taken only after the [MoU](#)'s signing has been completed.

The chapter is organised in three main sections. An overview of the Resource Management Plan is presented in Section [20.1](#). A set of tables and charts, summarising the required effort at Level-3 of the [PBS](#) and [WBS](#), is shown in Section [20.2](#). Note that for a given deliverable, the total required effort is estimated corresponding to the needed expertise and skill sets. However, in the tables the effort is divided into the typical professional categories consistent with the composition of the human resources available in those institutes that expressed their interests. More detailed information is being prepared for the [UCC](#) confidential material in the form of Excel spreadsheets. The last section of the chapter (Section [20.3](#)) documents in a few tables the interests and the aspirations of the participating Institutes (within and outside the current ATLAS [TDAQ](#) collaboration) to participate in the different elements of the [UPR](#) program.

### 20.1 Resource Management Plan

As part of the ATLAS organisation, the [TDAQ UPR](#) management relies on the Institutes and the Funding Agencies participating in the project to provide and manage directly the financial and human resources required to produce locally, and deliver successfully to ATLAS, the items for which institutes have agreed to be responsible.

## 20.2 Required Manpower Estimate

---

The Resource Management Plan is a process to develop and organise, through a set of documents, the resources needed based on the following:

- The UPR management through the UPR-RC and the Level-2/Level-3 managers of each system, determines the financial resources, the staffing and the staffing's competency required to complete the tasks in the project's WBS. An initial plan is prepared and documented for the approval of the process through the LHCC and UCG reviews.
- After the approval of the project, the UPL, assisted by the UPR-RC and the TDIB chair negotiate with the Institutes' Representative and the National Upgrade Programme Managers (or National Contact Physicists) responsibilities for specific tasks.
- The UPR-RC assists the UPL to verify that the resources and the staffing at a production site are sufficient to guarantee low risk of failures before the responsibility is assigned
- The final assignment of responsibilities to an Institute is a task of the UPL based on the CORE contribution, and shall be approved by the TDIB as part of the MoU document's preparation.
- The MoU documents and regulates not only the responsibilities of an Institution for each hardware deliverable, but also for firmware/software required to operate it.
- During the execution of the project, through specific reviews, the UPR-RC monitors the status and the progress of the design and of the construction, and verifies periodically the adequacy of the local resources.
- In case of failures or delays, the UPR management may decide, in concurrence with the TDIB chair, to intervene and negotiate with the Institutional Representative response strategies to recover schedule delays.

## 20.2 Required Manpower Estimate

### 20.2.1 Level-0 Trigger System

Table 20.1 summarises the required effort, in FTEs, for the Level-0 Trigger system in the years 2018-2026. The FTEs are subdivided in the following professional categories: scientists, electronics engineers, software engineers, technicians, and students. Figure 20.1 represents graphically, as stacked histograms, the same information as Table 20.1. Figures 20.4c-20.4f show the required effort for the six Level-3 sub-systems of the Level-0 Trigger system.

### 20.2.2 DAQ System

Table 20.2 summarises the required effort, in FTEs, of the DAQ system in the years 2018-2026 subdivided into the same personnel's categories as in Table 20.1. The same information

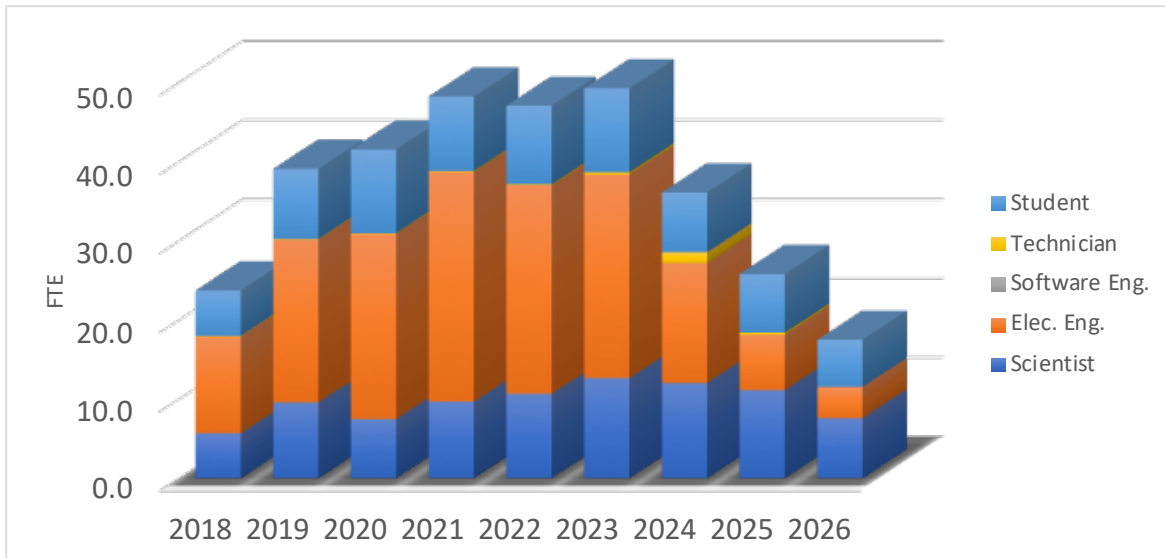


Figure 20.1: Estimated human resources required in the *LO* Trigger system, expressed as *FTEs* for each year of the upgrade project’s construction, installation and commissioning. Four professional categories are stacked on top of each other: Scientists (blue), Electronics Engineers (orange), Software Engineers (grey), Technicians (yellow), and Students (light blue).

Table 20.1: Required effort in the Level-0 Trigger system expressed in *FTEs* and divided by professional category for the duration of the *UPR*’s construction, installation and commissioning (2018-2026).

WBS: 1.1	Level-0 Trigger System						
	Year	Scientist [FTE]	Electronic Eng. [FTE]	Software Eng. [FTE]	Technician [FTE]	Student [FTE]	All [FTE]
	2018	5.7	12.2	0.0	0.1	5.7	<b>23.7</b>
	2019	9.6	20.6	0.0	0.1	8.9	<b>39.1</b>
	2020	7.5	23.4	0.0	0.1	10.6	<b>41.6</b>
	2021	9.7	29.0	0.0	0.1	9.4	<b>48.2</b>
	2022	10.7	26.4	0.0	0.1	9.9	<b>47.1</b>
	2023	12.7	25.7	0.0	0.3	10.6	<b>49.3</b>
	2024	12.1	15.1	0.0	1.4	7.6	<b>36.1</b>
	2025	11.1	7.0	0.0	0.2	7.4	<b>25.8</b>
	2026	7.6	4.0	0.0	0.0	6.0	<b>17.5</b>
<b>TOTAL</b>		<b>86.6</b>	<b>163.4</b>	<b>0.0</b>	<b>2.4</b>	<b>76.1</b>	<b>328.4</b>

## 20.2 Required Manpower Estimate

is shown in Figure 20.2. Figures 20.5a-20.5d show the required effort in each year of the

Table 20.2: Required effort in the DAQ system expressed in FTEs and divided by professional category for the duration of the UPR's construction, installation and commissioning (2018-2026).

Year	DAQ System						All [FTE]
	Scientist [FTE]	Electronic Eng. [FTE]	Software Eng. [FTE]	Technician [FTE]	Student [FTE]		
2018	1.4	1.2	0.7	0.0	1.1		4.3
2019	1.9	2.6	1.9	0.0	2.5		9.0
2020	2.0	1.9	1.9	0.0	2.9		8.8
2021	3.3	2.4	3.8	0.0	4.4		13.9
2022	6.6	2.8	3.6	0.0	5.4		18.4
2023	3.9	2.5	3.6	0.1	3.9		14.0
2024	2.6	1.1	3.4	0.9	2.5		10.4
2025	2.5	1.8	3.9	1.3	2.3		11.8
2026	0.7	0.3	0.6	0.2	0.4		2.1
<b>TOTAL</b>	<b>24.9</b>	<b>16.6</b>	<b>23.4</b>	<b>2.4</b>	<b>25.4</b>		<b>92.8</b>

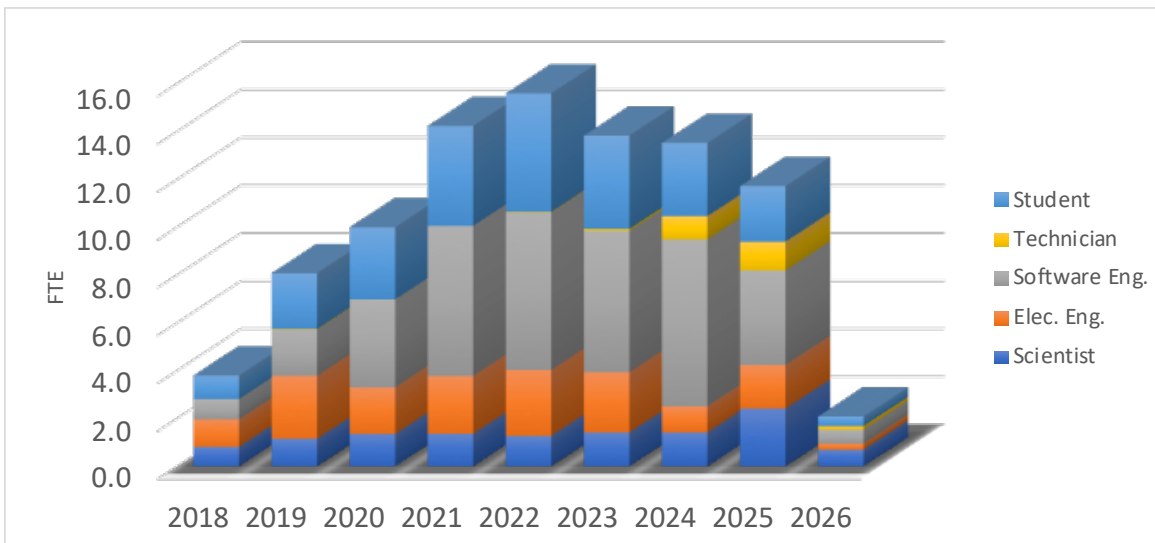


Figure 20.2: Estimated human resources required in the DAQ system, expressed as FTEs for each year of the upgrade project's construction, installation and commissioning. Four professional categories are stacked on top of each other: Scientists (blue), Electronics Engineers (orange), Software Engineers (grey), Technicians (yellow), and Students (light blue).

upgrade project's construction, installation and commissioning for the four (4) Level-3 sub-systems in the WBS.

### 20.2.3 EF System

Table 20.3 summarises the required resources for the UPR's EF system expressed in FTEs in the years 2018-2026. The FTEs are subdivided by the same professional categories as in Table 20.1. It should be noted that the EF software development requires scientists

Table 20.3: Required effort in the EF system expressed in FTE and divided by professional category for the duration of the UPR's construction, installation and commissioning (2018-2026).

WBS: 1.3 Year	Event Filter						All [FTE]
	Scientist [FTE]	Electronic Eng. [FTE]	Software Eng. [FTE]	Technician [FTE]	Student [FTE]		
2018	14.2	5.6	0.0	3.4	1.2		24.3
2019	17.1	6.7	0.0	4.2	1.5		29.6
2020	23.3	7.3	0.0	4.3	2.1		37.0
2021	28.7	6.9	0.0	3.3	1.7		40.6
2022	22.0	7.9	0.0	2.8	2.6		35.2
2023	26.6	5.8	0.0	2.8	3.5		38.6
2024	25.4	5.9	0.0	1.8	3.2		36.2
2025	10.6	5.4	0.0	1.8	3.6		21.4
2026	0.6	0.0	0.0	0.0	0.0		0.6
<b>TOTAL</b>	<b>168.4</b>	<b>51.4</b>	<b>0.0</b>	<b>24.4</b>	<b>19.4</b>		<b>263.5</b>

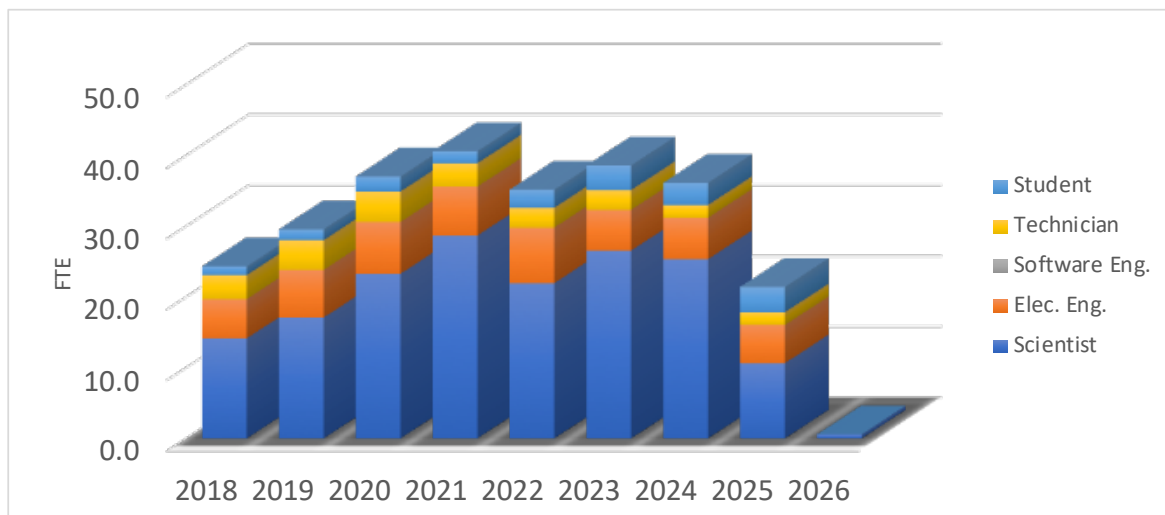
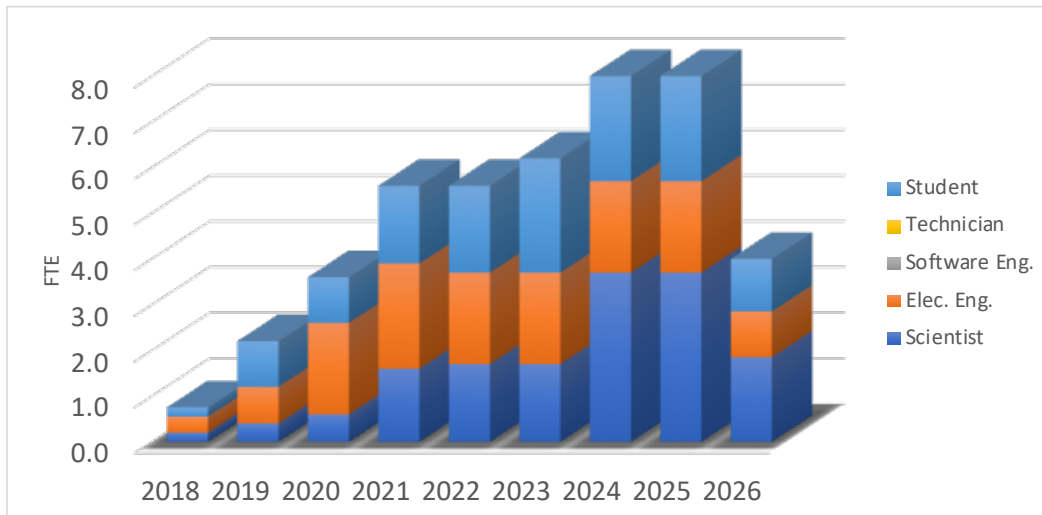


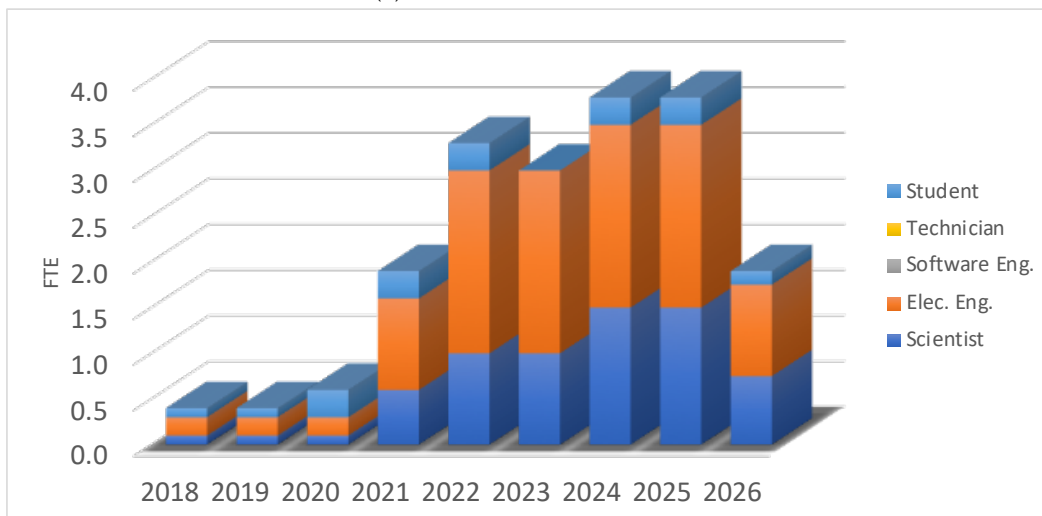
Figure 20.3: Estimated resources required in the EF system, expressed as FTEs for each year of the upgrade project's construction, installation and commissioning. Four professional categories are stacked on top of each other: Scientists (blue), Electronics Engineers (orange), Software Engineers (grey), Technicians (yellow), and Students (light blue).

## 20.2 Required Manpower Estimate

with strong software experience. In some cases the tasks may also be done by a software engineer. Figure 20.3 shows graphically the same information as in Table 20.3. Figures 20.6a, 20.6b show the required effort in each year of the upgrade project's construction, installation and commissioning for the two Level-3 sub-systems in the WBS.



(a) Level-0 CTP (WBS 1.1.1)



(b) Level-0 MUCTPI (WBS 1.1.2)

Figure 20.4: Required human resources expressed as FTEs for each year of the upgrade project's construction, installation and commissioning. Histograms are shown for the different sub-systems of the Level-0 Trigger system. Four professional categories are stacked on top of each other: Scientists (blue), Electronics Engineers (orange), Software Engineers (grey), Technicians (yellow), and Students (light blue).

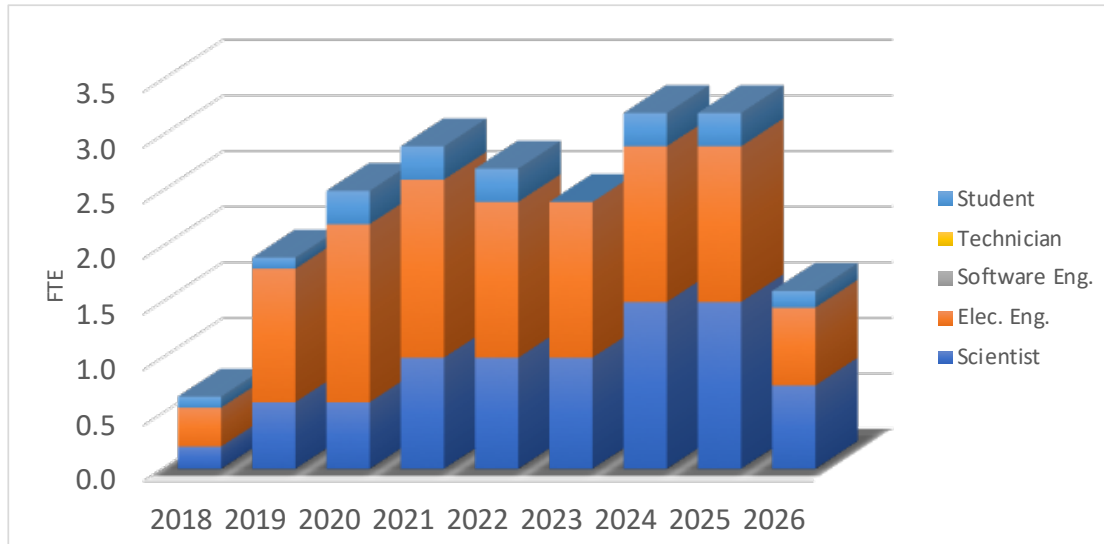
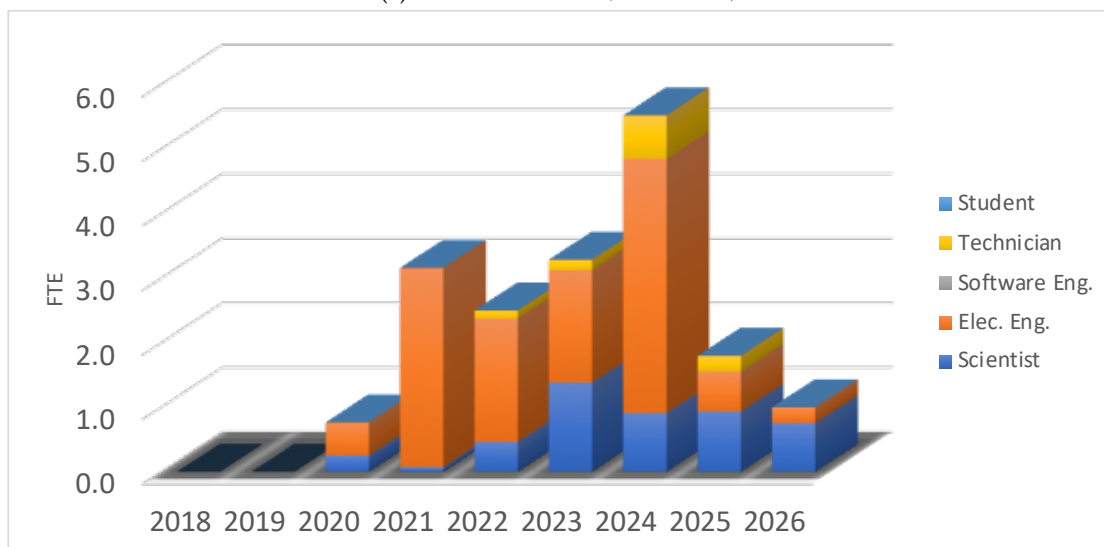
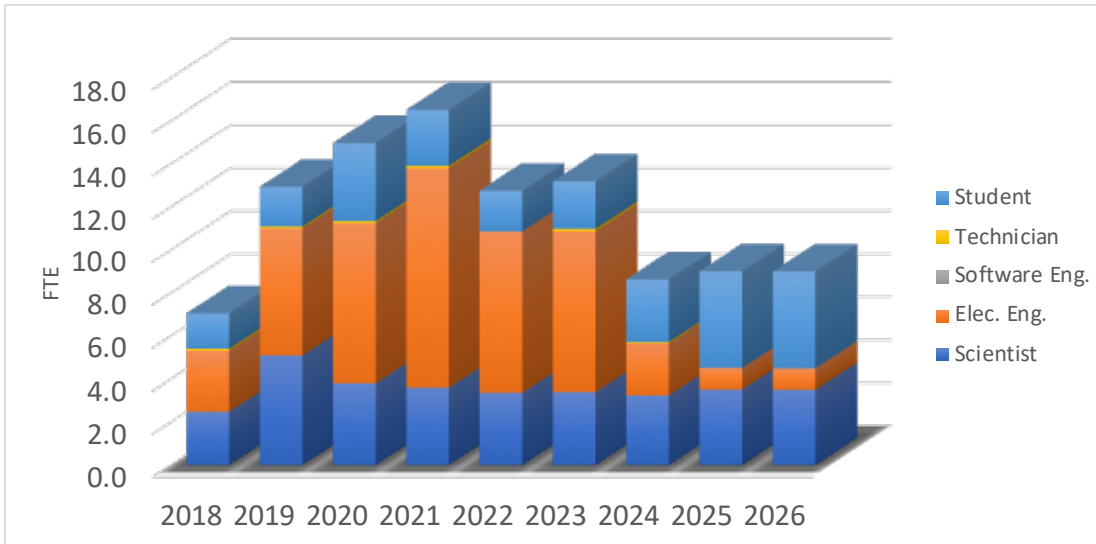
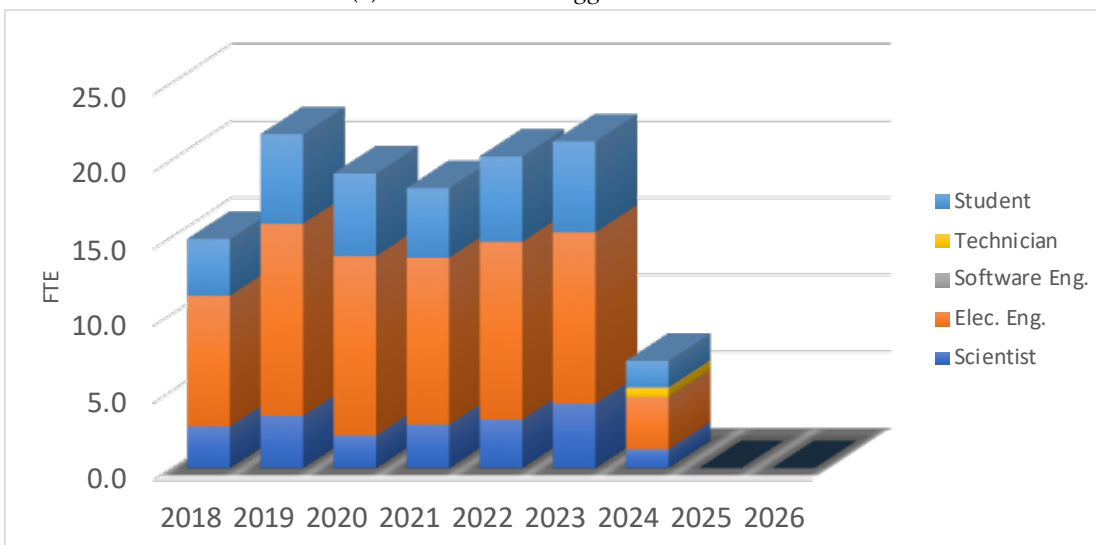
(c) *TTC distribution (WBS 1.1.3)*(d) *Level-0 Calorimeter Trigger (WBS 1.1.4)*

Figure 20.4: Required human resources expressed as *FTEs* for each year of the upgrade project's construction, installation and commissioning. Histograms are shown for the different sub-systems of the Level-0 Trigger system. Four professional categories are stacked on top of each other: Scientists (blue), Electronics Engineers (orange), Software Engineers (grey), Technicians (yellow), and Students (light blue).

## 20.2 Required Manpower Estimate



(e) Level-0 Muon Trigger (WBS 1.1.5)



(f) Level-0 Global Trigger (WBS 1.1.6)

Figure 20.4: Required human resources expressed as FTEs for each year of the upgrade project's construction, installation and commissioning. Histograms are shown for the different sub-systems of the Level-0 Trigger system. Four professional categories are stacked on top of each other: Scientists (blue), Electronics Engineers (orange), Software Engineers (grey), Technicians (yellow), and Students (light blue).



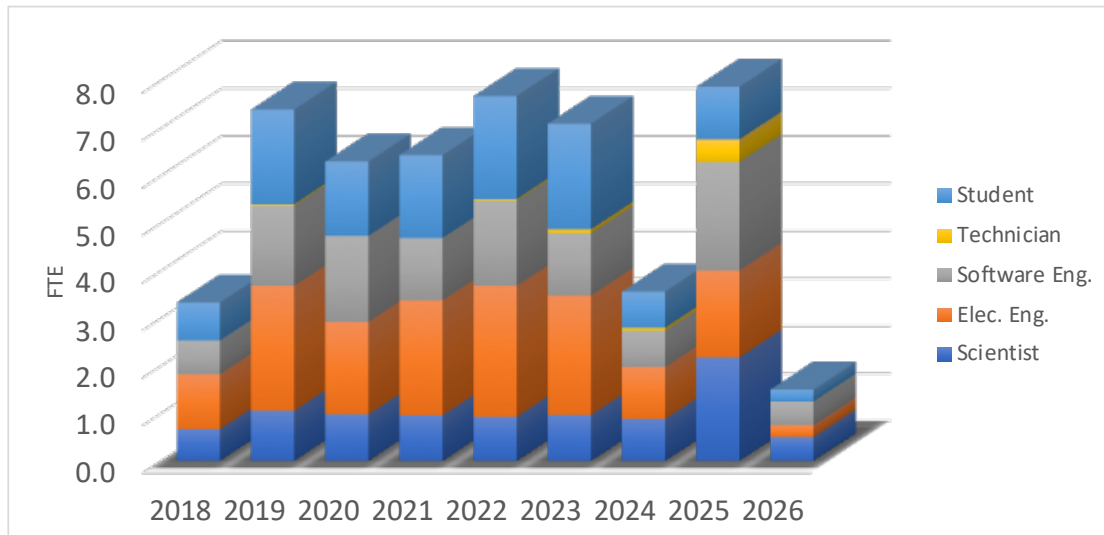
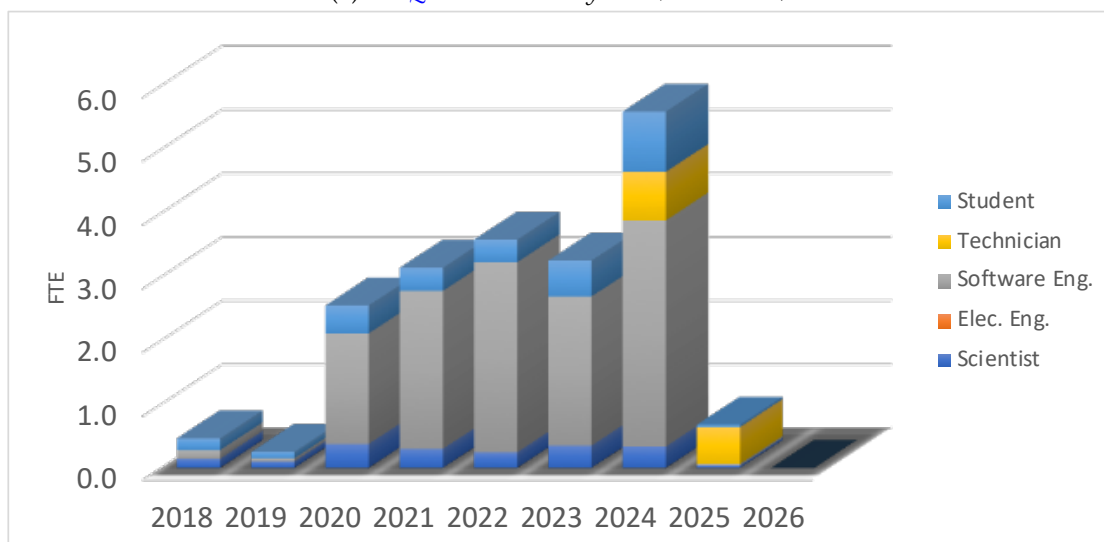
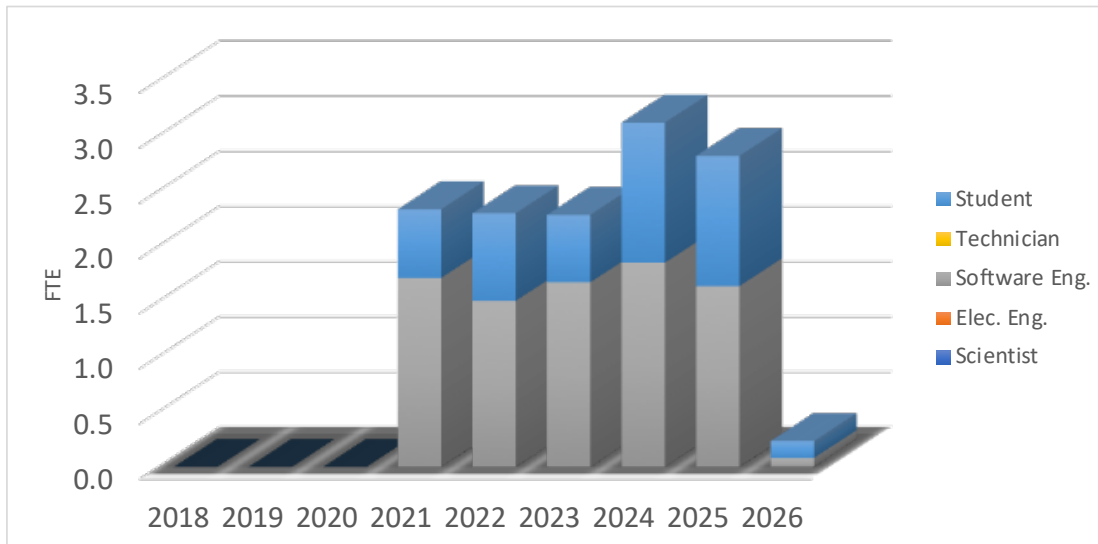
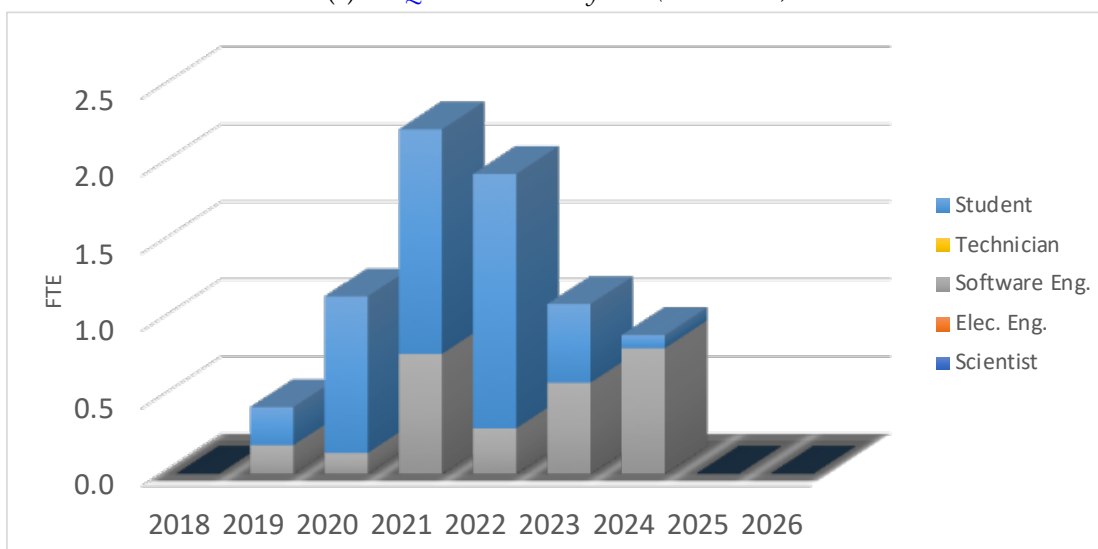
(a) *DAQ Readout sub-system (WBS 1.2.1)*(b) *DAQ Dataflow sub-system (WBS 1.2.2)*

Figure 20.5: Required resources expressed as FTEs for each year of the upgrade project's construction, installation and commissioning. Histograms are shown for the different sub-systems of the DAQ system. Four professional categories are stacked on top of each other: Scientists (blue), Electronics Engineers (orange), Software Engineers (grey), Technicians (yellow), and Students (light blue).

## 20.2 Required Manpower Estimate

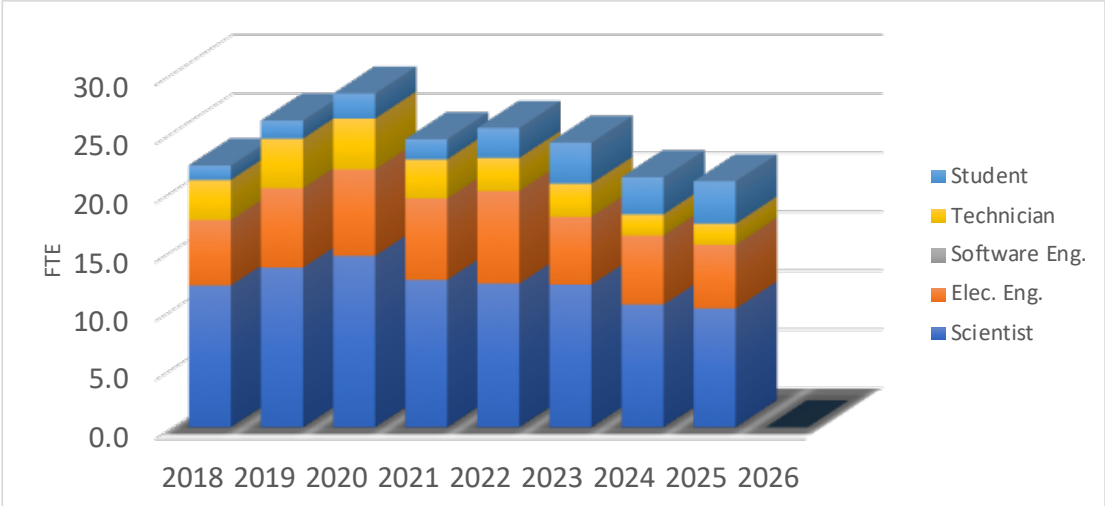


(c) DAQ Network sub-system (WBS 1.2.3)

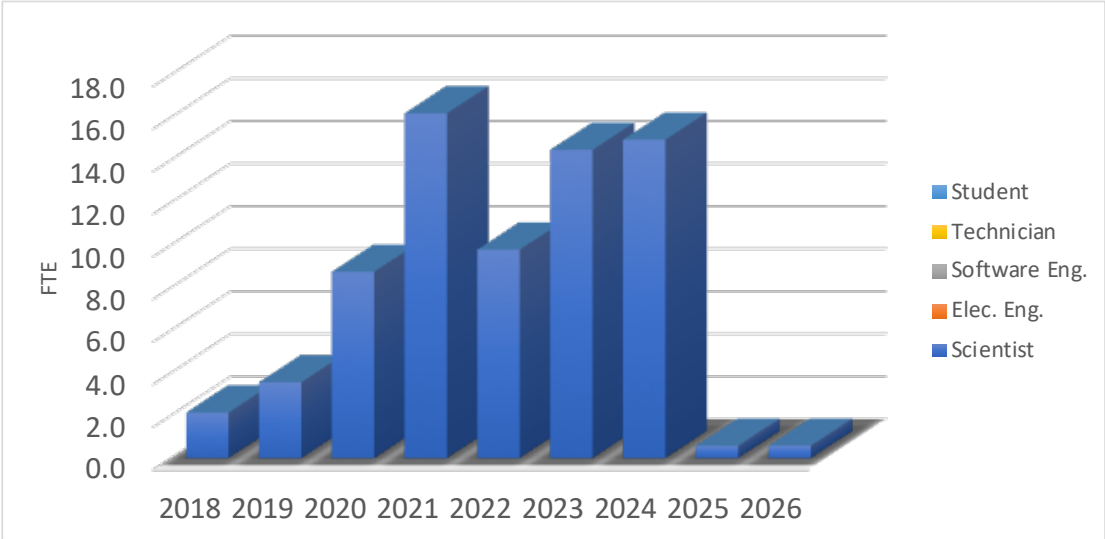


(d) DAQ Online Software sub-system (WBS 1.2.5)

Figure 20.5: Required resources expressed as FTEs for each year of the upgrade project's construction, installation and commissioning. Histograms are shown for the different sub-systems of the DAQ system. Four professional categories are stacked on top of each other: Scientists (blue), Electronics Engineers (orange), Software Engineers (grey), Technicians (yellow), and Students (light blue) (cont.)



(a) EF HTT sub-system (WBS 1.3.1)



(b) EF Processing sub-system (WBS 1.3.2)

Figure 20.6: Required resources expressed as FTEs for each year of the upgrade project’s construction, installation and commissioning. Histograms are shown for the two sub-systems (HTT, EFPU) of the EF system. Four professional categories are stacked on top of each other: Scientists (blue), Electronics Engineers (orange), Software Engineers (grey), Technicians (yellow), and Students (light blue).

## 20.3 Participating Institute Responsibilities

The formal institutional responsibilities, and the level of participation and financial support by each Institute and Funding Agency to the programme of the [TDAQ UPR](#), will be described in the [MoU](#) in preparation for the [RRB](#) meeting in the first half of 2018. In this [TDR](#), we limit ourselves to listing the the institutes participating in the [TDAQ Phase-II UPR](#) and their principal areas of interest.

Table 20.4 lists the Institutes interested in the upgrades of the Level-0 Trigger components. For each system and sub-system a ✓ indicates that the Institute plans to develop and/or provide hardware, firmware, and/or software deliverables. In this context any code that allows for the configuration, control, monitoring and operation of a hardware module during ATLAS data taking is categorised as software deliverable. For Phase-I legacy elements of the project, a ✓ indicates that the institute plans to maintain and/or upgrade firmware and software needed for them to operate in the HL-LHC environment.

Table 20.4: List of participating Institutes and their areas of interest within the Level-0 Trigger system.

Country Institute	Level-0 Trigger														
	Level-0 Calo					Level-0 Muon				Global		Central			
	eFEX	gFEX	jFEX	fFEX	Fibre Mgmt./Infrastr.	Barrel SL	Endcap SL	MDT TP	NSW TP	GCM, MUX, GEP	PFM	Fibre Mgmt./Infrastr.	CTP	MUCTPI	TTC
<b>Argentina</b> Buenos Aires La Plata										✓					
<b>Germany-BMBF</b> Heidelberg KIP Mainz	✓	✓	✓			✓	✓			✓			✓	✓	
<b>Germany-MPI</b> MPI-Munich								✓							
<b>Israel</b> Technion Tel Aviv Weizmann															

continue ...

... continued

Country Institute	Level-0 Trigger													
	Level-0 Calo					Level-0 Muon				Global		Central		
	eFEX	gFEX	jFEX	fFEX	Fibre Mgmt/Infrastr.	Barrel SL	Endcap SL	MDT TP	NSW TP	GCM, MUX, GEP	PFM	Fibre Mgmt./Infrastr.	CTP	MUCTPI
<b>Italy</b> Napoli Roma I Roma II						✓								
<b>Japan</b> KEK Kobe Kyoto Nagoya Shinshu Tokyo ICEPP Tokyo MU Tokyo Tech							✓	✓						
<b>Netherlands</b> NIKHEF										✓	✓			
<b>Poland</b> Cracow IFJ PAN Cracow AGH-UST/UJ										✓	✓			
<b>Romania</b> Bucharest									✓					
<b>United Kingdom</b> Birmingham Cambridge London QMUL RAL	✓				✓					✓	✓			
<b>USA-DOE</b> Argonne BNL		✓								✓	✓			

continue ...

## 20.3 Participating Institute Responsibilities

...continued

Country Institute	Level-0 Trigger														
	Level-0 Calo					Level-0 Muon				Global	Central				
	eFEX	gFEX	jFEX	fFEX	Fibre Mgmt./Infrastr.	Barrel SL	Endcap SL	MDT TP	NSW TP	GCM, MUX, GEP	PFM	Fibre Mgmt./Infrastr.	CTP	MUCTPI	TTC
<b>USA-NSF</b>															
Boston							✓								
Chicago		✓								✓					
Harvard									✓						
Indiana		✓								✓					
Michigan St.					✓					✓					
U Mass							✓	✓		✓					
Oregon		✓								✓					
Pittsburgh		✓								✓					
UC Irvine							✓			✓					
CERN													✓	✓	✓

Table 20.5 includes the Level-4 items of each DAQ sub-system in the PBS, and also elements of the ATLAS Online Software. The ATLAS Online Software is a large collection of software packages and libraries, which manages and controls the components of the DAQ system and the data-taking environment. Most of the software elements will be maintained and upgraded adiabatically in the next several years by the ATLAS DAQ Operation group. However, some of the core software infrastructure may be significantly redesigned to benefit from the availability of new technologies. Those elements are described in Section 11.7. In Table 20.5 a ✓ in the “Online” column indicates that either an institute is currently involved in operation, Phase-I or Phase-II upgrade tasks, or is planning to contribute in future.

Table 20.5: List of participating Institutes and their areas of interest within the DAQ system.

Country Institute	Data Acquisition (DAQ)						
	Readout		Dataflow			Networking	
	FELIX	Data Handlers	Storage Handler	Event Aggr.	Event Bld.	Readout Net. Dataflow Net. EF Net.	Online
<b>Australia</b> Adelaide Sidney	✓ ✓	✓ ✓					
<b>Czech Republic</b> Prague	✓						
<b>Germany-BMBF</b> Goettingen Wupperthal	✓ ✓	✓ ✓					
<b>Israel</b> Ben Gurion Tel Aviv Technion Weizmann	✓ ✓ ✓ ✓	✓ ✓ ✓ ✓		✓	✓		
<b>Italy</b> Bologna Pavia	✓	✓		✓	✓		
<b>Netherlands</b> NIKHEF	✓						
<b>Romania</b> IFIN-HH Bucharest							✓
<b>Russia</b> JINR Dubna NRC Kurchatov Inst. - PNPI							✓ ✓
<b>Switzerland</b> Bern	✓						
<b>United Kingdom</b> London RHUL London UC	✓	✓ ✓	✓				

continue ...

### 20.3 Participating Institute Responsibilities

... continued

Country / Institute		Data Acquisition (DAQ)								
		Readout		Dataflow			Networking			
		FELIX	Data Handlers	Storage Handler	Event Aggr.	Event Bld.	Readout Net.	Dataflow Net.	EF Net.	Online
<b>USA-DOE</b>										
Argonne		✓								
BNL		✓								
Winsconsin										✓
<b>USA-NSF</b>										
MSU										✓
UC Irvine		✓	✓							✓
<b>CERN</b>		✓		✓	✓	✓	✓	✓	✓	✓

Table 20.6 lists the Institutes interested in the components of the **HTT** sub-system and the **EF** hardware and software. A ✓ in the **EFPU** column indicates that the institute is contributing financially to the purchase of **EF CPUs**. Table 20.6 also lists Institutes contributing to and/or planning to contribute to the core **EF** software and to the algorithms for the online reconstruction of trigger-object candidates and studies of trigger menus. It is to be noted that both the **EF** reconstruction developments and the studies of trigger menus are activities not under the sole responsibility of the **TDAQ UPR**, but they are shared with the ATLAS Upgrade Physics and the ATLAS Trigger Activity groups.

Table 20.6: List of participating Institutes and their areas of interest within the Event Filter (EF) system.

Country / Institute		Event Filter								
		HTT				EF				
		Trk. Processors	RTM, HTTIF, Infr.	PRM	AM ASIC	TFM	SW/Monit.	Sim./Perf.	EFPU	Core SW
<b>Argentina</b>										

continue ...



... continued

Country Institute	Event Filter								
	HTT						EF		
	Trk. Processors RTM, HTTIF, Infr.	PRM	AM ASIC	TFM	SW/Monit.	Sim./Perf.	EPU	Core SW	Rec. SW/Menu
Buenos Aires La Plata									✓ ✓
<b>Australia</b> Melbourne		✓	✓						✓
<b>Canada</b> McGill Victoria									✓ ✓
<b>Denmark</b> Copenhagen NBI		✓							✓
<b>France-IN2P3</b> LPNHE Paris <b>Germany-BMBF</b> Heidelberg PI			✓						
<b>Israel</b> Ben Gourion Tel Aviv Technion Weizmann							✓ ✓ ✓ ✓		✓ ✓ ✓ ✓
<b>Italy</b> Genova Lecce Milano Pavia Pisa Roma 1		✓ ✓	✓		✓	✓			✓ ✓ ✓
<b>Japan</b> KEK Kobe Kyoto									✓ ✓ ✓

continue ...

### 20.3 Participating Institute Responsibilities

... continued

Country Institute	Event Filter								
	HTT						EF		
	Trk. Processors RTM, HTTIF, Infr. PRM AM ASIC TFM SW/Monit. Sim./Perf.						EFPU Core SW Rec. SW/Menu		
Nagoya Tokyo Tech									✓ ✓
<b>Netherlands</b> NIKHEF									✓
<b>Poland</b> Cracow IFJ PAN Cracow AGH-UST/UJ							✓ ✓	✓ ✓	✓ ✓
<b>Portugal</b> LIP							✓		✓
<b>Russia</b> Moscow State									✓
<b>Serbia</b> Belgrade									✓
<b>Spain</b> IFAE									✓
<b>Sweden</b> Uppsala			✓						
<b>Switzerland</b> Geneve			✓		✓	✓			✓
<b>United Kingdom</b> Edinburgh Glasgow London RHUL London UC Manchester RAL Sussex Warwick								✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓	✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓ ✓

continue ...

... continued

Country / Institute	Event Filter								
	HTT						EF		
	Trk. Processors RTM, HTTIF, Infr.	PRM	AM ASIC	TFM	SW/Monit.	Sim./Perf.	EPU	Core SW	Rec. SW/Menus
<b>USA-DOE</b>									
Argonne	✓					✓			
BNL							✓	✓	
SLAC							✓	✓	
SMU							✓		
<b>USA-NSF</b>									
Arizona	✓				✓				
Chicago					✓				
UIUC					✓				
NIU						✓			
NYU									
Oregon							✓	✓	
Penn	✓					✓			
SMU	✓						✓		
Stanford	✓								
<b>CERN</b>							✓	✓	✓



## 21 Risk Analysis and Mitigation Strategies

This chapter is an overview of the process developed by the [TDAQ UPR](#) together with the ATLAS Upgrade Project Office to assess, manage and mitigate risks associated with the project deliverables. At the time of the release of this [TDR](#), a documentation package on Risk Management, Risk Analysis and Control is being prepared for the [UCG Confidential Material](#), which will be made available for the [UCG](#) review in Q1-2018.

Risk Management has been fully integrated into the [TDAQ UPR](#) following the protocols and the guidelines summarised in [21.1], following accepted protocol standards outlined in Refs. [21.2], [21.3]. Risks in the [TDAQ UPR](#) are managed by a structured and integrated process for identifying, evaluating, tracking, mitigating, responding to, and managing project risks in terms of three risk categories: cost, schedule, and scope/technical performance.

The design and construction of the [UPR](#) are within the experience and expertise of the [TDAQ](#) participating collaborators, technical staff and scientists. It is a responsibility and a priority of the entire [UPR](#) management to reduce the risk to an acceptable low level for each hardware deliverable. Furthermore, as the [UPR](#) relies on a significant number of algorithms in both the Level-0 and the [EF](#) systems, risks associated to firmware and software are included in the Risk Management.

The [UCG Confidential Material](#) package contains a full description of the Risk Register, with risk identification, impact analysis and mitigation strategy. This chapter only summarises that information and it is organised in two overview sections: Section 21.1 is a summary of the [TDAQ UPR](#) Risk Management Plan [21.1]; Section 21.2 lists a few items extracted from the Risk Register, with a short description of the impact and of the possible mitigation strategies associated.

### 21.1 Risk Management Plan

The overall Risk Management approach is described in the Risk Management Plan [21.1], currently in preparation, and consists of a five-step process:

- Identifying potential project risks,
- analysing project risks,
- planning risk mitigation and response strategies,
- executing risk mitigation and response strategies, and

## 21.1 Risk Management Plan

---

- monitoring and tracking the results and revising the risk mitigation and response strategies.

In this context, a risk is any event not included in the project plan which can have a potential impact on the [TDAQ UPR](#). Every effort has been made to specify the project in a manner that reduces the risk to an acceptable low level. The technical risks to the project that are identified will be addressed as early as possible to assure that they do not impact the timely completion of the project or stress its budget in unexpected ways. Proactive risk identification and mitigation can therefore significantly reduce the probability of unexpected events that could require contingency and/or additional time to resolve.

The system (Level-2) and sub-system (Level-3) coordinators have the responsibility for managing and mitigating the risks associated with their respective [WBS](#) areas. They are responsible to identify potential sub-system risks, to analyse these risks, to develop mitigation and response strategies for these risks, and to monitor and track them. The [UPL](#) is assisted by the [UPR](#) Resource and Risk Manager ([UPR-RC](#)), who is formally delegated to coordinate this matter, oversees these processes at the project level, and also has the responsibility for implementing response plans. Risks are reviewed periodically (quarterly) meetings of the [eTDSG](#), in which risks are discussed, updated, and appropriate actions are taken if required.

- The mitigation steps that are/will be taken to minimise that risk from occurring;
- the response to the risk in the eventuality that the risk materialises;
- the impact of the risk on cost, schedule, and performance: the risk impact is classified as Negligible, Low, Medium, or High based on its potential impact on cost, schedule, or performance, see [Table 21.2](#). The performance impact is based on identifying a set of Key Performance Parameters ([KPP](#)) and assessing the scientific impact of the risks on those [KPPs](#);
- the risk probability, identified as Low, Moderately Low, Moderately High, or High based on the probability ranges identified in [Table 21.1](#).
- the most probable impact on the cost and schedule, and optimistic/pessimistic impacts quantifying the best/worst case scenarios;
- the overall assessment of the risk based on [Table 21.3](#) that correlates the probability of the risk to occur and the impact of the risk. The overall risk is classified as High, Medium, or Low based on the product of the risk impact and the risk probability.

Table 21.1: *Classification of the risk occurrence's probability.*

Risk Probability	Probability range	
	Min [%]	Max [%]
Low	0	5
Moderately Low	5	15
Moderately High	15	30
High	>30	

Table 21.2: Classification of the risk impact based on its impact on the cost, the schedule, and the scope/performance.

Risk Impact		Cost Impact Range [kCHF]		Schedule Impact Range [months]		Scope and Performance Impact Assessment: Objective KPP variance
		Min	Max	Min	Max	
Negligible	0	0	20	0	1	<1%
Low	1	20	100	1	3	1–5%
Medium	2	100	500	3	6	5–10%
High	3	>500		>6		>10%

<sup>1</sup> for finding maxima

Table 21.3: Correlation of Risk Probability (Rows, classified as 1-4) and Risk Impact (columns, classified as Negligible, Low, Medium). The identified risk is subsequently classified as Low (if the correlation falls within the green shaded area), Medium (yellow shade), or High (red shade).

High	4	0	4	8	12
Moderately high	3	0	3	6	9
Moderately low	2	0	2	4	6
Low	1	0	1	2	3
		0	1	2	3
		Negligible	Low	Medium	High

## 21.2 Risk Register

The process of identifying risks for the [TDAQ UPR](#) has followed the procedures summarised in Section 21.1: initial brainstorming sessions and dedicated management meetings between the [UPR-RC](#) and the Level-2,3 managers have performed a top-down high-level risk assessment, quantitative risk analysis with evaluation of probabilities and impacts and discussions of mitigation strategies. A detailed set of risks, as well as additional material on the risk mitigation and response plans is being prepared in a Risk Register [21.4] that will be made available at the beginning of 2018 in preparation for the [UCG](#) review. For illustrative purposes only the Network part of the Risk Register is shown in Figure 21.1.





In this Section only the elements identified at this stage with potentially particular significance are presented.

### 21.2.1 Detector Readout Limitations

The design of the detector's Front-End electronics shall satisfy the 1 MHz L0 rate requirement of the baseline architecture with sufficient margins. However, for certain detector systems, in particular for the innermost ITk-pixel layers, the front-end ASIC's design is specified assuming an estimated occupancy potentially subject to large uncertainties. Should the occupancy greatly exceed expectations, the increase in event size would exceed the bandwidth for the pixel detector front-end readout. Similarly, there are risks that the ITk-pixel front-end data transmission will not operate successfully at the specified 5 Gb/s max. throughput for a Level-0 trigger rates of 1 MHz. The mitigation strategy adopted is explained in detail in Chapter 14, with the system evolving to a dual L0/L1 trigger architecture using regional tracks built from the ITk strip detector and outer pixel detector layer information at Level-1.

### 21.2.2 Projected rates for hadronic trigger signatures

Another risk mitigated by the same strategy mentioned in Section 21.2.1 is due to the uncertainty in the projected trigger rates for hadronic objects at  $\langle \mu \rangle = 200$ . The use of tracking at the L1 trigger stage would open up the possibility for a rudimentary primary vertex selection for multijet triggers, allowing for additional hadronic background rejection and a lower overall readout rate for the inner detector pixel layers. Additional details regarding the motivation and criteria for evolution are presented in Section 14.1, followed by the requirements for the evolved system in Section 14.2 and a description of the resulting "evolved" architecture design in Section 14.3.

### 21.2.3 Resource availability of the EF Processing Unit farms

The estimate of the required EF compute power is subject to large uncertainties as described in Chapter 12. The probability of such a risk is relatively low, once the expected luminosity growth in the initial phases of the HL-LHC are taken into account. However, there is a risk that the ATLAS M&O rolling replacement doesn't provide the expected computing power, as described in Chapter 12. To mitigate the risk, the resources required for  $\langle \mu \rangle = 200$  are planned to be available from the beginning of Run 4. There are also software development milestones to check the software performance and ensure that it is on track, and an evaluation of hardware accelerators prior to the decision of which commodity processing technologies to purchase.

### 21.2.4 Hardware Tracking ASIC

Two possible risks of the [HTT ASIC](#) have been considered: possible failures during the prototype and design phase may require an additional prototype cycle. The impact might be significant on the schedule of the project, but wouldn't have any impact on the CORE costs of the project. Production-related failures, e.g. very low yield due to marginal design, might require an iteration of the design's cycle, or the submission of a new set of masks, or the purchase of extra wafers. The impact might be high in terms of costs and of schedule. Possible mitigations are schedule advancing, carefully planning prototype, and pre-production, and including two mask-sets in the cost planning.

### 21.2.5 FPGA resource usage

During the design phase of a sub-system that makes significant use of [FPGAs](#), the resource consumption of algorithms to be deployed onto the device might exceed the expected usage. Careful design, timely development of the firmware, and detailed analysis of the resource's performance and usage should keep the probability of this risk low. In case of occurrence, the impact might be significant on the costs of the project, as [FPGAs](#) are likely to be main cost-driver. The mitigation strategy would be to design the sub-system suitable for [FPGA](#) resource changing and to define the specifications for algorithms such that they respect the resources of the chosen [FPGAs](#).

### 21.2.6 Interfaces and Link speeds

The link speed foreseen in the [TDAQ](#) electronics is largely available in today's available technology. However, some sub-systems are planning to use higher link speeds, up to 30+ Gb/s: examples are the interfaces between the LAr and Tile calorimeter Pre-processors and the Global Trigger designed to operate in the 16-25.8 Gb/s range, and the internal links between the Global [MUX](#) and [GEP](#) modules running at 30+ Gb/s. Significant engineering developments are needed to evaluate this type of [MGT](#) measuring [Bit Error Rate \(BER\)](#) in dedicated system tests. Possible mitigation is to design the sub-system which is feasible either to downscale the [MGT](#) bandwidth and increase the number of connections, or to increase the number of fibres and modules, with possible significant impacts on costs.

### 21.2.7 Latency limitations

The design described in this [TDR](#) defines clear envelopes for the latency available at each stage of the trigger data processing, guaranteeing sufficient margins on overall [L0](#) latency. Some of the object reconstruction in the processing [FPGAs](#) might risk to exceed the envelope, reducing or eliminating all the pre-defined margins. The probability of the risk should

be moderately low. There shouldn't be any impact on the cost. However, the system's performance penalty could be high as less powerful algorithms might need to be deployed for the trigger's candidate reconstruction.

### 21.2.8 FPGA obsolescence

The CORE cost estimates are based on [FPGA](#) devices with sufficient resources currently on the market, with sufficient resources for the functions required to be implemented in each application. There is a moderately low probability that the pre-selected [FPGA](#) may become obsolescent and unavailable by the time procurement has to be placed. There may be some moderate impact on the cost of the sub-system if a new generation and more performant device has to be selected.

### 21.2.9 Commissioning time of the Level-0 Muon Trigger sub-system

The [L0](#) Muon contains hardware and firmware elements that may require a commissioning time longer than anticipated before becoming fully operational. The resources available at the [TFM](#) should minimise the probability of occurrence of this risk for what concerns firmware. If required, as further mitigation strategy, the sub-system should guarantee additional experienced manpower to finalise the firmware design to deploy on the production units in [USA15](#). The impact of the risk may be limited by the fact that the [L0](#) Muon sub-system may concentrate the effort on commissioning and operating the Sector Logic alone.

### 21.2.10 Commissioning time of the Global Trigger sub-system

Similarly, the Global sub-system is totally a new set of hardware and firmware elements that may require a commissioning time longer than anticipated before becoming fully operational. The risk impact on the ATLAS operations may be limited by the fact that the legacy [FEXs](#) and the legacy [L1Topo](#) modules will be fully available and operational at the beginning of Run 4 if needed.

### 21.2.11 Readout's I/O bandwidth margins

The [FELIX](#) design is optimised to accommodate and manage with large margins the full input bandwidth from the detector's point-to-point sources. The probability of the risk of having the data input bandwidth close to the limit should be negligible by design. The risk of having a less performing system, i.e. the data throughput from the custom I/O cards to the hosting server's [CPU](#) doesn't match the input bandwidth, shall be low. The mitigation

would be to redimension the Readout sub-system accordingly with a possible significant impact on the costs of the sub-system.

### 21.2.12 Data Handler processing capabilities for detector-specific functionality

As the detector requirements for data management in the second stage of the Readout sub-systems have not been defined yet, the system described in Chapter 11 may not be sufficient and additional Data Handler nodes would be required. The impact in terms of CORE costs might not be negligible. However, as the architecture is network-based, and therefore fully scalable, and therefore no technical or schedule impact.

### 21.2.13 Event Size

The risk that the ATLAS upgrade would result into larger event size is moderately high. There are several examples that could cause an increase of the event size: (i) the scope of the ATLAS upgrades increases with new detectors and upgrade project; (ii) data compression in the front-end ASICs doesn't perform as expected. Cost impact on the Dataflow and Network sub-systems might be non-negligible. Mitigation technique would be to develop stronger compression at the level of Readout, e.g. in the Data Handlers.

### 21.2.14 Simulation software schedule

The EF software depends on timely delivery of simulation of the trigger upgrades and detector data. EF software components require realistic input data in order to develop and test them. In the case where the sources of these data are new or upgraded trigger or detector subsystems, their data needs to be provided through simulation. This is mitigated by connecting the plans of the EF software with those for simulating the upgrades. However, with the number of new data sources for the EF in the overall upgrade, the risk is moderately high that a delay will occur in at least one of them. This would have an impact on the schedule and/or performance of the EF software, which depends on the length of the delay to the simulation and which components are delayed.

## References

- [21.1] The TDAQ Group, *TDAQ Phase-II Upgrade Risk Management Plan*, Tech. Rep. ATL2-D-MG-0004, EDMS Id 1874388 v.1, CERN, Geneva, Nov, 2017.  
<https://edms.cern.ch/document/1874388/1>.

- [21.2] J. Makepeace, *DOE G 413.3-7A, Risk Management Guide*, tech. rep., DOE, 2011.  
<https://www.directives.doe.gov/directives-documents/400-series/0413.3-EGuide-07a>.
- [21.3] *A Guide To The Project Management Body Of Knowledge (PMBOK Guides)*. Project Management Institute, 2004.
- [21.4] The TDAQ Group, *TDAQ Phase-II Upgrade Risk Register*, Tech. Rep. ATL2-D-ES-00002, EDMS Id 1856380 v.1, CERN, Geneva, Oct, 2017.  
<https://edms.cern.ch/document/1856380/1>.



## **Part IV**

### **Appendix: Glossary**





---

<b>AFP</b>	ATLAS Forward Proton Detector . . . . .	105, 295
<b>AM</b>	Associative Memory . . . . .	69, 93, 102, 111, 112, 118–120, 216, 221–223, 228, 343, 344, 346, 348, 349, 355, 362, 370, 372–378, 382, 387–390, 411, 413–416, 486, 490, 510, 528–531, 545
<b>AM06</b>	Fast-Track Associative Memory ASIC . . . . .	376–378, 381
<b>AM07</b>	Associative Memory 07 prototype ASIC . . . . .	119, 376, 381
<b>AM08</b>	Associative Memory 08 prototype ASIC . . . . .	120, 375–383, 390, 510
<b>AM09</b>	Associative Memory 09 ASIC . . . . .	118, 120, 376–383, 390, 510
<b>AM09pre</b>	Associative Memory 09 pre-production ASIC . . . . .	378, 390, 510
<b>AMTP</b>	AM Trigger Processor . . . . .	118, 119, 345–348, 362, 364–366, 369, 370, 373, 374, 386–390, 411, 414, 416, 418, 419, 443, 486
<b>APE</b>	Accelerator Process Extension . . . . .	334
<b>ARM</b>	Advanced RISC Machines . . . . .	311
<b>ASIC</b>	Application-Specific Integrated Circuits . . . . .	62, 64, 76, 77, 81, 82, 98, 99, 111, 112, 118–120, 197, 225, 343, 344, 346, 348, 349, 362, 370, 373–378, 382, 387–390, 416, 418, 419, 423, 424, 431, 479, 486, 490, 491, 494, 500, 510, 528–531, 537, 538, 540
<b>ASSO</b>	ATLAS activity Systems Status Overview review . . . . .	498
<b>ATCA</b>	Advanced Telecommunications Architecture . . . . .	xix, 82, 83, 92–99, 118, 119, 196, 204, 205, 212, 213, 216, 226, 258–260, 272–274, 279, 282–284, 311, 344, 345, 362, 366, 367, 369, 370, 372, 386–389, 404, 416, 425–427, 430, 431, 434–438, 442, 443, 445–449, 483, 484, 486, 487, 493, 494
<b>ATCN</b>	ATLAS Control Network . . . . .	430, 431
<b>BC</b>	Bunch Crossing . . . . .	78, 80–82, 236–240, 262, 545
<b>BCID</b>	Bunch Crossing IDentifier . . . . .	82, 297
<b>BCMS</b>	Batch Compression Merging and Splitting . . . . .	7
<b>BDT</b>	Boosted Decision Tree . . . . .	132, 133
<b>BE</b>	Back-End . . . . .	11, 421, 422, 425, 427, 428, 430, 431
<b>BER</b>	Bit Error Rate . . . . .	538
<b>BOE</b>	Basis of Estimate . . . . .	480
<b>BSM</b>	Beyond the Standard Model . . . . .	9, 10, 33, 36, 416
<b>CAM</b>	Content-Adresseable Memory . . . . .	275
<b>CAN</b>	Controller Area Network . . . . .	421, 423, 426
<b>CB</b>	ATLAS Collaboration Board . . . . .	458–460

---

<b>CBE</b>	Current Best Estimate: the value of a resource or of a parameter extracted through simulations and/or extrapolation of existing data to the conditions foreseen in the <a href="#">HL-LHC</a> phase and/or best current knowledge of what will be the ATLAS experimental layout . . . . .	<a href="#">226, 227, 262, 263, 446, 481, 551</a>
<b>CORE</b>	COsting REview; costing policy established by the LHCC at the time of the original construction . . . . .	<a href="#">xxi, 463, 470, 478–481, 486–495</a>
<b>COTS</b>	Commercial Off-The-Shelf electronics component . . . . .	<a href="#">421, 481, 493</a>
<b>CPU</b>	Central Processing Unit . . . . .	<a href="#">xviii, 2, 13, 15, 40, 65, 69, 71, 89, 111, 112, 114–116, 118, 157, 158, 219, 220, 306, 307, 312–314, 329–340, 342, 343, 480, 492, 528, 539, 548</a>
<b>CSC</b>	Cathode Strip Chamber . . . . .	<a href="#">194</a>
<b>CSM</b>	Chamber Service Module . . . . .	<a href="#">213, 225</a>
<b>CTP</b>	Central Trigger Processor . . . . .	<a href="#">2, 15, 17, 40, 42, 58, 60, 62, 63, 81, 87–89, 91–100, 102, 105, 113, 163, 167, 173, 235, 237–240, 257, 258, 262, 263, 265, 269–273, 278–286, 295, 399, 404–406, 408, 430, 436, 442, 446, 469, 472, 482, 483, 507, 518, 524–526, 546, 551, 554, 555</a>
<b>CTP Interface</b>	Global-to-CTP Interface	<a href="#">95, 99, 100, 235, 237–242, 255–258, 260, 262, 265, 402</a>
<b>CTPCORE</b>	<a href="#">CTP CORE</a> Processor . . . . .	<a href="#">483, 487</a>
<b>CTPIN</b>	<a href="#">CTP Input</a> Interface . . . . .	<a href="#">483, 487</a>
<b>CTPMI</b>	<a href="#">CTP Machine</a> Interface . . . . .	<a href="#">483, 487</a>
<b>CUDA</b>	Compute Unified Device Architecture . . . . .	<a href="#">334</a>
<b>DAQ</b>	Data Acquisition System	<a href="#">2, 13–17, 19, 20, 36, 39, 42, 44, 63, 83, 87–89, 101, 103, 105, 158, 225, 226, 289–291, 295–297, 306, 307, 310, 311, 318, 323, 324, 333, 397, 407, 433, 434, 436, 438, 439, 446, 450, 453, 457, 463, 465, 468–470, 472, 473, 482, 484, 485, 488, 489, 491–494, 502, 509, 510, 514, 516, 521, 522, 526–528</a>
<b>Data Handler</b>	Data Handler subsystem . . . . .	<a href="#">88, 89, 101–103, 418</a>
<b>Dataflow</b>	Dataflow subsystem . . . . .	<a href="#">2, 14, 15, 17, 84, 85, 88, 89, 113, 290, 436</a>
<b>DC</b>	“Don’t Care” . . . . .	<a href="#">351, 352, 354, 355</a>
<b>DCS</b>	Detector Control System . . . . .	<a href="#">16, 83, 84, 92–97, 101–103, 109, 172, 175, 176, 241, 284, 290–292, 297, 306, 308, 316, 323, 421–428, 430, 431, 470, 471, 485</a>
<b>DFS</b>	Distributed FileSystem . . . . .	<a href="#">107, 304–306</a>
<b>DO</b>	Data Organiser . . . . .	<a href="#">370, 373, 384, 413</a>
<b>DSP</b>	Digital Signal Processor . . . . .	<a href="#">248, 250, 252, 385</a>
<b>DSS</b>	Detector Safety System . . . . .	<a href="#">421</a>

<b>DUPL</b>	Deputy Upgrade Project Leader . . . . .	467, 468
<b>e-link</b>	Electrical chip-to-chip interconnect . . . . .	424, 425
<b>EB</b>	ATLAS Executive Board . . . . .	458–460
<b>EDMS</b>	Engineering & Equipment Data Management Service . . . . .	482, 497, 504
<b>EF</b>	Event Filter . . . . .	xviii, 2, 13, 15–17, 75, 84–89, 104, 107, 109–118, 236, 240, 244, 249, 289–291, 293, 297, 300, 307, 312, 313, 318–320, 329–334, 336, 337, 339, 340, 343–345, 347, 362, 387, 390, 399, 409–411, 417, 419, 430, 433, 434, 436, 438, 439, 443, 453, 465, 468–470, 473, 474, 482, 485, 486, 488, 490–492, 495, 509, 511, 517, 523, 528–531, 533, 537
<b>eFEX</b>	electron Feature EXtractor . . . . .	15, 17, 40, 45–47, 83, 90–92, 99, 100, 124, 126, 127, 132, 134, 165, 166, 168–175, 240, 242, 257, 259, 430, 436, 442
<b>EFPU</b>	Event Filter Processing Unit . . . . .	113, 313, 315, 330, 344, 347, 348, 386, 411, 415, 443, 450, 469, 473, 488, 491, 494, 511, 523, 528–531
<b>ELMB</b>	Embedded Local Monitor Board . . . . .	421, 423
<b>EM</b>	electromagnetic . . . . .	47, 52, 394
<b>EMEC</b>	LAr Electromagnetic EndCap Calorimeter . . . . .	90, 91, 167, 169, 171, 172
<b>eTDSG</b>	Extended TDAQ Steering Group . . . . .	465, 467, 468, 470, 498, 503, 534
<b>Event Filter</b>	Event Filter . . . . .	235, 236, 239, 241, 290
<b>FA</b>	Funding Agency . . . . .	459, 460
<b>FCal</b>	LAr Forward Calorimeter . . . . .	12, 48, 51, 52, 80, 90, 164, 167, 169, 171, 172, 244
<b>FCNC</b>	Flavour-Changing Neutral Currents . . . . .	10
<b>FDR</b>	Final Design Review . . . . .	230, 414, 440, 500–502, 511
<b>FE</b>	Front End . . . . .	xix, 11, 75, 76, 80, 82, 101, 103, 168, 290–293, 297, 395, 396, 398, 400, 421–425, 427, 428
<b>FEB</b>	Front End Board . . . . .	64
<b>FELIX</b>	Front-End LInk eXchange . . . . .	15–17, 42, 43, 81, 82, 88, 89, 92–97, 101–105, 109, 172, 174–176, 180, 197–199, 205, 213, 225, 239, 241, 257, 258, 264, 265, 269, 270, 272, 278, 279, 281, 283, 284, 286, 290–299, 306, 308, 311, 324, 386, 396–405, 407, 408, 410, 412, 418, 419, 422–425, 427, 433, 436, 442–444, 446, 472, 483, 485, 539
<b>FEX</b>	L1Calo Feature EXtractor . . . . .	42, 44, 45, 58, 66, 77, 79–82, 90, 91, 95, 102, 163, 164, 166–175, 236, 237, 239, 241, 242, 401, 437, 442, 449, 468, 507, 539
<b>fFEX</b>	forward Feature EXtractor . . . . .	1, 15, 17, 33, 87, 90–92, 99, 100, 148, 163–165, 167–169, 171, 172, 174–176, 244, 257, 430, 436, 442
<b>FMC</b>	FPGA Mezzanine Card . . . . .	367, 369, 370, 372

---

<b>FOX</b>	Fibre-Optic eXchange Plant .....	165, 168–170, 175
<b>FPGA</b>	Field Programmable Array .....	48, 58, 69, 90, 93, 95, 98, 99, 102–104, 111, 112, 114, 118, 142, 171–173, 182, 196–199, 203–206, 212, 213, 216, 218–220, 222, 223, 226, 231, 235, 236, 238, 239, 242–244, 246, 248–252, 255, 258, 260, 265, 271, 273–277, 279–285, 294, 297, 311, 329, 332, 333, 343, 344, 346, 347, 352, 353, 360, 362, 367–370, 372–378, 380, 382, 383, 385, 386, 388, 390, 403, 414–416, 419, 431, 440, 443, 481, 487, 490, 491, 500–502, 538, 539, 547
<b>FSM</b>	Finite State Machine .....	422
<b>FTE</b>	Full Time Equivalent .....	514–523
<b>FTK</b>	Fast TracKer	40, 42, 68–71, 117, 118, 221, 222, 235, 329, 333, 334, 338, 344, 345, 347–349, 376, 385, 419, 463
<b>GBT</b>	Gigabit Bidirectional Trigger and Data Link ..	76, 213, 292–294, 403, 422, 424, 425, 551
<b>GBT-SCA</b>	GBT Slow Control Adapter .....	423, 424
<b>GCM</b>	Global Common Module .....	99, 235, 255–261, 263, 264, 419, 509, 524–526
<b>GCS</b>	Global Control Station .....	421, 422
<b>GEP</b>	Global Event Processor ..	80, 95, 99, 100, 235–244, 246, 247, 249–251, 255–262, 264–266, 401–403, 419, 440, 443, 487, 524–526, 538
<b>gFEX</b>	global Feature EXtractor ...	15, 17, 40, 45, 48–51, 90–92, 99, 100, 135–137, 145, 147, 165, 167–175, 240, 242, 246, 253, 257, 259, 430, 436, 442
<b>gHTT</b>	Full-scan global hardware-based tracking for the trigger that is part of the Event Filter System .....	2, 15, 17, 20, 110, 114, 117–119, 133, 140, 145, 146, 151–153, 155, 299, 330, 333, 337, 338, 340, 343, 344, 347, 354, 357, 360, 362, 363, 366, 385, 387, 388, 409–411, 413, 417
<b>Global Trigger</b>	Global Trigger system .....	1, 2, 15, 17, 20, 33, 79, 81, 85, 87–96, 98–100, 102, 105, 113, 116, 117, 123, 124, 126, 132, 133, 135, 142, 145, 147, 150, 157, 158, 163–165, 167–169, 171–176, 235–243, 249, 255, 257, 260–266, 269–271, 275, 281–286, 295, 333, 334, 339, 396, 398–405, 408, 411, 418, 419, 436, 440, 446, 449, 487
<b>GPGPU</b>	A General Purpose Graphics Processing Unit is a <a href="#">Graphics Processing Unit (GPU)</a> which performs non-specialised calculations that would normally be performed by a <a href="#">CPU</a> . .....	111, 114–116, 329, 332–337, 343
<b>GPU</b>	Graphics Processing Unit .....	548
<b>HCM</b>	Hit Count Memory .....	370, 384
<b>HDD</b>	Hard Disk Drive .....	108, 300, 301
<b>HEC</b>	LAr Hadronic EndCap Calorimeter .....	90, 91, 167, 169, 172

---

<b>HGTD</b>	High-Granularity Timing Detector . . . . .	12, 79
<b>HI</b>	Heavy-Ion . . . . .	35, 36
<b>HL-LHC</b>	High Luminosity LHC . . . . .	iii, 1, 2, 5–11, 13, 19, 20, 22, 24, 29, 32, 35, 37, 39, 40, 44, 50–53, 62, 65, 66, 68, 70, 75, 79, 81, 83, 111, 112, 144, 156–158, 177, 178, 180–182, 191, 193–195, 207, 222, 223, 227–230, 269, 343, 393, 394, 396, 438, 441, 443, 465, 466, 469, 474, 479, 488, 491, 493, 537, 546
<b>HLM</b>	Hit List Memory . . . . .	370, 384
<b>HLP</b>	Hit List Pointer . . . . .	370, 384
<b>HLT</b>	High-Level Trigger . . . . .	39, 40, 42, 43, 45, 48, 54, 55, 67–72, 84, 117, 124, 126, 236, 299, 313, 314, 316, 333, 338, 418, 441, 450, 463, 469, 472, 473
<b>HPC</b>	High Performance Computing . . . . .	307, 308
<b>HTT</b>	Hardware-based Tracking for the Trigger . . . . .	2, 13, 15–17, 83, 86, 88, 89, 111–115, 117–119, 123, 150, 155, 157, 158, 221, 330, 334, 338–340, 343–349, 354, 360, 362–366, 369, 374, 376, 378, 381, 383, 386–390, 397, 399, 410–413, 415, 417–419, 430, 433, 434, 436, 443, 446, 469, 473, 474, 485, 486, 490, 494, 510, 523, 528–531, 538, 549, 553, 555
<b>HTTIF</b>	HTT Interface . . . . .	113, 118, 119, 299, 345–348, 362–367, 369, 386–388, 413, 443, 528–531
<b>Hub</b>	Common readout infrastructure for <b>L1Calo</b> . The <b>ROD</b> is a daughter card.	401
<b>I<sup>2</sup>C</b>	I2C . . . . .	431
<b>IB</b>	System Institutional Board . . . . .	460, 461
<b>ID</b>	Inner Detector . . . . .	337–340
<b>IDR</b>	Initial Design Review . . . . .	13, 460
<b>IPBus</b>	IP-based protocol implementing register-level access over Ethernet for module control & monitoring. . . . .	426, 427, 431
<b>IPC</b>	Interprocess communication . . . . .	336
<b>IPMC</b>	IPMI Management Controller . . . . .	82, 83, 212, 213, 284, 427, 430, 431
<b>IPMI</b>	Intelligent Platform Management Interface . . . . .	309, 316, 430
<b>ISR</b>	Initial State Radiation . . . . .	10, 24, 26–30
<b>ITk</b>	Inner Tracker . . . . .	5, 11, 13, 16, 46, 69, 75–77, 82, 84–87, 89, 101, 103, 105, 111–115, 117, 118, 123, 142, 151, 155, 157, 294, 295, 297, 330, 333, 339, 340, 343, 344, 346, 347, 349, 350, 352, 354, 355, 360, 362–364, 370, 383, 386, 393, 395–403, 407, 408, 411, 419, 444, 460, 480, 493, 537
<b>iUPL</b>	Interim Upgrade Project Leader . . . . .	467, 468

---

<b>JCOP</b>	Joint COntrols Project . . . . .	426, 428
<b>jFEX</b>	jet Feature EXtractor . . . . .	15, 17, 40, 45, 47–51, 90–92, 99, 100, 135–137, 165–175, 240, 242, 246, 253, 257, 430, 436, 442
<b>JTAG</b>	Joint Test Action Group (developer of IEEE Standard 1149.1-1990) . . . . .	502
<b>KPP</b>	Key Performance Parameter. . . . .	534, 535
<b>L0</b>	Level-0 Trigger 82, 93, 103, 112, 113, 343, 433, 434, 436, 438, 446, 449, 465, 468, 515, 537–539	
<b>L0/L1</b>	Level-0/Level-1 Trigger Architecture . . . . .	434
<b>L0A</b>	Level-0 Trigger Accept . . . . .	80–82, 88, 89, 93, 95, 96, 104, 118, 163, 174, 225, 241, 272, 275, 277–279, 281, 283, 284, 293, 296, 299, 343, 396–399, 401–404, 407, 411–413, 415, 416
<b>L0Calo</b>	Level-0 Calorimeter Trigger (Run 4 and beyond) . . . . .	1, 2, 15, 17, 77, 79, 81, 82, 87–92, 94, 95, 99, 100, 102, 105, 163, 165, 236, 239, 241, 262, 263, 295, 430, 436, 442, 443, 446, 449
<b>L0CTP</b>	Level-0 Central Trigger Processor . . . . .	272, 280, 292, 402–406, 408
<b>L0ID</b>	Level-0 IDentifier . . . . .	297
<b>L0Muon</b>	Level-0 Muon Trigger System (Run 4 and beyond) . . . . .	1, 2, 15, 17, 20, 36, 78, 81, 82, 87–90, 93, 98–100, 102, 236, 239–241, 262, 263, 269, 282, 283, 286, 436, 442, 446–448, 487
<b>L1</b>	Level-1 Trigger . . . . .	112, 343, 463, 537
<b>L1A</b>	Level-1 Trigger Accept . . . . .	42, 65, 69, 225, 397–403, 405, 407–409
<b>L1Calo</b>	Level-1 Calorimeter Trigger (through Run 3) . . . . .	42, 43, 51, 58, 66, 87, 90, 103, 163, 270, 271, 437, 438, 549
<b>L1CTP</b>	Level-1 Central Trigger Processor . . . . .	272, 280, 400, 402, 404–406, 408, 418, 419
<b>L1Muon</b>	Level-1 Muon Trigger (through Run 3) . . . . .	42, 103
<b>L1Topo</b>	Level-1 Topological Processor . . . . .	42, 58, 59, 80, 94–96, 164, 167, 168, 171, 173–176, 235, 237, 265, 266, 539
<b>L1Track</b>	Level-1 Track Trigger (Phase-II) 257, 360, 374, 399–403, 407–412, 414–416, 418, 419, 553	
<b>LAN</b>	Local Area Network . . . . .	430, 556
<b>LAr</b>	Liquid Argon Calorimeter 11, 12, 42–44, 59, 64–66, 77, 84, 87, 89–91, 102, 103, 105, 124, 126, 163–165, 168, 170, 172, 174, 240, 243, 262, 295, 449	
<b>LASP</b>	LAr Signal Processor . . . . .	77, 91, 94, 105, 169, 170, 176, 240, 246, 257, 295
<b>LCG</b>	LHC Computing Grid . . . . .	311

---

<b>LCS</b>	Local Control Station.....	421, 431
<b>LDPB</b>	LAr Digital Processing Blade.....	105, 295
<b>LDPS</b>	LAr Digital Processing System.....	91, 168–170
<b>LHC</b>	Large Hadron Collider ..	5–7, 20, 40, 59, 64, 71, 95, 96, 240, 270, 307, 331, 405, 421, 449, 483, 491
<b>LHCC</b>	LHC Experiments Committee.....	459, 460, 478, 482, 514
<b>LLP</b>	Long-Lived Particle.....	196
<b>lpGBT</b>	Low Power Gigabit Bidirectional Trigger and Data Link	76, 83, 103, 104, 213, 292–294, 422, 424
<b>LS3</b>	Long Shutdown 3 .....	468, 491–493
<b>LTI</b>	Local Trigger Interface.....	81, 89, 95–97, 99, 100, 102, 265, 269, 270, 272, 277, 279–281, 283–286, 291, 293, 405, 408, 418, 472, 483, 488
<b>LUCID</b>	LUMinosity Cherenkov In-tegrating Detector.....	105, 295
<b>LUT</b>	Lookup Table. An array that replaces runtime computation with a simpler array indexing operation.....	213, 218, 226, 248, 250–252, 275
<b>MC</b>	Monte-Carlo.....	188, 191–193, 208, 210, 211, 217, 218, 221–224
<b>MDT</b>	Monitored Drift Tube .....	2, 11, 12, 15, 17, 32, 55, 77, 78, 80, 81, 87, 91–94, 98–100, 102, 105, 117, 128, 129, 131, 177–180, 182, 183, 190, 194, 196–203, 207, 208, 210–216, 219–223, 225–231, 295, 401, 430, 436, 442, 447, 448, 468, 483, 484, 488, 508, 524–526
<b>MEV</b>	Maximum Expected Value: the CBE value of a parameter when contingency is added.....	227, 551
<b>MGT</b>	Multi-Gigabit Transceiver.....	98, 99, 258, 259, 275, 538
<b>MIP</b>	Minimum Ionising Particle.....	77
<b>MLM</b>	Multi-Layer Masks .....	378
<b>MM</b>	Micromega .....	54, 92, 93, 177, 178, 181, 202–207, 230
<b>MoU</b>	Memorandum of Understanding .....	440, 514
<b>MoU</b>	Memorandum of Understanding .....	457, 460, 464, 477, 478, 513, 524
<b>MPV</b>	Maximum Possible Value: the MEV value of a parameter when margins are added.....	227, 259, 262, 263, 446
<b>MUCTPI</b>	Muon CTP Interface	15, 17, 40, 42, 87–89, 91, 93–96, 99, 100, 102, 105, 180, 190, 197–200, 202, 203, 225, 228, 229, 236, 237, 240, 242, 256, 257, 269–275, 281–286, 293, 295, 405, 430, 436, 442, 446, 469, 472, 483, 484, 487, 488, 507, 518, 524–526

---

<b>MUX</b>	Multiplexer Processor . . . . .	95, 99, 100, 102, 235–238, 240, 241, 246, 255, 256, 258–264, 402, 403, 418, 419, 440, 442, 443, 487, 524–526, 538
<b>NCP</b>	ATLAS National Contact Physicist . . . . .	459, 460
<b>NIC</b>	Network Interface Card . . . . .	421
<b>NRE</b>	Non-Recursive Engineering refers to the one-time cost to research, design, develop and test a new product or product enhancement. . . . .	479, 491
<b>NSW</b>	New Small Wheel . . . . .	2, 15–17, 32, 40, 42, 43, 54, 55, 60, 77, 78, 82, 87, 91–94, 98–100, 102, 103, 105, 128, 129, 177–180, 182, 191–195, 198, 200, 202–207, 211, 216, 227–231, 295, 397, 399, 401, 407, 430, 436, 447, 448, 468, 484, 488, 508, 524–526
<b>NVMe</b>	Non-Volatile Memory Express . . . . .	300
<b>OLT</b>	Optical Line Terminal . . . . .	96
<b>ONU</b>	Optical Network Unit . . . . .	96, 104
<b>OPC UA</b>	Open Platform Communications Unified Architecture . . . . .	422, 424–431
<b>PAR</b>	Production Advancement Review . . . . .	498
<b>PBS</b>	Product Breakdown Structure . . . . .	13, 468, 469, 482–488, 490, 492–495, 504, 505, 513, 526
<b>PC</b>	Personal Computer . . . . .	2, 15, 104, 114, 118
<b>PCB</b>	Printed Circuit Board . . . . .	388, 479, 500, 501
<b>PCIe</b>	Peripheral Component Interconnect Express . . . . .	65, 98, 104, 118, 293, 294, 297, 298, 301, 335, 425, 481
<b>PDF</b>	Parton Distribution Function . . . . .	31, 36
<b>PDR</b>	Preliminary Design Review . . . . .	230, 440, 500, 501, 503, 511
<b>PFM</b>	Production Firmware Deployment Module . . . . .	260, 261, 440, 443, 484, 488, 509, 524–526
<b>PLL</b>	Phase-Locked Loop . . . . .	383
<b>PO</b>	ATLAS Technical Coordination Project Office . . . . .	459
<b>PON</b>	Passive Optical Network. PON is a telecommunication’s network technology that uses point-to-multipoint bi-directional connections. A PON consists of a central office node, called an optical line terminal (OLT), one or more user nodes, called optical network units (ONUs) . . . . .	96–99, 103, 104, 270, 272, 279–281, 283, 291, 293, 405, 483
<b>PPES</b>	Physics, Performance and Event Selection group . . . . .	473



---

<b>PRM</b>	Pattern Recognition Mezzanine . . . . 345, 346, 349, 352, 353, 362–365, 367, 370, 372–376, 378, 382–385, 387, 388, 390, 412, 414–416, 486, 490, 491, 510, 528–531
<b>PRR</b>	Production Readiness Review . . . . . 433, 440, 441, 500, 502, 503
<b>QGP</b>	Quark-Gluon Plasma . . . . . 35, 36
<b>quasar</b>	Quick OPC UA Server Generation Framework . . . . . 428–431
<b>R3</b>	Regional Readout Request; Level-0 regions of interest for <a href="#">L1Track</a> . . 399, 402, 403, 408, 418
<b>RAM</b>	Random Access Memory . . . . . 258, 264, 265, 301, 314, 370, 373, 374, 402, 416
<b>RB</b>	CERN Research Board. The Research Board receives the recommendations from all the CERN Experiment Committees, and takes decisions on them. Once approved, the proposals become part of the CERN experimental programme. The Research Board also decides on the accelerator schedules and requests for “Recognised Experiments” at CERN . . . . . 461, 467, 477
<b>RC</b>	ATLAS Resource Coordinator . . . . . 459, 467
<b>Readout</b>	Readout subsystem . . . . . 2, 14–17, 89, 92, 93, 436
<b>rHTT</b>	Regional hardware-based tracking for the trigger that is part of the Event Filter System . . . . . 2, 15, 17, 20, 110, 113, 114, 117–119, 123, 127, 130, 134, 139, 142, 151, 153, 299, 330, 331, 333, 337–340, 343, 344, 347, 356–358, 360, 363, 366, 387, 388, 404, 409–411, 418, 419
<b>ROD</b>	ReadOut Driver . . . . . 42, 65, 104, 174, 296, 437, 549
<b>RoI</b>	Region of Interest . . 42, 71, 107, 113, 114, 213–215, 218–220, 223, 226, 228, 241, 242, 253, 292, 329, 333, 334, 338, 347, 360, 361, 398, 399, 401, 418
<b>RoIE</b>	Region of Interest Engine . . . . . 399–406, 408, 418, 419
<b>ROL</b>	ReadOut Link . . . . . 43, 65, 66
<b>ROS</b>	ReadOut System . . . . . 42, 43, 65
<b>RPC</b>	Resistive Plate Chamber . . . . . 2, 11, 12, 15, 53–55, 64, 77, 78, 80, 81, 87, 91–93, 99, 100, 102, 105, 128–131, 177–192, 196–199, 201, 202, 207, 208, 210, 211, 213, 215, 227–231, 295, 401, 436, 448, 468, 483, 484, 488, 507, 508
<b>RPV</b>	R-parity violating . . . . . 10, 154
<b>RRB</b>	LHC Resources Review Board. The Resources Review Board comprises the representatives of each Experiment’s Funding Agencies and the managements of CERN and of each Experiment’s Collaboration. It is chaired by the CERN Director for Research and Computing. . . . . 460, 478, 513, 524
<b>RTM</b>	Rear Transition Module in the <a href="#">HTT</a> . . . . . 83, 367, 369, 386, 486, 488, 490, 510, 528–531

---

<b>SCA</b>	Switched Capacitor Array .....	64
<b>SCADA</b>	Supervisory Control And Data Acquisition.....	421, 428
<b>SCS</b>	Local Control Station.....	421, 422
<b>SCT</b>	Semiconductor Tracker .....	44
<b>SDX1</b>	Surface counting room .....	16, 434–436, 450–453, 471, 479, 480, 485
<b>SL</b>	Sector Logic .....	99, 100, 102, 442, 447, 483, 484, 488, 506, 507, 524–526
<b>SM</b>	Standard Model .....	196
<b>sMDT</b>	Small-diameter Monitored Drift Tube.....	55, 177
<b>SMP</b>	Schedule Management Plan .....	497
<b>SNMP</b>	Simple Network Management Protocol.....	316, 427, 431
<b>SoC</b>	System-on-Chip .....	83, 219, 220, 426, 427, 431, 556
<b>SR</b>	Specification Validation Review.....	500, 501, 503
<b>SSD</b>	Solid State Disk Drive .....	108, 300, 301
<b>SSID</b>	SuperStrip Identifier .....	349, 370, 373, 375, 384, 413
<b>SSTP</b>	Second-Stage Tracking Processor ...	118, 119, 345–348, 353, 362, 364–367, 369, 383, 385–390, 443
<b>sTGC</b>	small-strip <b>TGC</b> .....	54, 92, 93, 177, 178, 181, 202–207, 230
<b>Storage Handler</b>	Storage Handler .....	85, 331, 333
<b>super cell</b>	LAr calorimeter region formed by summing $E_T$ from cells that are adjacent in $\eta$ and $\phi$ .....	45–47, 51, 52, 66, 77, 90, 164, 168, 171, 172
<b>SUSY</b>	Supersymmetry.....	10, 19, 22, 24, 25, 27, 30, 31, 50, 71, 149, 154, 416, 417
<b>TAP</b>	Trigger items after prescale in the <b>CTP</b> .....	277, 278
<b>TAV</b>	Trigger items after veto in the <b>CTP</b> .....	277, 278, 404, 405
<b>TBP</b>	Trigger items before prescale in the <b>CTP</b> .....	276, 278
<b>TC</b>	ATLAS Technical Coordinator.....	459, 470
<b>TDAQ</b>	ATLAS Trigger and Data Acquisition System .....	1, 2, 5, 9–14, 16, 17, 19, 37, 39–44, 56, 59, 62, 66, 68, 75–80, 82, 85, 87, 88, 90, 92–98, 101–103, 115, 123, 163, 164, 169, 279, 282, 290, 292, 297, 312, 318, 320, 323, 331, 332, 343, 362, 364, 374, 393, 396, 398–400, 407, 409, 411, 414, 415, 418, 433–438, 440, 441, 443, 444, 446, 449, 450, 453, 457, 461–472, 474, 477–479, 482, 485–487, 489–492, 497, 499, 500, 503, 504, 513, 524, 528, 533–535, 538, 547, 554, 555
<b>TDC</b>	Time-to-Digital Converter .....	78, 213, 215, 225
<b>TDIB</b>	<b>TDAQ</b> Institutional Board .....	462–464, 467, 468, 470, 478, 498, 513, 514

---

<b>TDMT</b>	TDAQ Management Team . . . . .	463, 464, 468
<b>TDR</b>	Technical Design Report . . . . .	9–12, 16, 17, 59, 75, 98, 177, 179, 227, 364, 390, 418, 463, 466, 467, 474, 480, 497, 524, 533, 538
<b>TDSG</b>	TDAQ Steering Group . . . . .	463–465, 467, 468
<b>TFM</b>	Track-Fitting Mezzanine . . . . .	346, 347, 353, 362–365, 367, 383–388, 390, 443, 486, 490, 491, 510, 528–531, 539
<b>TGC</b>	Thin Gap Chamber . . . . .	2, 12, 15, 53–56, 77, 78, 80, 81, 87, 91–93, 100, 102, 105, 128, 129, 177–180, 182, 190–200, 202–204, 207, 208, 210, 211, 215, 227–231, 295, 401, 436, 448, 468, 483, 484, 488, 507, 508, 554
<b>TIC</b>	Trigger pattern of the trigger Inputs Contributing to the formation of the L0A . . . . .	278, 401, 403
<b>Tile</b>	Tile Calorimeter . . . . .	77, 87, 90–93, 102, 105, 163, 168, 177, 179, 182, 184, 191, 192, 194, 195, 198, 200, 227–230, 240, 262, 295, 449
<b>TIP</b>	Trigger Input signals as they enter the CTP before the trigger logic . . . . .	95, 100, 237, 239–241, 262–265, 278, 402
<b>TLA</b>	Trigger-object Level Analysis . . . . .	26, 29
<b>TMF</b>	TDAQ Maintenance Facility . . . . .	438–440, 442, 449, 504
<b>TOB</b>	Trigger Object . . . . .	91, 95, 163, 173–175, 236, 237, 240–242, 402, 403
<b>ToR</b>	Top of Rack network switch . . . . .	485
<b>TP</b>	Tracking Processor main board in the HTT . . . . .	118, 119, 345, 347, 348, 354, 362, 363, 367–369, 371, 374, 490, 491, 510
<b>TPPr</b>	Tile PreProcessor System . . . . .	77, 94, 169, 257
<b>TRT</b>	Transition Radiation Tracker . . . . .	65
<b>TT</b>	Calorimeter Trigger Tower . . . . .	169, 171
<b>TTC</b>	Trigger, Timing, and Control system . . . . .	15, 17, 40, 62, 81, 89, 95–97, 101, 103, 104, 197, 199, 205, 206, 213, 241, 257, 265, 269, 270, 272, 278–281, 283, 285, 286, 290–293, 403, 405, 436, 449, 469, 472, 482, 483, 507, 519, 524–526
<b>UAB</b>	ATLAS Upgrade Advisory Board . . . . .	459, 460
<b>UC</b>	ATLAS Upgrade Coordinator . . . . .	459–461, 467, 498
<b>UCG</b>	CERN Upgrade Cost Group. The UCG reviews the CORE cost of Technical Design Reports (TDR). The UCG reports to the Research Board via the LHCC Chairman . . . . .	459, 477, 480, 482, 487, 488, 490, 497, 507, 513, 514, 533, 535
<b>UE</b>	Underlying Event . . . . .	36

---

<b>UPL</b>	Upgrade Project Leader . . . <a href="#">459</a> , <a href="#">461</a> , <a href="#">464</a> , <a href="#">465</a> , <a href="#">467</a> , <a href="#">468</a> , <a href="#">470</a> , <a href="#">477</a> , <a href="#">478</a> , <a href="#">498</a> , <a href="#">503</a> , <a href="#">513</a> , <a href="#">514</a> , <a href="#">534</a>
<b>UPMT</b>	Upgrade Project Management Team . . . . . <a href="#">465</a> , <a href="#">467</a> , <a href="#">468</a> , <a href="#">470</a>
<b>UPO</b>	ATLAS Upgrade Project Office . . . . . <a href="#">470</a> , <a href="#">480</a> , <a href="#">498</a>
<b>UPR</b>	Upgrade Project . <a href="#">457</a> , <a href="#">459–461</a> , <a href="#">464–474</a> , <a href="#">477–480</a> , <a href="#">482</a> , <a href="#">486</a> , <a href="#">487</a> , <a href="#">489</a> , <a href="#">490</a> , <a href="#">497</a> , <a href="#">498</a> , <a href="#">500</a> , <a href="#">503–505</a> , <a href="#">513–517</a> , <a href="#">524</a> , <a href="#">528</a> , <a href="#">533–535</a>
<b>UPR-RC</b>	Upgrade Project Resource Coordinator and Risk Manager <a href="#">477</a> , <a href="#">478</a> , <a href="#">498</a> , <a href="#">514</a> , <a href="#">534</a> , <a href="#">535</a>
<b>UPR-TC</b>	Upgrade Project Technical Coordinator . . . . . <a href="#">470</a>
<b>UPS</b>	Uninterruptible Power Supply . . . . . <a href="#">450</a>
<b>URD</b>	User’s Requirement Document . . . . . <a href="#">500</a> , <a href="#">501</a> , <a href="#">503</a>
<b>USA15</b>	Main underground electronics cavern . . <a href="#">16</a> , <a href="#">78</a> , <a href="#">92</a> , <a href="#">98–100</a> , <a href="#">279</a> , <a href="#">286</a> , <a href="#">433–446</a> , <a href="#">449</a> , <a href="#">450</a> , <a href="#">471</a> , <a href="#">479</a> , <a href="#">480</a> , <a href="#">485</a> , <a href="#">504</a> , <a href="#">539</a>
<b>USC</b>	ATLAS Upgrade Steering Committee . . . . . <a href="#">459–461</a> , <a href="#">464</a> , <a href="#">467</a> , <a href="#">498</a>
<b>UX15</b>	ATLAS underground experimental cavern . . . . . <a href="#">435</a> , <a href="#">444</a> , <a href="#">450</a>
<b>VBF</b>	Vector Boson Fusion . . . . . <a href="#">10</a> , <a href="#">24</a> , <a href="#">25</a> , <a href="#">27</a> , <a href="#">28</a> , <a href="#">33</a> , <a href="#">34</a> , <a href="#">144</a> , <a href="#">148–152</a> , <a href="#">154</a> , <a href="#">416</a> , <a href="#">417</a>
<b>VLAN</b>	Virtual <a href="#">LAN</a> . . . . . <a href="#">430</a>
<b>VME</b>	Versa Module Europa; a computer bus standard . . . . . <a href="#">xix</a> , <a href="#">69</a> , <a href="#">311</a> , <a href="#">425–427</a>
<b>WBS</b>	Work Breakdown Structure . . . . . <a href="#">468</a> , <a href="#">469</a> , <a href="#">471</a> , <a href="#">482</a> , <a href="#">497–499</a> , <a href="#">504</a> , <a href="#">513</a> , <a href="#">514</a> , <a href="#">516</a> , <a href="#">518–523</a> , <a href="#">534</a>
<b>WinCC OA</b>	Simatic WinCC Open Architecture . . . . . <a href="#">421</a> , <a href="#">426–429</a>
<b>WTA</b>	Winner-Take-All . . . . . <a href="#">250–252</a>
<b>ZDC</b>	Zero Degree Calorimeters . . . . . <a href="#">105</a> , <a href="#">295</a>
<b>Zynq®</b>	A Xilinx <a href="#">SoC</a> composed of an FPGA and an ARM processor. . . . . <a href="#">197</a> , <a href="#">199</a> , <a href="#">205</a> , <a href="#">213</a> , <a href="#">226</a> , <a href="#">231</a>

# The ATLAS Collaboration

## The ATLAS Collaboration

### **Argentina**

#### *Buenos Aires*

M.A. Bonaventura, J.D. Bossio Sola, R.D. Castro, M.F. Daneri, M.R. Devesa,  
D.J. Foguelman, A. Laurito, G. Marceca, G. Otero y Garzon, R. Piegaia

#### *La Plata*

M.J. Alconada Verzini, F. Alonso, F.A. Arduh, M.T. Dova, J. Hoya, F. Monticelli,  
H. Wahlberg

### **Armenia**

#### *Yerevan*

T. Mkrtchyan

### **Australia**

#### *Adelaide*

D. Duvnjak, P. Jackson, J.L. Oliver, A. Petridis, A. Qureshi, A.S. Sharma, M.J. White

#### *Melbourne*

E.L. Barberio, A.J. Brennan, E. Dawe, S. Goldfarb, T. Kubota, B. Le, L.H. Mason,  
E.F. McDonald, P.C. McNamara, M. Milesi, F. Nuti, T. Pham, P. Rados, F. Scutti, S. Shojaii,  
L.A. Spiller, G.N. Taylor, P.T.E. Taylor, F.C. Ungaro, P. Urquijo

#### *Sydney*

A. Limosani, C.J.E. Suster, K.E. Varvell, J. Wang, B. Yabsley

### **Austria**

#### *Innsbruck*

S. Jakobsen, E. Kneringer, W. Lukas, A. Manousos

### **Azerbaijan**

#### *Baku*

F. Ahmadov, N. Huseynov, N. Javadov, F. Khalil-zada

**Brazil**

*Juiz de Fora UF*

A.S. Cerqueira, R. Goncalves Gama, L. Manhaes de Andrade Filho, B.S. Peralva,  
M.V. Silva Oliveira

*Rio de Janeiro UF*

Y. Amaral Coutinho, V. Araujo Ferraz, R. Araujo Pereira, M. Begalli, L.P. Caloba,  
R. Coura Torres, W.S. Freund, C. Maidantchik, F. Marroquim, J.M. Seixas

*Sao Joao del Rei UF*

M.A.B. do Vale

*Sao Paulo*

M. Donadelli, J.L. La Rosa Navarro, M.A.L. Leite

**Canada**

*Alberta*

N. Dehghanian, D.M. Gingrich, S. Jabbar, R.W. Moore, J.L. Pinfold, X. Sun, H. Wang

*Carleton*

A. Bellerive, C.C. Chau, G. Cree, D. Di Valentino, D. Gillberg, J. Heilman, R.F.H. Hunter,  
A.M. Hupe, J.S. Keller, T. Koffas, I. Nomidis, F.G. Oakham, A. Ruiz-Martinez, N. Sherafati,  
M.G. Vincter, S.A. Weber

*McGill*

A.J. Chuinard, F. Corriveau, R.A. Keyes, B. Lefebvre, R. Mantifel, S. Prince, S.H. Robertson,  
A. Robichaud-Veronneau, H.L. Russell, B. Vachon, T. Vazquez Schroeder, A. Warburton

*Montreal*

J-F. Arguin, G. Azuelos, T.R.V. Billoud, F. Dallaire, O.A. Ducu, B. Freund, L.G. Gagnon,  
C. Leroy, K. Mochizuki, T. Nguyen Manh, D. Shoaleh Saadi

*Simon Fraser Burnaby*

H. Bahrasemani, E. Dreyer, A.J. Horton, A. Montalbano, D. Mori, D.C. O'Neil, K. Pachal,  
B. Stelzer, D. Temple, J. Van Nieuwkoop, M.C. Vetterli

**TRIUMF**

G. Azuelos, S.V. Chekulaev, D.M. Gingrich, F. Guescini, N.P. Hessey, N. Hod, J. Jovicevic,  
L.L. Kurchaninov, F.G. Oakham, P. Savard, O. Stelzer-Chilton, R. Tafirout, I.M. Trigger,  
M.C. Vetterli

*Toronto*

S.H. Abidi, O. Arnaez, S.J. Batista, A.F. Casha, K.J.R. Cormier, D.A. DeMarco, K. Dette, M. Diamond, R. Di Sipio, P. Krieger, R. Les, A. Liblong, G. Mc Goldrick, A. Milic, R.S. Orr, V.R. Pascuzzi, P. Savard, P. Sinervo, R.J. Teuscher, W. Trischuk, L.M. Veloce

*Vancouver UBC*

F. Cormier, M. Danninger, W. Fedorko, C. Gay, A. Held, S. Henkelmann, A. Lister, S. Rettie, V.W.S. Wong

*Victoria*

J. Albert, Y.H. Chiu, A.A. Elliot, K. Hamano, E. Hill, R. Keeler, R. Kowalewski, E.S. Kuwertz, T. Kwan, M. Lefebvre, K.D. McLean, R.A. McPherson, R. Sobie, M. Trovatelli, M. Venturi

*York*

M.J. Kareem, A.M. Rodríguez Vera, W. Taylor

**CERN**

*European Organization for Nuclear Research, Geneva, Switzerland*

M. Aleksa, B. Alvarez Gonzalez, S. Amoroso, L. Aperio Bella, A.J. Armbruster, G. Avolio, M-S Barisits, T.A. Beermann, J. Bortfeldt, J. Boyd, Q. Buat, W. Buttinger, P. Butti, S. Camarda, S. Campana, M.D.M. Capeans Garrido, T. Carli, G.D. Carrillo-Montoya, A. Catinaccio, A. Cattai, M.A. Chelstowska, A.S. Chisholm, A. Cortes-Gonzalez, P. Czodrowski, V. Dao, A. Dell'Acqua, P.O. Deviveiros, A. Di Girolamo, R. Di Nardo, F. Dittus, A. Dudarev, M. Dührssen, T. Eifert, N. Ellis, M. Elsing, P. Farthouat, P. Fassnacht, D. Fazio, E.J. Feng, D. Francis, S.M. Fressard-Batraneanu, D. Froidevaux, L. Georgopoulos, L. Goossens, B. Gorini, J. Guenther, S. Guindon, C. Gumpert, S. Haas, R.J. Hawkings, L. Helary, C. Hensens, A.M. Henriques Correia, L. Hervas, A. Hoecker, J. Hrdinka, M. Huhtinen, P. Iengo, P. Jenni, G.E. Jerezcek, M. Joos, T. Klioutchnikova, B. Kolobara, A. Krasznahorkay, S. Kuehn, M.K. Kvist, R.J. Langenberg, M. Lassnig, F. Le Goff, G. Lehmann Miotto, B. Lenzi, S. Malyukov, L. Mapelli, C.A. Marin Tobon, B. Martin, A. Marzin, J.A. Mcfayden, J. Montejo Berlingen, M. Moreno Llácer, S. Morgenstern, A.K. Morley, G. Mornacchi, A.M. Nairz, M. Nessi, M. Nordberg, S. Palestini, P. Pani, T. Pauly, H. Pernegger, M. Pernigotti, B.A. Petersen, R. Polifka, J. Poveda, M.E. Pozo Astigarraga, M. Raymond, C. Rembser, B. Ristić, E. Ritsch, S. Roe, N. Ruthmann, V. Ryjov, A. Salzburger, S. Schaepe, E.J. Schioppa, S. Schlenker, K. Schmieden, J. Schovancova, J. Schumacher, C.A. Solans Sanchez, G. Spigo, R. Spiwoks, H.J. Stelzer, G.A. Stewart, S. Stärz, H. Ten Kate, L. Tremblet, G. Unal, W. Vandelli, N. Venturi, P. Vichoudis, R. Vuillermet, P.S. Wells, T. Wengler, S. Wenig, H.G. Wilkens, O.J. Wyszynski, C.J.S. Young, D. Zanzi, L. Zwalinski

**Chile**

*Santiago*

Dr. Blunier, M.A. Diaz, F.M. Garay Walls, G.R. Lee, J.P. Ochoa-Ricoux, S.A. Olivares Pino, J.C. Rivera Vergara

*Valparaiso*

W.K. Brooks, E. Carquin, S. Kuleshov, J.A. Lopez, R. Pezoa, F. Prokoshin, J.E. Salazar Loyola, S. Tapia Araya, G.A. Vasquez, N. Viaux Maira, R. White

**China**

*Beijing IHEP*

M.K. Ayoub, J. Barreiro Guimarães da Costa, C. Bertella, Y. Fang, J.A. García Pascual, Y. Huang, Z. Liang, J. Llorente Merino, X. Lou, F. Lyu, J.D. Mansour, Q. Ouyang, L.Y. Shan, D. Xu, X. Zhuang, H. Zhu

*Beijing Tsinghua*

X. Chen, L. Xia

*Beijing UCAS*

H.J. Cheng, S. Han, M.G. Kurth, Q. Li, C. Peng, H. Ren, Y. Zhang, M. Zhou

*Hefei*

R. Avramidou, Z. Barnovska-Blenessy, C. Chen, J. Chen, Y. Guo, K. Han, L. Han, Y. Jiang, B. Li, C.-Q. Li, Q. Li, J.B. Liu, M. Liu, Y. Liu, Y.L. Liu, R. Ospanov, H. Peng, R. Wang, W. Wang, Y. Wang, Y. Wu, H. Xu, G. Zhang, L. Zhang, R. Zhang, Z. Zhao, H. Zhu, Y. Zhu

*Hong Kong CUHK*

Y.L. Chan, M.C. Chu, L.R. Flores Castillo, J.M. Iturbe Ponce, T.S. Lau, H. Lu, A. Salvucci, K.W. Tsang

*Hong Kong HKU*

C.Y. Lo, N. Orlando, D. Paredes Hernandez, A. Salvucci, Y. Tu

*Hong Kong HKUST*

K. Lie, T.Y. Ng, K. Prokofiev, A. Salvucci

*Nanjing*

S. Chen, S. Jin, W. Wang, H. Zhang, P. Zhang

*Shandong*

M.J. Da Cunha Sargedas De Sousa, Y. Du, C. Feng, H. Li, L.L. Ma, Y. Ma, C. Wang, D. Zhang, X. Zhang, Y. Zhao, C.G. Zhu

*Shanghai*



S.L. Barnes, M. Cano Bret, J. Guo, S. Hu, N. Kondrashova, L. Li, S. Li, X. Li, N. Nishu,  
B. Parida, Z. Wang, H. Yang, N. Zhou

*TDLI*

S. Li, H. Yang

### **Colombia**

*Bogota UAN*

M. Losada, D. Moreno, G. Navarro, C. Sandoval

### **Czech Republic**

*Olomouc*

L. Chytka, P. Hamal, M. Hrabovsky, J. Kvita, L. Nozka

*Prague AS*

J. Chudoba, J. Hejbal, O. Hladik, P. Jacka, T. Jakoubek, O. Kepka, J. Kroll, A. Kupco,  
M. Lokajicek, R. Lysak, M. Marcisovsky, M. Mikestikova, S. Nemecek, O. Penc, P. Sicho,  
P. Staroba, M. Svatos, M. Tasevsky

*Prague CTU*

B. Ali, K. Augsten, D. Caforio, P. Gallus, M. Havranek, Z. Hubacek, M. Myska, R. Novotny,  
S. Pospisil, T. Slavicek, K. Smolek, M. Solar, A. Sopczak, M. Suk, V. Vacek, P. Vokac, V. Vrba

*Prague CU*

I. Carli, T. Davidek, J. Dolejsi, Z. Dolezal, J. Faltova, P. Kodys, T. Kosek, R. Leitner,  
M. Mlynarikova, V. Pleskot, P. Reznicek, D. Scheirich, R. Slovak, M. Spousta, T. Sykora,  
P. Tas, V. Vorobel

### **Denmark**

*Copenhagen NBI*

A. Alonso, M. Bajic, H. Bertelsen, G.J. Besjes, M. Dam, F.A. Dias, G. Galster, J.B. Hansen,  
J.D. Hansen, P.H. Hansen, R. Ignazzi, J. Monk, T.C. Petersen, S.H Stark, F. Thiele,  
C. Wiglesworth, S. Xella

### **France**

*Annecy LAPP*

N. Berger, A.M. Burger, O. Dartsis, M. Delmastro, L. Di Ciaccio, P.J. Falke, S. Falke, C. Goy,  
T. Guillemin, T. Hryn'ova, S. Jézéquel, O. Kivernyk, I. Koletsou, R. Lafaye, J. Levêque,  
N. Lorenzo Martinez, P. Mastrandrea, S. Raspopov, E. Sauvan, B.H. Smart,  
S. Todorova-Nova, A. Vallier, I. Wingerter-Seez, E. Yatsenko

*Clermont-Ferrand*

S. Angelidakis, W.M. Barbe, D. Boumediene, D. Calvet, S. Calvet, J. Donini, Ph. Gris, F.A. Jimenez Morales, R. Madar, M. Marjanovic, E. Nibigira, D. Pallin, C. Santoni, S. Senkin, F. Vazeille

*Grenoble LPSC*

S. Berlendis, C. Camincher, J. Collot, S. Crépe-Renaudin, P.A. Delsart, M.H. Genest, M. Kuna, F. Ledroit-Guillon, A. Lleres, A. Lucotte, F. Malek, T. Meideck, E. Petit, N.P. Readioff, J. Stark, B. Trocmé

*Lyon CC-IN2P3*

G. Rahal

*Marseille CPPM*

G. Aad, M.I. Alstaty, M. Barbero, N. Brahimi, A. Calandri, Y. Coadou, K. De Vasconcelos Corga, C. Diaconu, F. Djama, A.E. Dumitriu, A. Duperrin, R. El Kosseifi, V. Ellajosyula, L. Feligioni, Z. Guo, A. Hadeif, G.D. Hallewell, D. Hoffmann, F. Hubaut, E.B.F.G. Knoops, R. Kukla, E. Le Guirriec, C. Meessen, E. Monnier, S. Muanza, E. Nagy, H.D.N. Nguyen, P. Pralavorio, Y. Rodina, A. Rozanov, M. Talby, G. Tarna, S. Tisserant, J. Toth, F. Touchard, L. Vacavant, C. Wang, R. Wolff, R. Zhang

*Orsay LAL*

B. Abeloos, C. Allaire, A. Bassalat, C. Bourdarios, A.R. Chomont, D. Delgove, C. Delporte, J.B. De Vivie De Regie, L. Dufлот, M. Escalier, L. Fayard, D. Fournier, C.R. Goudet, J.-F. Grivaz, A. Guerguichon, K. Han, D. Hohov, J. Hrivnac, L. Iconomidou-Fayard, M. Kado, A. Laudrain, A. Lounis, N. Makovec, N. Morange, P. Petroff, L. Poggioli, P. Puzo, T. Rashid, D. Rousseau, G. Rybkin, S. Sacerdoti, A.C. Schaffer, L. Serin, S. Simion, R. Tanaka, D. Varouchas, D. Zerwas, Z. Zhang, Y. Zhao

*Paris LPNHE*

T. Beau, G. Bernardi, M. Bomben, G. Calderini, F. Crescioli, L. D'Eramo, S. De Cecco, F. Derue, A. Ducourthial, M. Garci, R. Hankache, M.W. Krasny, D. Lacour, B. Laforge, S. Laplace, O. Le Dortz, kliu Liu, A. Lopez Solis, I. Luise, P.M. Luzzi, B. Malaescu, G. Marchiori, I. Nikolic-Audit, J. Ocariz, D.M. Portillo Quintero, M. Ridet, L. Roos, A. Tarek Abouelfadl Mohamed, S. Trincaz-Duvoid, R.-J. Wang

*Saclay CEA*

H. Bachacou, F. Balli, F. Bauer, N. Besson, M. Boonekamp, L. Chevalier, F. Deliot, A.I. Etienvre, A. Formica, P.F. Giraud, C. Guyot, S. Hassani, F. Jeanneau, W. Kozanecki, J.F. Laporte, E.P. Le Quilleuc, A.A.J. Lesage, B. Mansoulie, J.-P. Meyer, T.J. Neep, R. Nicolaidou, A. Ouraou, L. Pacheco Rodriguez, M.M. Perego, A. Peyaud, L. Schoeffel, Ph. Schune, Ph. Schwemling, T. Xu

**Georgia**

*Tbilisi IP*

J. Jejelava, E.G. Tskhadadze

*Tbilisi SU*

T. Djobava, A. Durglishvili, J. Khubua, I.A. Minashvili, M. Mosidze

**Germany***Berlin HU*

D. Biedermann, J. Dietrich, S. Grancagnolo, G.H. Herbert, I. Hristova, O.M. Kind, H. Lacker, T. Lohse, S. Mergelmeyer, Y.S. Ng, F. Peri, L. Rehnisch, F. Schenck, D. Sperlich, S. Stamm, M. zur Nedden

*Bonn*

O. Arslan, A. Bandyopadhyay, P. Bechtle, F.U. Bernlochner, A. Betti, I. Brock, J. Caudron, I.A. Cioarã, M. Cristinziani, W. Davey, K. Desch, J. Dingfelder, G. Gaycken, M. Ghneimat, C.A. Gottardo, C. Greife, S. Hageböck, M.C. Hansen, S. Heer, D. Hohn, M. Huebner, F. Huegging, R.M. Jacobs, J. Janssen, T. Klingl, V.V. Kostyukhin, J. Kroseberg, H. Krüger, K. Lantzs, T. Lenz, A. Melzer, R. Moles-Valls, T. Obermann, D. Pohl, O. Ricken, L.K. Schildgen, E. Schopf, A. Sciandra, P. Seema, E. von Toerne, P. Wagner, N. Wermes, B.T. Winter, K.H. Yau Wong, S.P.Y. Yuen, R. Zhang

*DESY*

N. Asbah, J.K. Behr, D. Berge, C. Bertsche, M. Bessner, I. Bloch, F. Braren, K. Brendlinger, L. Brenner, L. Bryngemark, Y.-H. Chen, T. Daubney, C. David, C. Deterre, B. Dutta, M. Dyndal, S. Díez Cornell, C. Eckardt, J. Ferrando, N. Flaschel, K. Gasnikova, P.C.F. Glaysher, A. Glazov, I.M. Gregor, K. Grevtsov, S. Heim, B. Heinemann, K.H. Hiller, J. Jeong, J. Katzy, V. Kitali, T. Kuhl, J. Lacey, V.S. Lang, W.A. Leight, E.M. Lobodzinska, X. Lou, A. Madsen, K. Mönig, R.F. Naranjo Garcia, T. Naumann, A.A. O'Rourke, K.A. Parker, K. Peters, A. Poley, C.S. Pollard, K. Potamianos, M. Queitsch-Maitland, D.M. Rauch, O. Rifki, J.E.M. Robinson, E.M. Rüttinger, M. Saimpert, C.O. Sander, S. Schmitt, D. South, M.M. Stanitzki, M. Stegler, N.A. Styles, K. Tackmann, T. Theveneaux-Pelzer, A. Trofymov, F. Tsai, L. Valéry, A. Vishwakarma, C. Wanotayaroj, Y.C. Yap, N. Zakharchuk

*Dortmund*

I. Burmeister, D. Cinca, J. Erdmann, G. Gessner, C. Gössling, M. Homann, R. Klingenberg, K. Kroeninger, T. Kupfer, O. Nackenhorst, I. Nitsche

*Dresden*

C. Bittrich, D. Duschinger, L. Hauswald, M. Hils, P. Horn, F. Iltzsche, D. Kirchmeier, M. Kobel, W.F. Mader, N. Madysa, J. Manjarres Ramos, O. Novgorodova, F. Siegert, F. Socher, A. Straessner, S. Todt, H. Torres, S. WAhrmund

*Freiburg*

K. Becker, F. Buehrer, D. Büscher, F. Cardillo, E. Coniavitis, A. Di Simone, G. Gonella, R. Gugel, K.K. Heidegger, B. Heinemann, G. Hertzen, K. Jakobs, M. Javurkova, T. Javůrek, P. Jenni, S. Jiggins, T. Klapdor-Kleingrothaus, A. Knue, K. Köneke, U. Landgraf, A. Loesle, C. Luedtke, V. Magerl, T. Megy, P. Mogg, M. Nagel, U. Parzefall, K. Rosbach, Z. Rurikova, F. Rühr, D. Sammel, K.E. Schleicher, U. Schnoor, M. Schumacher, P. Tornambe, C. Weiser, L.A.M. Wiik-Fuchs, L. Zhang, S. Zimmermann

*Göttingen*

H.C. Beck, M. Bindi, T. Bisanz, P. Bokan, E. Buschmann, T. Dado, A. De Maria, E. Drechsler, T. Dreyer, J. Geisen, J. Grosse-Knetter, M. Janus, G. Kawamura, S. Lai, B. Lemmer, M. Mantoani, G. Mchedlidze, J.D. Mellenthin, C. Nellist, I. Pokharel, A. Quadt, N.-A. Rosien, G.F. Rzehorz, P. Sabatini, E. Shabalina, J.W. Smith, P. Stolte, R.E. Ticse Torres, J. Veatch, J. Weingarten, K. Zoch

*Giessen*

M. Düren, C. Heinz, K. Kreutzfeldt, H. Stenzel

*Heidelberg KIP*

R. Achenbach, V. Andrei, C. Antel, A.E. Baas, O. Brandt, J.I. Djuvsland, M. Dunford, S. Franchino, M.P. Geisler, P. Hanke, J. Jongmanns, A. Kugel, H. Meyer Zu Theenhausen, F. Napolitano, D.I. Narrias Villar, M. Sahinsoy, A. Santos, K. Schmitt, H.-C. Schultz-Coulon, T.M. Spieker, R. Stamen, P. Starovoitov, S. Suchek, S.M. Weber, M. Wessels, X. Yue

*Heidelberg PI*

C.F. Anders, A.E. Bolz, D. Britzger, D.E. Ferreira de Lima, M. Giulini, M. Kolb, M. Lisovyi, A. Schoening, D. Sosa, J. Wang

*Mainz*

S. Artz, B. Bauss, M. Becker, P. Berta, K. Bierwagen, B. Brickwedde, A.S. Brogna, V. Büscher, J. Cúth, J.F. Damp, A.Chr. Dudder, F. Fiedler, V. Filimonov, M. Geisen, S. Groh, H. Herr, K.B. Jakobi, C. Kahra, A. Kaluza, K. Kleinknecht, T.H. Lin, L. Masetti, S. Patariaia, S. Rave, A. Reiss, E. Rocco, J. Schaeffer, B. Schlag, C. Schmitt, S. Schmitz, M. Schott, N. Schuh, A. Schulte, U. Schäfer, E. Simioni, R. Simoniello, M. Simon, A. Sydorenko, D. Ta, S. Tapprogge, P. Urrejola, J. Vieira De Souza, S. Webb, M. Weirich, A. Wolf, E. Yildirim, M. Zinser

*Munich LMU*

M. Adersberger, M. Bender, O. Biebel, C. Bock, D. Bogavac, G. Duckeck, B.M. Flierl, D.M. Handl, N.M. Hartmann, J.J. Heinrich, R. Hertzenberger, F. Hoenig, M. Holzbock, F.F. Klitzner, F. Legger, J. Lorenz, P.J. Lösel, T. Maier, A. Mann, S. Mehlhase, J. Mitrevski, R.S.P. Mueller, F. Rauscher, B.M. Schachtner, D. Schaile, C. Valderanis, J. Wagner-Kuhr, R. Walker

*Munich MPI*

S. Abovyan, T. Barillari, S. Bethke, J. Beyer, D. Cieri, V. Danielyan, K.M. Ecker, M. Fras, P. Gadow, A. Hönle, J.J. Junggeburth, A.E. Kiryunin, S. Kluth, O. Kortner, S. Kortner, D. Krauss, H. Kroha, N.M. Köhler, A. La Rosa, A. Macchiolo, T.G. McCarthy, S. Menke, F. Mueller, R. Nisius, S. Nowak, H. Oberlack, B. Pearson, R. Richter, P. Rieck, R. Röhrig, D. Salihagic, N. Savic, P. Schacht, K.R. Schmidt-Sommerfeld, L.M. Scyboz, M. Spalla, F. Spettel, S. Stonjek, E.H. Takasugi, A. Wildauer, Z. Zinonos

*Siegen*

N.B. Atlay, P. Buchholz, A. Campoverde, I. Fleck, S. Ghasemi, I. Ibragimov, Y. Li, W. Walkowiak, M. Ziolkowski

*Wuppertal*

A.A.E. Bannoura, D. Boerner, G. Brandt, C. Dülsen, F. Ellinghaus, J. Fischer, T. Flick, G. Gilles, K. Hamacher, T. Harenberg, D. Hirschbuehl, G. Jäkel, S. Kersten, J.T. Kuechler, P. Mättig, M. Neumann, C.J. Riegel, M. Sandhoff, F. Tepel, M. Vogel, W. Wagner, M. Wensing, C. Zeitnitz

*Würzburg*

M. Haleem, V. Herget, F. Kuger, A. Redelbach, O. Sidiropoulou, G. Siragusa, R. Ströhmer, T. Trefzger, A. Zibell

**Greece**

*Athens NKUA*

D. Fassouliotis, I. Gkialas, C. Kourkoumelis, K. Papageorgiou, N. Tsirintanis

*Athens NTU*

T. Alexopoulos, N. Benekos, E.N. Gazis, P. Gkountoumis, E. Karentzos, A. Koulouris, S. Maltezos, P. Moschovakos, I. Panagoulas, S. Vlachos, G. Zacharis

*Thessaloniki*

E. Bochoridis, G. Bourlis, S. Chouridou, C. Gentsos, S. Gkaitatzis, D. Iliadis, K. Kordas, A. Leisos, A. Marantis, I. Maznas, K. Mitrogeorgos, S. Nikolaidis, M. Ntogramatzi, C. Petridou, D. Sampsonidis, D. Sampsonidou, P. Tsilias

**Israel**

*Technion Haifa*

H. Abreu, Y. Afik, R. Barnea, M. Bellomo, O. Gabizon, E. Gozani, B.J. Gutelman, E. Kajomovitz, N. Lupu, J. Meyer, E. Musto, Y. Rozen, S. Tarem

*Tel Aviv*

H. Abramowicz, G. Alexander, A. Ashkenazi, L. Barak, G. Bella, Y. Benhammou, T. Cao, J. Duarte-Campderros, E. Etzion, O. Gueta, O. Kuprash, Y. Oren, A. Soffer, J. Taenzer, N. Taiblum

*Weizmann Rehovot*

P. Balek, M. Birman, S. Bressler, Z.H. Citron, A.E.C. Coimbra, E. Duchovni, M. Dumancic, D.M. Front, S. Ganguly, I. Grayzman, E. Gross, A. Ivina, D. Lellouch, L.J. Levinson, G. Mikenberg, A. Milov, J. Narevicius, M. Pitt, I. Ravinovich, A. Roich, J. Shlomi, V. Smakhtin, D. Turgeman

**Italy**

*Bologna*

G.L. Alberghi, G. Balbi, L. Bellagamba, S. Biondi, D. Boscherini, F. Brasolin, A. Bruni, G. Bruni, M. Bruschi, G. Cabras, A. Cervelli, G. D'amen, S. De Castro, F. Fabbri, L. Fabbri, D. Falchieri, M. Franchini, A. Gabrielli, B. Giacobbe, N. Giangiacomi, F.M. Giorgi, F. Lasagni Manghi, A. Mengarelli, M. Negrini, A. Polini, L. Rinaldi, M. Romano, C. Sbarra, A. Sbrizzi, N. Semprini-Cesari, A. Sidoti, M. Sioli, R. Spighi, R. Travaglini, G. Ucchielli, S. Valentinetti, M. Villa, C. Vittori, A. Zoccoli

*Cosenza*

G. Callea, M. Capua, G. Crosetti, M. Del Gaudio, L. La Rotonda, F. La Ruffa, A. Mastroberardino, E. Meoni, S. Palazzo, A. Policicchio, D. Salvatore, M. Schioppa, M. Scornajenghi, E. Tassi

*Frascati*

P. Albicocco, M. Antonelli, M. Beretta, V. Chiarella, M. Gatta, G. Maccarrone, G. Mancini, A. Sansoni, M. Testa, E. Vilucchi

*Genova*

D. Barberis, A. Coccaro, G. Darbo, A. Favareto, G. Gagliardi, A. Gaudiello, C. Gemme, A. Lapertosa, P. Morettini, H. Oide, F. Parodi, S. Passaggio, L.P. Rossi, M. Sannino, C. Schiavi, C. Varni

*Lecce*

M. Aliev, K. Bachas, G. Chiodini, E. Gorini, L. Longo, A. Mirto, M. Primavera, M. Reale, S. Spagnolo, A. Ventura

*Milano*

G. Alimonti, A. Andreani, A. Andreatza, A. Camplani, L. Carminati, S. Carrá, M. Citterio, M. Fanti, L. Frontini, D. Giugni, T. Lari, M. Lazzaroni, V. Liberali, S. Manzoni, C. Meroni, S. Monzani, A. Murrone, L. Perini, F. Ragusa, M.G. Ratti, S. Resconi, L. Rossini, A. Stabile, G.F. Tartarelli, C. Troncon, R. Turra, M. Villaplana Perez

*Napoli*

A. Aloisio, M.G. Alviggi, V. Canale, G. Carlino, F. Ciotto, F. Conventi, R. de Asmundis, M. Della Pietra, C. Di Donato, A. Doria, A. Giannini, R. Giordano, V. Izzo, M. Lavorgna, L. Merola, S. Perrella, E. Rossi, G. Sekhniaidze

*Pavia*

A. Abed Abud, E.M. Farina, R. Ferrari, G. Gaudio, G. Introzzi, A. Kourkoumeli-Charalampidi, A. Lanza, M. Livan, A. Negri, R. Poggi, G. Polesello, D.M. Rebuzzi, A. Rimoldi, S. Sottocornola, V. Vercesi

*Pisa*

A. Annovi, R. Beccherle, N.V. Biesuz, M. Calvetti, V. Cavasinni, G. Chiarelli, M. Dell'Orso, A. De Maria, P. Francavilla, P. Giannetti, S. Leone, M. Piendibene, C. Roda, F. Scuri, C.L. Sotiropoulou

*Roma I*

F. Anulli, P. Bagnaia, M. Bauce, C. Bini, M. Corradi, D. De Pedis, A. De Salvo, S. Falciano, S. Gentile, S. Giagu, V. Ippolito, F. Lacava, C. Luci, L. Luminari, A. Messina, A. Nisati, E. Pasqualucci, E. Petrolo, L. Pontecorvo, M. Rescigno, S. Rosati, F. Safai Tehrani, C.D. Sebastiani, R. Vari, S. Veneziano

*Roma II*

G. Aielli, V. Bortolotto, S. Bruno, P. Camarri, R. Cardarelli, L. Cerrito, U. De Sanctis, A. Di Ciaccio, B. Liberti, S. Loffredo, L. Massa, M. Pinamonti, A. Salamon, R. Santonico, M. Vanadia

*Roma Tre*

A. Baroncelli, M. Biglietti, F. Ceradini, B. Di Micco, A. Farilla, M. Iodice, D. Orestano, F. Petrucci, G. Salamanna, M. Sessa, C. Stanescu, V. Vecchio, M. Verducci

*Trento*

R. Iuppa

*Trieste ICTP*

B.S. Acharya, L. Serkin, K. Shaw

*Udine*

M. Cobal, M.P. Giordani, G. Giugliarelli, N. Kimura, A. Sanchez Pineda, R. Soualah

**Japan**

*Hiroshima IT*

Y. Nagasaka

*KEK*

M. Aoki, K. Hanagaki, S. Higashino, T. Honda, Y. Ikegami, M. Ikeno, H. Iwasaki, J. Kanzaki, T. Kondo, T. Kono, Y. Makida, A. Mizukami, R. Nagai, K. Nagano, K. Nakamura, Y. Ninomiya, S. Odaka, T. Okuyama, O. Sasaki, S. Suzuki, Y. Takubo, S. Terada, K. Tokushuku, S. Tsuno, I. Ueda, Y. Unno, J. Usui, A. Yamamoto, Y. Yasu

*Kobe*

Y. Chen, K. Kawade, S. Kido, H. Kurashige, J. Maeda, A. Ochi, S. Shimizu, K. Takeda, R. Tanioka, K. Wakamiya, F. Yamane, Y. Yamazaki

*Kyoto*

S. Akatsuka, T. Kunigo, Y. Noguchi, Y. Okazaki, T. Sumida, T. Tashiro

*Kyoto UE*

R. Takashima

*Kyushu*

K. Kawagoe, D. Kobayashi, S. Oda, H. Otono, S. Shirabe, J. Tojo

*Nagasaki*

T. Fusayasu, M. Shimojima

*Nagoya*

Y. Horii, T. Kawaguchi, Y. Nakahama, K. Onogi, Y. Sano, M. Tomoto

*Okayama*

I. Nakano

*Osaka*

K. Hanagaki, M. Hirose, N. Ishijima, M. Nomachi, Y. Sugaya, J.J. Teoh, K. Yajima

*Shinshu*

Y. Hasegawa, T. Takeshita

*Tokyo ICEPP*

S. Adachi, S. Asai, I. Chiu, Y. Enari, R. Iguchi, M. Ishino, Y. Kano, Y. Kataoka, C. Kato, T. Kawamoto, T. Kishimoto, A. Kobayashi, T. Kobayashi, C. Kozakai, T. Mashimo, T. Masubuchi, Y. Minegishi, T. Mori, T. Nakamura, T. Nobe, Y. Okumura, M. Saito, T. Saito, H. Sakamoto, R. Sawada, J. Tanaka, K. Terashi, K. Uno, T. Yamanaka, M. Yamatani, T. Yamazaki, Y. Yang

*Tokyo MU*

U. Bratzler, C. Fukunaga, T. Kumita



*Tokyo Tech*

M. Ishitsuka, O. Jinnouchi, S. Kawaguchi, M. Kuze, K. Motohashi, M. Tanaka, K. Todome, D. Yamaguchi, Y. Yamaguchi, Y. Yazawa

*Tsukuba*

M. Hagihara, K. Hara, S. Honda, F. Ito, K. Kasahara, S.H. Kim, K. Nagata, H. Okawa, K. Sato, F. Ukegawa

*Waseda*

T. Iizawa, T. Ikai, T. Kaji, Y. Kawaguchi, T. Mitani, M. Morinaga, T. Nitta, R. Watari, K. Yorita

**Morocco**

*Casablanca*

D. Benchekrout, A. Chafaq, A. Hoummada

*Marrakesh*

M. El Kacimi, D. Goujdami

*Oujda*

M. Aaboud, J.E. Derkaoui, M. Ouchrif

*Rabat*

S. Batlamous, R. Cherkaoui El Moursli, S. Dahbi, M. Ezzi, F. Fassi, N. Haddad, Z. Idrissi, Y. Tayalati

**Netherlands**

*Nijmegen*

S. Alderweireldt, S. Caron, L. Colasurdo, N. De Groot, M. Donszelmann, V. Fabiani, F. Filthaut, O. Igonkina, A.C. König, S. Nektarijevic, J.F.P. Schouwenberg

*Nikhef*

I. Angelozzi, H. Arnold, P.J. Bakker, M. Bedognetti, S. Bentvelsen, G.J. Bobbink, A. Borga, H. Boterenbrood, L.S. Bruni, C.D. Burgard, R. Castelijm, W.S. Chan, Y.S. Chow, A.P. Colijn, P. de Jong, D. Denysiuk, D. Duda, P. Ferrari, O. Igonkina, P. Kluit, E. Koffeman, S. Liem, M. Morgenstern, K.P. Oussoren, G. Sabato, D. Salek, F.P. Schreuder, B.S. Stapf, W. Van Den Wollenberg, H. van der Graaf, P. Vankov, I. van Vulpen, M.C. van Woerden, W. Verkerke, A.T. Vermeulen, J.C. Vermeulen, M. Vreeswijk, T.M.H. Wolf

**Norway**

*Bergen*

T. Buanes, O. Dale, G. Eigen, N. Fomin, A. Lipniacka, S. Maeland, B. Martin dit Latour, B. Stugu, Z. Yang, J. Zalieckas

*Oslo*

M.K. Bugge, D. Cameron, J.R. Catmore, S. Feigl, L. Franconi, V. Garonne, E. Gramstad, S. Hellesund, V. Morisbak, H. Oppen, F. Ould-Saada, A.L. Read, O. Røhne, H. Sandaker, C. Serfon, K.O.H. Vadla

## **Poland**

*Krakow AGH-UST*

L. Adamczyk, T. Bold, W. Dabrowski, G.P. Gach, I. Grabowska-Bold, M.P. Guzik, P.A. Janus, S. Koperny, J.A. Kremer, B. Mindur, A. Ogrodnik, M. Przybycien

*Krakow IFJ PAN*

M. Bahmani, E. Banas, P.A. Bruckman de Renstrom, K. Burka, J.J. Chwastowski, S. Czekierda, D. Derendarz, B.S. Dziejczak, A. Kaczmarska, K. Korcyl, A.B. Kowalewska, Pa. Malecki, A. Olszewski, J. Olszowska, M. Slawinska, E. Stanecka, R. Staszewski, M. Trzebinski, A. Trzupek, M.W. Wolter, B.K. Wosiek, K.W. Woźniak, B. Zabinski

*Krakow Jagiellonian*

G. Korcyl, M. Palka, E. Richter-Was, P. Strzempek

## **Portugal**

*Coimbra*

S.P. Amor Dos Santos, B. Galhardo, F. Veloso, H. Wolters

*LIP*

J.A. Aguilar-Saavedra, S.P. Amor Dos Santos, J.P. Araque, N.F. Castro, P. Conde Muiño, A.M. Da Silva Pina, T. Dias do Vale, R.M. Faisca Rodrigues Pereira, M.C.N. Fiolhais, B. Galhardo, A. Gomes, R. Gonçalo, P.M. Jorge, A. Maio, J. Maneira, E.D. Mendes Gouveia, L.F. Oleiro Seabra, A. Onofre, R. Pedro, A.P. Pereira Peixoto, J.C. Rufino Amaro, H. Santos, J.G. Saraiva, J. Silva, A. Tavares Delgado, F. Veloso, H. Wolters

*Lisboa*

A. Gomes, P.M. Jorge, A. Maio, J. Maneira, J.A. Soares Augusto, A. Tavares Delgado

*Minho*

J.P. Araque, N.F. Castro, A.M. Da Silva Pina, T. Dias do Vale, E.D. Mendes Gouveia, A. Onofre, A.P. Pereira Peixoto

## **Republic of Belarus**

*Minsk AC*

S. Harkusha, Y. Kulchitsky, Y.A. Kurochkin, P.V. Tsiareshka

*Minsk NC*

A. Hrynevich

**Romania**

*Brasov TU*

T.T. Tulbure

*Bucharest IFIN-HH*

C. Alexa, M. Caprini, A. Chitan, S. Constantinescu, M. Dobre, O.A. Ducu, A.E. Dumitriu,  
C. Galea, A. Jinaru, V.S. Martoiu, J. Maurer, A. Olariu, M. Rotaru, G. Stoicea, G. Tarna,  
A. Tudorache, V. Tudorache

*ITIM*

G.A. Popeneciu

*Iasi UAIC*

C. Agheorghiesei

*Politehnica Bucharest*

A.C. Contescu

*Timisoara WU*

P.M. Gravila

**Russia**

*Moscow FIAN*

A.V. Akimov, F. Dubinin, I.L. Gavrilenko, R. Mashinistov, P.Yu. Nechaeva, A. Shmeleva,  
A.A. Snesev, V.V. Sulin, V.O. Tikhomirov, K. Zhukov

*Moscow ITEP*

A. Artamonov, A. Gavriilyuk, P.A. Gorbounov, P.B. Shatalov, I.I. Tsukerman

*Moscow MEPhI*

K. Belotskiy, N.L. Belyaev, O. Bulekov, V.A. Kantserov, D. Krasnopevtsev,  
D. Ponomarenko, N. Proklova, A. Romaniouk, E. Shulga, N. Smirnov, S.Yu. Smirnov,  
Y. Smirnov, E.Yu. Soldatov, V.O. Tikhomirov, S. Timoshenko, K. Vorobev

*Moscow SU*

A.S. Boldyrev, L.K. Gladilin, V.A. Kramarenko, A.S. Maevskiy, S.Yu. Sivoklokov,  
L.N. Smirnova

*Novosibirsk State University*

A.V. Anisenkov, E.M. Baldin, V.S. Bobrovnikov, A.G. Bogdanchikov, A.R. Buzykaev, S. Dubrov, V.F. Kazanin, A.G. Kharlamov, T. Kharlamova, A.A. Korol, A.L. Maslennikov, D.A. Maximov, N. Melnikova, S.V. Peleganchuk, P. Podberezko, O.L. Rezanova, A.M. Soukharev, A.A. Talyshev, Yu.A. Tikhonov, V. Zhulanov

*Petersburg NPI*

A. Basalava, A. Ezhilov, O.L. Fedin, V. Gratchev, A. Kazarov, M. Levchenko, V.P. Maleev, I. Naryshkin, S. Olechko, Y.F. Ryabov, V.A. Schegelsky, A. Shcherbakova, V. Solovyev

*Protvino IHEP*

A. Borisov, E. Cheremushkina, S.P. Denisov, R.M. Fakhruddinov, A.B. Fenyuk, D. Golubkov, A. Kamenshchikov, A.N. Karyukhin, A.S. Kozhin, A.A. Minaenko, A.G. Myagkov, V. Nikolaenko, A. Ryzhov, A.A. Solodkov, O.V. Solovyanov, E.A. Starchenko, A.M. Zaitsev, O. Zenin

*Tomsk SU*

A. Khodinov, A. Vaniachine

**JINR**

*Joint Institute for Nuclear Research, Dubna, Russia*

F. Ahmadov, I.N. Aleksandrov, E. Alexandrov, V.A. Bednyakov, I.R. Boyko, I.A. Budagov, G.A. Chelkov, A. Cheplakov, M.V. Chizhov, D.V. Dedovich, M. Demichev, A. Gongadze, M.I. Gostkin, N. Huseynov, N. Javadov, Z.M. Karpova, S.N. Karpov, E. Khramov, V.M. Kotov, U. Kruchonak, V. Kukhtin, E. Ladygin, V. Lyubushkin, M. Mineev, E. Plotnikova, I.N. Potrap, N.A. Rusakovich, R. Sadykov, A. Sapronov, M. Shiyakova, L. Simic, A. Soloshenko, S. Turchikhin, V.B. Vinogradov, I. Yeletsikh, A. Zhemchugov, N.I. Zimine

**Serbia**

*Belgrade IP*

J. Krstic, Dj. Sijacki, M. Vranjes Milosavljevic, N. Vranjes, L. Živković

**Slovak Republic**

*Bratislava*

R. Astalos, P. Bartos, T. Blazek, T. Dado, O. Majersky, M. Melo, J. Smiesko, I. Sykora, S. Tokár, T. Ženiš

*Kosice*

D. Bruncko, E. Kladiva, P. Murin, F. Sopkova, P. Strizenec, J. Urban

**Slovenia**

*Ljubljana*

V. Cindro, A. Filipčič, A. Gorišek, B. Hiti, L. Kanjir, B.P. Kerševan, G. Kramberger,  
I. Mandić, B. Maček, M. Mikuž, M. Muškinja, T. Novak, G. Sokhrannyi, T. Šfiligoj

**South Africa**

*Cape Town*

R.J. Atkin, A. Hamilton, C.J. Lee, C. Mwewa, S. Yacoob

*Johannesburg*

S. Ballestrero, D. Casadei, S.H. Connell, N. Govender, L. Truong

*Witwatersrand*

Y. Hernández Jiménez, D.R. Hlaluku, H. Jivan, D. Kar, J.E. Mdhluli, B.R. Mellado Garcia,  
R.G. Reed, D. Roy, X. Ruan, E. Sideras Haddad, S.E. von Buddenbrock

**Spain**

*Barcelona*

M. Bosman, M.P. Casado, M. Casolino, E. Cavallaro, M. Cavalli-Sforza, C. Fischer,  
F.A. Förster, D. Gerbaudo, E.L. Gkougkousis, J. Glatzer, S. Grinstein, A. Juste Rozas,  
I. Korolkov, J.C. Lange, I. Lopez Paz, M. Martinez, L.M. Mir, A. Pacheco Pages,  
C. Padilla Aranda, I. Riu, C. Rizzi, Y. Rodina, A. Rodriguez Perez, S. Terzo, M.F. Tripiana,  
S. Tsiskaridze, T.R. Van Daalen, D. Vazquez Furelos, G. Volpi, R. Zaidan

*Granada*

J.A. Aguilar-Saavedra, D. Melini

*Madrid UA*

F. Barreiro, S. Calvente Lopez, A. Cueto, J. Del Peso, C. Glasman, J. Terron

*Valencia*

A.J. Bailey, L. Barranco Navarro, S. Cabrera Urbán, V. Castillo Gimenez, L. Cerda Alberich,  
M.J. Costa, C. Escobar, O. Estrada Pastor, A. Ferrer, L. Fiorini, J. Fuster, C. García,  
J.E. García Navarro, S. González de la Hoz, E. Higón-Rodríguez, J. Jimenez Pena,  
C. Lacasta, J.J. Lozano Bahilo, D. Madaffari, J. Mamuzic, S. Marti-Garcia, D. Melini,  
V.A. Mitsou, S. Pedraza Lopez, S. Rodriguez Bosca, D. Rodriguez Rodriguez,  
E. Romero Adam, J. Salt, U. Soldevila, J. Sánchez, A. Valero, J.A. Valls Ferrer, M. Vos,  
D. Álvarez Piqueras

**Sweden**

*Lund*

S.S. Bocchetta, E.E. Corrigan, C. Doglioni, E. Hansen, V. Hedberg, G. Jarlskog,  
C.W. Kalderon, E. Kellermann, B. Konya, E. Lytken, K.H. Mankinen, J.U. Mjörnmark,  
R. Poettgen, T. Poulsen, O. Smirnova, O. Viazlo, T.P.A. Åkesson

*Stockholm*

G. Bertoli, O. Bessidskaia Bylund, C. Boehm, R.M.D. Carney, C. Clement, K. Gellerstedt,  
S. Hellman, K. Jon-And, D.A. Milstead, T. Moa, S. Molander, P. Pasuwan, N.W. Shaikh,  
S.B. Silverstein, J. Sjölin, S. Strandberg, M. Ughetto, E. Valdes Santurio, V. Wallangen

*Stockholm KTH*

A. Kastanas, B. Lund-Jensen, C.C. Ohm, G. Ripellino, P.E. Sidebo, J. Strandberg

*Uppsala*

E.M. Asimakopoulou, E. Bergeaas Kuutmann, P. Bokan, R. Brenner, T. Ekelof, M. Ellert,  
A. Ferrari, P.O.J. Gradin, M.F. Isacson, M.U.F. Martensson, H. Ohman, P.H. Sales De Bruin

**Switzerland**

*Bern*

J.K. Anders, Y. Bandi, H.P. Beck, A. Ereditato, S. Haug, F. Meloni, C. Merlassino, A. Miucci,  
G.A. Mullier, M. Rimoldi, M.S. Weber, T.D. Weston

*Geneva*

E. Akilli, C.S. Amrouche, L.S. Ancu, M. Benoit, N. Calace, A. Clark, D. della Volpe,  
F.A. Di Bello, A. Dubreuil, W.J. Fawcett, D. Ferrere, S. Gadatsch, T. Golling,  
S. Gonzalez-Sevilla, G. Iacobucci, R. Jansky, A. Katre, T.J. Khoo, M. Kiehn,  
M.C. Lanfermann, A.L. Lioni, L. March, P. Mermod, M. Nessi, C.E. Pandini, L. Paolozzi,  
D. Salamani, S. Schramm, A. Sfyrla, D.M.S. Sultan, M. Valente, X. Wu

**Taiwan**

*Hsinchu NTHU*

K. Cheung, P.J. Hsu, Y.J. Lu

*Taipei AS*

A. Buzatu, S. Hou, S.C. Lee, R. Mazini, L. Shi, D.A. Soh, P.K. Teng, S.M. Wang, W. Wang,  
Y. Yang, G. Zhang

*Taipei ASGC*

S.C. Lin

**Turkey**

*Ankara*

O. Cakir, H. Duran Yildiz

*Bahcesehir*

A.J. Beddall

*Bilgi*

E. Celebi, S.A. Cetin

*Bogazici Istanbul*

S. Gurbuz, S. Istin, V.E. Ozcan

*Gaziantep*

A. Beddall, A. Bingul

*Istanbul Aydin*

S. Kunday, I. Turk Cakir

*TOBB*

S. Sultansoy

**United Kingdom**

*Birmingham*

P.P. Allport, N. Andari, M.J. Baca, J. Bracinik, D.L. Briglin, J.H Broughton, D.G. Charlton, A.C. Daniells, A.G. Foster, L. Gonella, F. Gonnella, C.M. Hawkes, S.J. Hillier, J.J. Kempster, J. Kendrick, A. Mazurov, P.R. Newman, K. Nikolopoulos, R.E. Owen, E. Reynolds, M. Slater, R.J. Staley, J.P. Thomas, P.D. Thompson, R.J. Turner, R.A. Vallance, P.M. Watkins, A.T. Watson, M.F. Watson, J.A. Wilson, S.D. Worm

*Cambridge*

M. Arratia, N. Barlow, J.R. Batley, B.H. Brunt, J.D. Chapman, G. Cottin, E.H. Flaherty, P.J. Garsed, J.C. Hill, C.G. Lester, C.Y. Lin, C. Malone, M.A. Parker, C.J. Potter, D. Robinson, J.H.N. Rosten, R.S. Shaw, C.P. Ward, S. Williams, I. Yusuff

*Edinburgh*

T.M. Bristow, P.J. Clark, M. Fauci Giannelli, A. Hasib, M.P. Heath, X. Hoad, C. Leonidopoulos, V.J. Martin, L. Mijović, A. Strubig, A. Søgaard, A.J. Taylor, A. Washbrook, B.M. Wynne

*Glasgow*

A.A. Alshehri, R.L. Bates, A. Blue, S.K. Boutle, W.D. Breaden Madden, D. Britton, A.G. Buckley, P. Bussey, C.M. Buttar, S.J. Crawley, S. D'Auria, A.T. Doyle, A.K. Duncan, M.J. Fenton, C. Gray, U. Gul, P. Mullen, V. O'Shea, M. Owen, A. Robson, A.S. Thompson, K. Wraight

*Lancaster*

A.E. Barton, M.D. Beattie, I.A. Bertram, G. Borissov, E.V. Bouhova-Thacker, H. Fox, R.C.W. Henderson, R.W.L. Jones, V. Kartvelishvili, R.E. Long, P.A. Love, D. Muenstermann, A.J. Parker, M.B. Skinner, M. Smizanska, A.S. Tee, J. Walder, A.M. Wharton, B.W. Whitmore

*Liverpool*

S. Burdin, M. D'Onofrio, P. Dervan, Y. Gao, E.C. Graham, C.B. Gwilliam, H.S. Hayward, T.J. Jones, E.F. Kay, B.T. King, M. Klein, U. Klein, J. Kretzschmar, P. Laycock, S.J. Maxfield, A. Mehta, L. Meng, N. Rompotis, J.H. Vosseveld

*London QMUL*

L.J. Armitage, A.J. Bevan, U. Blumenschein, M. Bona, E. Eisenhandler, R. Gamboa Goni, J.M. Hays, M.P.J. Landon, D. Lewis, S.L. Lloyd, D.A. Millar, J.D. Morris, T. Nooney, E. Rizvi, T.J. Stevenson, E. Tahirovic

*London RHBNC*

D.K. Abhayasinghe, R.E. Ardell, T. Berry, V. Boisvert, A.J. Bozson, T. Brooks, G. Cowan, S. George, S.M. Gibson, B. Green, C.R. Kilby, J.G. Panduro Vazquez, Fr. Pastore, B.C. Sowden, F. Spanò, P. Teixeira-Dias, D.W. Thomas, L.J. Wilkins

*London UC*

A.S. Bell, J.M. Butterworth, M. Campanelli, V. Christodoulou, G.J. Crone, G. Facini, D. Freeborn, K. Gregersen, Z.J. Grout, N.G. Gutierrez Ortiz, C. Gutsche, G.G. Hesketh, N. Konstantinidis, A. Korn, K.J.C. Leney, A.C. Martyniuk, L.I. McClymont, M. Montella, E. Nurse, S. Richter, T. Scanlon, P. Sherwood, B. Simmons, A. Vaidya, D.R. Wardrope, M. Warren, B.M. Waugh, D.P. Yallup

*Manchester*

A. Bethani, R. Bielski, I.A. Connelly, B.E. Cox, J. Crane, N.D. Dann, C. Da Via, G.T. Forcolin, A.C. Forti, J. Howarth, D.P.J. Lack, J. Masik, S.B. Menary, F.J. Munoz Sanchez, A. Oh, E.C. Orgill, J.R. Pater, R.F.Y. Peters, A.D. Pilkington, D. Price, Y. Qin, J.A. Raine, J.H. Rawling, R.T. Roberts, N. Scharmberg, H. Schweiger, S.M. Shaw, S. Watts, F. Wilk, T.R. Wyatt

*Oxford*

L. Ambroz, G. Artoni, M. Backes, A.J. Barr, L. Beresford, D. Bortoletto, J.T.P. Burr, A.M. Cooper-Sarkar, J.A. Frost, E.J. Gallas, F. Giuli, C. Gwenlan, C.P. Hays, T.B. Huffman, C. Issever, J.K.K. Liu, L. Marchese, K. Nagai, M.E. Nelson, R.B. Nickerson, N. Norjoharuddeen, M. Petrov, M.A. Pickering, A. Sharma, I.P.J. Shipsey, C. Tosciri, J.C.-L. Tseng, G.H.A. Viehhauser, L. Vigani, A.R. Weidberg, G. Zemaityte, M. Zgubič

*RAL*



T. Adye, J.T. Baines, B.M. Barnett, I.P. Brawn, S. Burke, A. Dewhurst, J. Dopke, D. Emelianov, B.J. Gallop, C.N.P. Gee, S.J. Haywood, J. Kirk, S. Martin-Haugh, S.J. McMahon, R.P. Middleton, W.J. Murray, P.W. Phillips, W. Qian, D.P.C. Sankey, C. Sawyer, M.J. Siyad, W. Song, S. Taghavirad, F.J. Wickens, M. Wielers

*Sheffield*

C. Anastopoulos, M.T. Anthony, D. Costanzo, T. Cuhadar Donszelmann, I. Dawson, S. Fracchia, G.N. Hamity, M.C. Hodgkinson, P. Johansson, E.V. Korolkova, E. Kourlitis, K. Lohwasser, C.M. Macdonald, P.S. Miyagawa, H.J. Moss, B. Ravina, P. Sommer, D.R. Tovey, O.E. Vickey Boeriu, T. Vickey, T.G. Zorbas

*Sussex*

N.L. Abraham, B.M.M. Allbrooke, L. Asquith, A. Cerri, C.A. Chavez Barajas, A. De Santo, M. Grandi, S.D. Jones, D. Kelsey, G. Lerner, F. Miano, B. Safarzadeh Samani, F. Salvatore, I. Santoyo Castillo, C.Y. Shehu, M. Spina, K. Suruliz, M.R. Sutton, F. Tresoldi, F. Trovato, I. Vivarelli, E. Winkels, O.J. Winston, I. Xiotidis

*Warwick*

J.S. Ennis, S.M. Farrington, P.F. Harrison, A. Jelinskas, C. Jeske, T.A. Martin, C.J. McNicol, W.J. Murray, E. Pianori, M. Spangenberg

**United States of America**

*Albany*

V.M.M. Cairo, V. Jain, S.P. Swift

*Argonne*

Y. Abulaiti, R.E. Blair, S. Chekanov, G. Drake, T. LeCompte, B. Liu, J. Love, D. Malon, J. Metcalfe, M.B. Oberling, A. Paramonov, J. Proudfoot, S. Ryu, P. van Gemmeren, R. Wang, J.S. Webster, J. Zhang

*Arizona*

E. Cheu, C.M. Delitzsch, K.A. Johns, S. Jones, W. Lampl, M. LeBlanc, R. Leone, P. Loch, R. Nayyar, J.P. Rutherford, E.W. Varnes, Y. Zhou

*Arlington UT*

A. Brandt, D. Bullock, S. Darmora, K. De, A. Farbin, L. Feremenga, J. Griffiths, H.K. Hadavand, L. Heelan, H.Y. Kim, J.D. Little, N. Ozturk, E. Sarkisyan-Grinbaum, G. Usai, A. White, J. Yu

*Austin*

T. Andeen, R. Narayan, N. Nikiforou, P.U.E. Onyisi, H. Potti, A.F. Webb

*Berkeley LBNL*

B.T. Amadio, B. Axen, R.M. Barnett, J. Beringer, P. Calafiura, F. Cerutti, A. Ciocio, R.N. Clarke, J. Dickinson, A. Dimitrievska, E.M. Duffield, K. Einsweiler, S. Farrell, A. Gabrielli, M. Garcia-Sciveres, H.M. Gray, C. Haber, T. Heim, I. Hinchliffe, L. Jeanty, K. Krizka, C. Leggett, P. Liu, Z. Marshall, B.P. Nachman, S. Pagan Griso, M. Shapiro, V. Tsulaia, S. Viel, H. Wang, H. Yang, W-M. Yao

*Boston*

K.M. Black, J.M. Butler, L. Dell’Asta, K.D. Finelli, D.E. Gastler, E. Hazen, B.A. Long, J.T. Shank, A.D. Sherman, Z. Yan, E. Yigitbasi

*Brandeis*

C. Amelung, G. Barone, J.R. Bensinger, L.J. Bergsten, C. Blocker, S. Dhaliwal, D. Dodsworth, M. Goblirsch-Kolb, H. Herde, K.M. Loew, K. O’Connor, Z.M. Schillaci, G. Sciolla

*Brookhaven BNL*

K. Assamagan, M. Beigel, V. Cavaliere, H. Chen, K. Chen, J. Elmsheuser, H.A. Gordon, Q. Hu, G. Iakovidis, A. Klimentov, V. Kouskoura, F. Lanni, E. Lançon, C.A. Lee, H. Liu, D. Lynn, T. Maeno, H. Ma, E. Mountricha, P. Nilsson, M.A. Nomura, D. Oliveira Damazio, D.V. Perepelitsa, M.-A. Pleier, V. Polychronakos, S. Protopopescu, S. Rajagopalan, G. Redlinger, S. Snyder, P. Steinberg, S.A. Stucci, H. Takai, S. Tang, A. Tricoli, A. Undrus, T. Wenaus, W. Wu, H. Xu, L. Xu, S. Ye

*Chicago*

J. Alison, M. Bogdan, P. Bryant, R. Camacho Toro, R.W. Gardner, K. Hildebrand, T.R. Holmes, L.A. Horyn, Y.K. Kim, D.W. Miller, M.J.R. Olsson, M.J. Oreglia, J.E. Pilcher, D. Schaefer, T. Seiss, M.J. Shochet, G.H. Stark, M. Swiatlowski, F. Tang, I. Vukotic, M. Wu, R. Zou

*Columbia*

A. Angerami, G. Brooijmans, R.M. Carbone, J. Chen, M.R. Clark, B. Cole, L.B. Havener, E.W. Hughes, K. Iordanidou, S. Mohapatra, I. Ochoa, J.A. Parsons, M.N.K. Smith, R.W. Smith, Y. Tian, P.M. Tuts, T. Wang

*Dallas SMU*

M. Feickert, R. Gupta, J.W. Hetherly, S. Kama, R. Kehoe, F. Lo Sterzo, M.A. McKay, S.J. Sekula, R. Stroynowski, T. Varol, P. Wang, J. Ye, X. Zhao, L. Zhou

*Dallas UT*

J.M. Izen, B. Meirose, K. Reeves

*Duke*

A.T.H. Arce, D.P. Benjamin, D.M. Bjergaard, A. Bocci, D.R. Davis, M.B. Epland, M. Feng,  
A.T. Goshaw, A. Kotwal, M.C. Kruse, S.H. Oh

*Harvard*

S.K. Chan, K.F. Di Petrillo, M. Franklin, P. Giromini, J.L. Gonski, J. Huth, L. Lee, M. Morii,  
C.S. Rogan, J. Roloff, B. Tong, A.N. Tuna, A.M. Wang, S. Zambito

*Indiana*

P. Calfayan, K. Choi, H. Evans, J.C. Gass, C.A. Johnson, R. Kopeliansky, B.M. Kunkler,  
S. Lammers, R.A. Linck, F. Luehring, G. Palacino, J. Penwell, C.P.A. Roland,  
R. Van Kooten, B. Weinert, D. Zieminska

*Iowa*

S. Argyropoulos, J. Benitez, U. Mallik, S. Yang

*Iowa State*

C. Chen, J. Cochran, F. De Lorenzi, H. Jiang, N. Krumnack, B. Liu, D. Pluth, S. Prell,  
M.D. Werner, J. Yu

*Louisiana Tech*

D. Bakshi Gupta, A.K. Elliott, Z.D. Greenwood, G.C. Grossi, D.K. Jana, L. Sawyer,  
M. Wobisch

*Massachusetts*

N.R. Bernard, B. Brau, R. Caminal Armadans, C. Dallapiccola, M.S. Lutz,  
V.I. Martinez Outschoorn, Z.A. Meadows, E.J.W. Moyse, N.E. Pettersson, A. Picazio,  
S. Willocq

*Michigan*

D. Amidei, H.C. Cheng, T. Dai, E.B. Diehl, R.C. Edgar, C. Ferretti, P. Fleischmann, C. Geng,  
L. Guan, W. Guo, Y. Guo, R. Hyneman, M.H. Klein, D. Levin, B. Li, H. Liu, N. Lu,  
D.E. Marley, A. McCarn, S.P. Mc Kee, X.T. Meng, H.A. Neal, J. Qian, T.A. Schwarz,  
J. Searcy, K. Sekhon, I. Siral, S. Sun, A.S. White, Z. Xi, W. Xu, D. Zhang, B. Zhou, J. Zhu

*Michigan SU*

G. Arabidze, R. Brock, A. Chegwidden, H. De la Torre, D. Edmunds, I. Ermoline,  
T. Farooque, W.C. Fisher, R.R. Gregory, G. Halladjian, R. Hauser, D. Hayden, J. Huston,  
P. Laurens, K. Lin, J.T. Linnemann, M.C. Mondragon, F.H. Phillips, P. Plucinski,  
I. Pogrebnyak, B.G. Pope, R. Schwienhorst, D. Shooltz, C. Willis

*NYU New York*

C. Becot, K. Cranmer, V. Croft, A. Haas, L. Heinrich, B. Kaplan, R. Konoplich, A.I. Mincer,  
P. Nemethy, M. Ronzani, A. Sakharov, C.J. Treado

*New Mexico*

A. Grummer, M.R. Hoferkamp, N.C. Mc Fadden, S.C. Seidel, A.C. Taylor, K. Toms

*Northern Illinois*

J. Adelman, E. Brost, T.J. Burch, B. Burghgrave, DC Chakraborty, P. Klimek, P. Saha

*Ohio SU*

J.B. Beacham, A. Boveia, S. Che, K.K. Gan, B. Gui, H. Kagan, R.D. Kass, K.A.Looper, C.B. Martin, S. Shrestha, B.B. Tannenwald, E. Tolley

*Oklahoma*

B. Abbott, M. Alhroob, A. De Benedetti, G. Gustavino, P. Gutierrez, H. Severini, Y. Shen, D.R. Shope, P. Skubic, M. Strauss, J. Stupak, Q. Wang

*Oklahoma SU*

J. Cantero, J. Haley, D.O. Jamin, A. Khanov, F. Rizatdinova

*Oregon*

B.W. Allen, J.T. Barkeloo, J.S. Bonilla, J.E. Brau, A. Dattagupta, C. Dudley, W.H. Hopkins, S. Majewski, J. Myers, N.B. Sinev, I.M. Snyder, M.C. Stockton, D.M. Strom, E. Torrence, K. Whalen, F. Winklmeier

*Pennsylvania*

W.K. Balunas, S. Chen, R.A. Creager, J.R. Dandoy, W.K. Di Clemente, R.R.M. Fletcher, L.M. Flores, B. Haney, T.C. Herwig, J. Kroll, E. Lipeles, J. Machado Miguens, C. Meyer, K.P. Mistry, J. Reichert, E.D. Resseguie, L. Schaefer, E. Thomson, H.H. Williams, K. Yoshihara

*Pittsburgh*

R.M. Bianchi, J. Boudreau, N. Brusino, B.T. Carlson, T.M. Hong, J. Mueller, J. Su

*SLAC*

T. Barklow, R. Bartoldus, H.S. Bawa, C. Bernius, J.E. Black, A.R. Cukierman, Y.S. Gao, N. Garelli, P. Grenier, K. Grimm, N. Ilic, Z. Jiang, M. Kagan, M. Kocian, T. Koi, J. Moss, R. Mount, F. Rubbo, A. Salnikov, A. Schwartzman, D. Su, L. Tompkins, M. Wittgen, C. Young, Q. Zeng

*Santa Cruz UC*

O.S. AbouZeid, A.A. Affolder, M. Battaglia, C. Debenedetti, M. Gignac, A.A. Grillo, M. Hance, A.T. Law, A.M. Litke, S.M. Mazza, J. Nielsen, J.M.P. Pasner, R. Reece, P. Rose, H.F-W. Sadrozinski, S. Schier, B.A. Schumm, A. Seiden, J.D. Shahinian, N.L. Woods

*Seattle Washington*

S.P. Alkire, C. Alpigiani, A.G. Goussiou, A. Hostiuc, S.-C. Hsu, W.J. Johnson, H.J. Lubatti, S. Meehan, R. Rosten, J. Rothberg, J. Schaarschmidt, E. Torr o Pastor, M.S. Twomey, G. Watts, N.L. Whallon

*Stony Brook*

C.P. Bee, A. Behera, T.P. Calvet, C. Hayes, J. Hobbs, P. Huo, J. Jia, B.E. Lindquist, L. Morvaj, G. Piacquadio, S.K. Radhakrishnan, M. Rijssenbeek, R.D. Schamberger, V. Tsiskaridze, D. Tsybychev, M. Zhou

*Tufts*

P.H. Beauchemin, F. Sforza, K. Sliwa, H. Son

*UCI*

D.J.A. Antrim, A. Armstrong, K.T. Bauer, D.W. Casper, A. Corso-Radu, M. Frate, J. Gramling, D. Guest, S. Kolos, A.J. Lankford, A.S. Mete, Y.W.Y. Ng, K. Ntekas, D.A. Scannicchio, M. Schernau, A.M. Soffa, I. Soloviev, A. Taffard, G. Unel, D. Whiteson, S.C. Yildiz

*Urbana UI*

M. Atkinson, H. Cai, Y. Cao, P. Chang, S. Errede, B.H. Hooberman, M. Khader, Y.P. Kulinich, T.M. Liss, L. Liu, J.D. Long, M.S. Neubauer, A. Puri, M. Rybar, R. Shang, A.M. Sickles, J.C. Zeng, M. Zhang, D. Zhong

*Wisconsin*

Sw. Banerjee, N.P. Dang, W. Guan, A.S. Hard, Y. Heng, H. Ji, X. Ju, L.S. Kaplan, L. Kashif, Y. Ming, A. Pathak, M. Silva Jr., F. Wang, W. Wiedenmann, S.L. Wu, F. Zhang, C. Zhou, G. Zobernig

*Yale*

O.K. Baker, E. Benhar Noccioli, S. Demers, M. Paganini, M. Pettee, C.O. Shimmin, S.J. Thais, L.A. Thomsen, P. Tipton, J.G. Vasquez, C. Weber





0100101101110101 0100101101110101 0100101101110101  
0100101101110101 0100101101110101