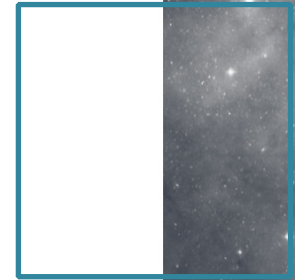


A hierarchical approach to Big Data



Hierarchical Progressive Surveys (HiPS)
and Multi-Order Coverage (MOC) Maps



Mark G. Allen
Pierre Fernique



□ Introduction

- Scientific exploitation of Big Data requires practical solutions for:
 - Data Access
 - Visualisation
 - Analysis
- Hierarchical sky tessellation techniques
 - Working now!
 - Scalable to Big Data



□ Introduction

- Starting points
 - Hundreds of image surveys, and thousands of catalogues at CDS
 - ~2 TB per band for all-sky surveys
 - Catalogues up to ~2 billion rows
 - Need for interactive display/comparison
 - Need to maintain scientific integrity of data, and links to original data
 - Ease of use and interoperability

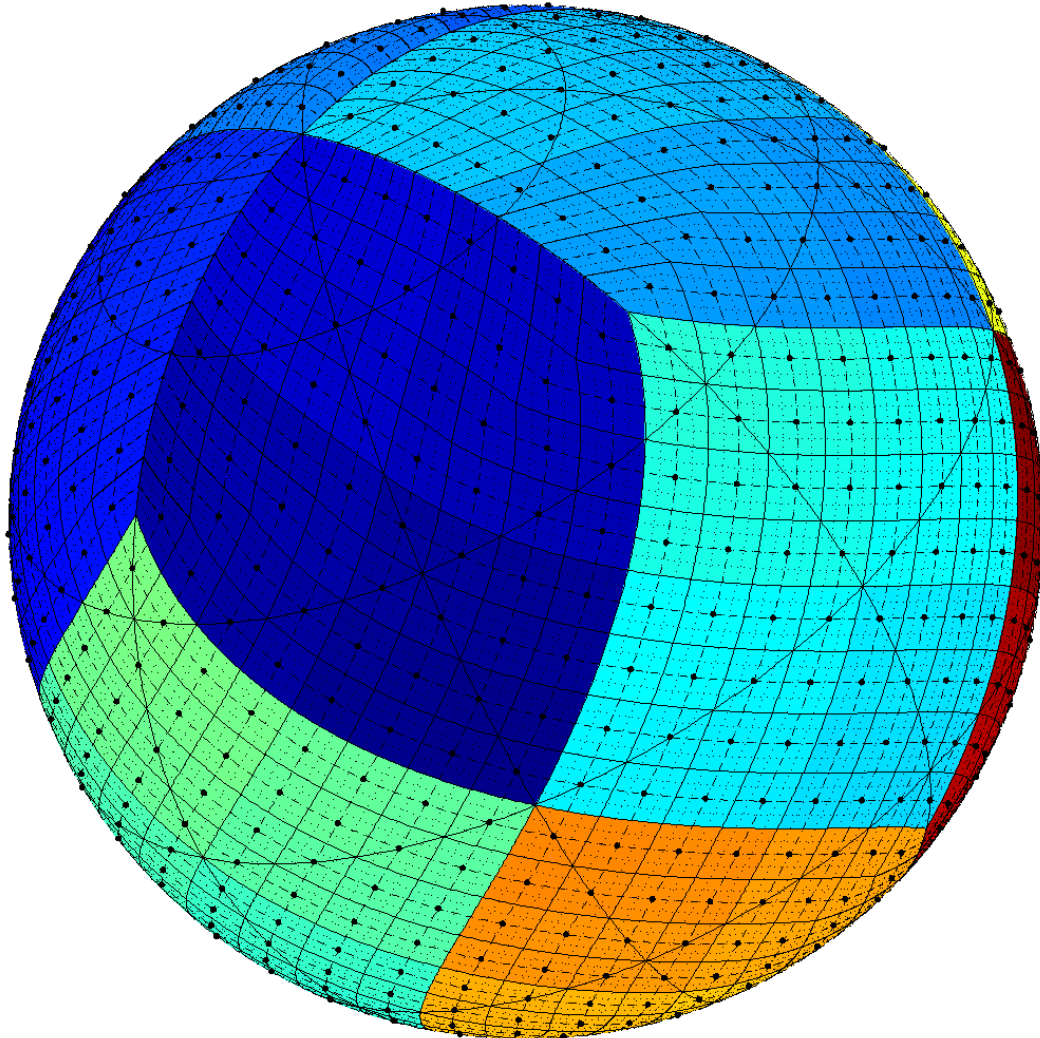


□ HiPS

- HiPS: **H**ierarchical **P**rogressive **S**urveys
 - Multi-resolution HEALPix data structure for
 - *Images*
 - *Catalogues*
 - *3-dimensional data cubes*
 - Conserves scientific data properties alongside visualisation considerations
 - Implemented for ~250 data sets and growing
 - New levels of interoperability

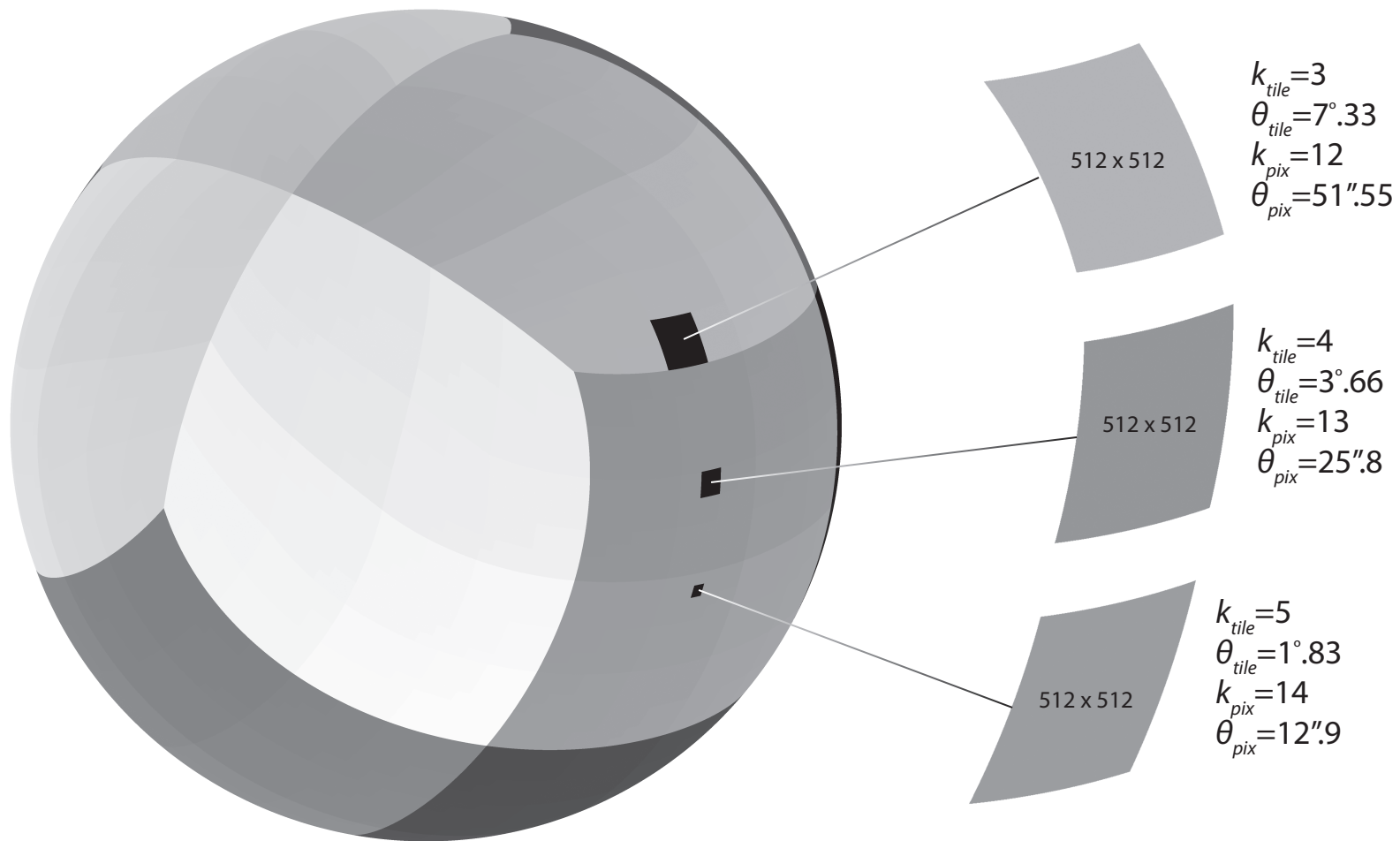


□ HEALPix*



- 12 quadrilateral pixels
- 2x2 division at each level
- Equal area
- Iso-latitude
- Nested index scheme encodes inheritance
- Libraries

HiPS – Tiles and Pixels



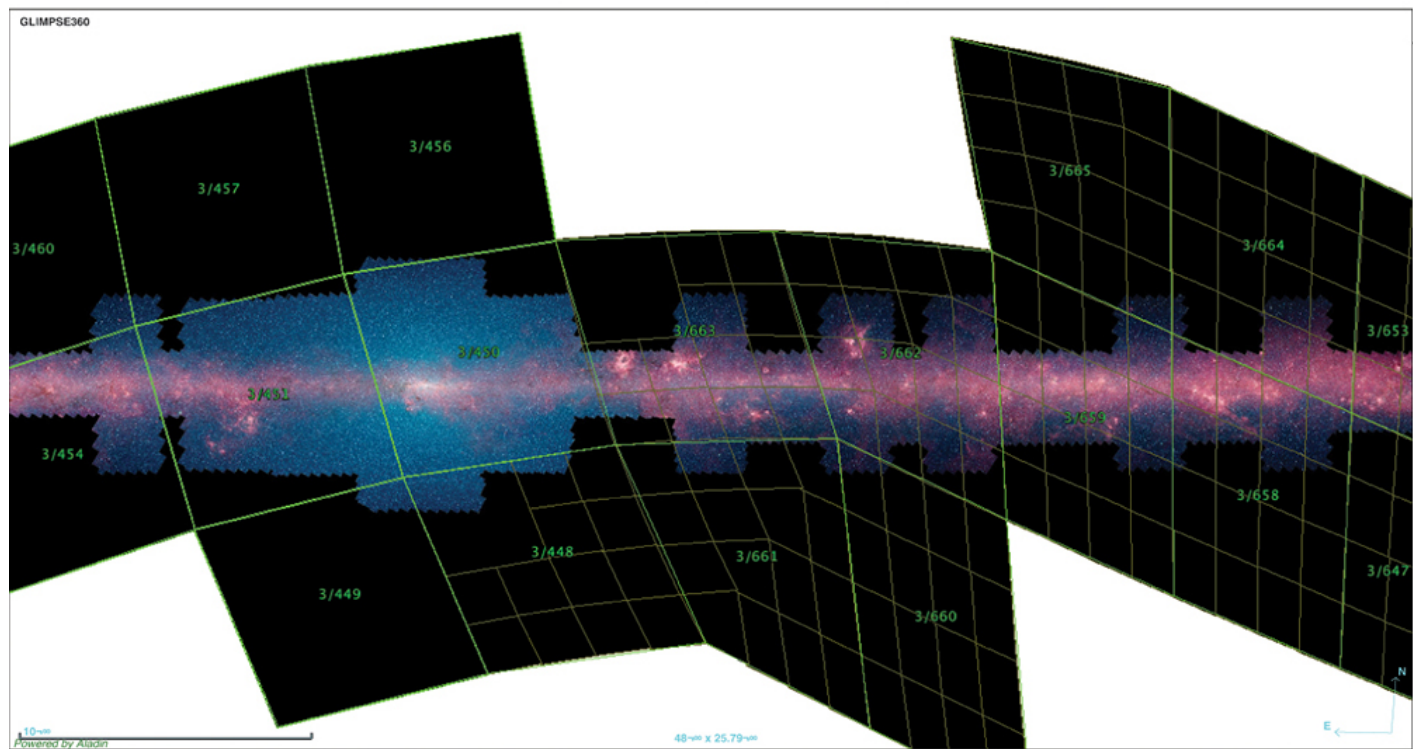
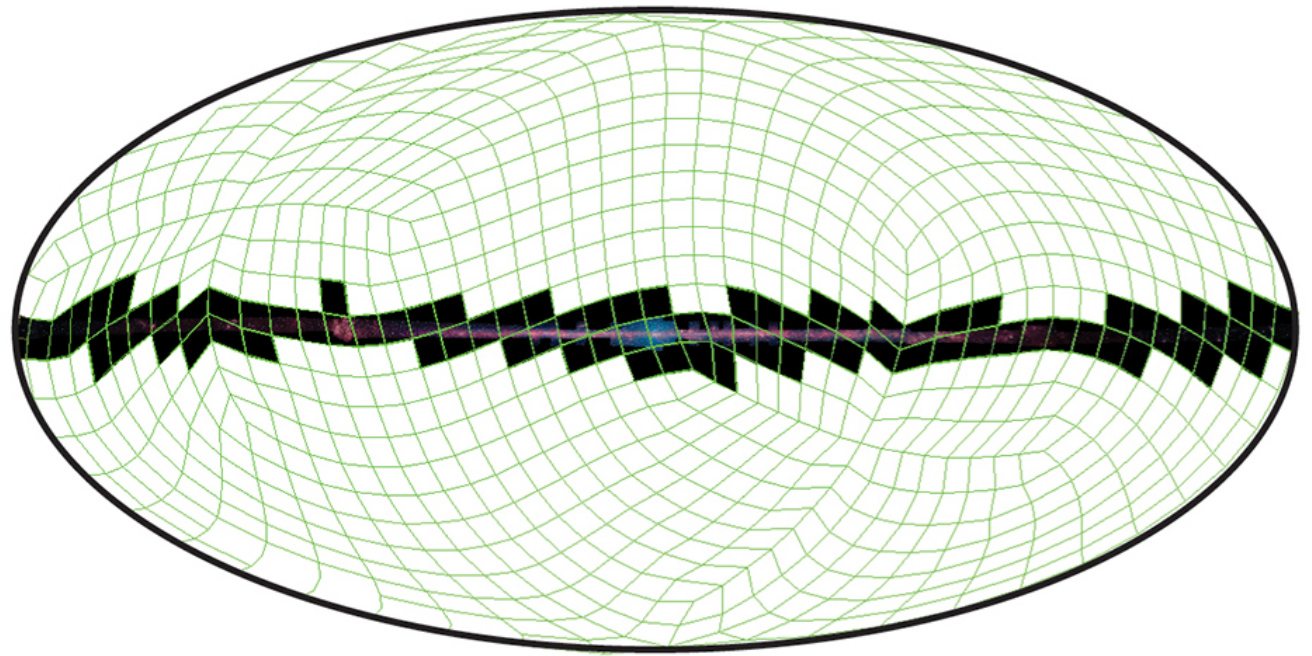
k	$N_{side} = 2^k$	N_{pix}	θ_{pix}	$k_{tile,512}$	$N_{tile,512}$	$\theta_{tile,512}$	
0	1	12	58°6				
1	2	48	29°3				
2	4	192	14°7				
3	8	768	7°33				
4	16	3072	3°66				
5	32	12,288	1°83				
6	64	49,152	55'0				
7	128	196,608	27'5				
8	256	786,432	13'7				
9	512	3,145,728	6'87	0	12	58°6	- WMAP
10	1024	12,582,912	3'44	1	48	29°3	- PLANCK HFI
11	2048	50,331,648	1'72	2	192	14°7	- IRAS
12	4096	201,326,592	51''5	3	768	7°33	
13	8192	805,306,368	25''8	4	3072	3°66	
14	2^{14}	3.22×10^9	12''9	5	12288	1°83	- NVSS
15	2^{15}	1.29×10^{10}	6''44	6	49152	55'0	
16	2^{16}	5.15×10^{10}	3''22	7	196608	27'5	- SCUBA
17	2^{17}	2.06×10^{11}	1''61	8	786432	13'7	
18	2^{18}	8.25×10^{11}	0''81	9	3,145,728	6'87	- DSS
19	2^{19}	3.30×10^{12}	0''40	10	12,582,912	3'44	- SDSS
20	2^{20}	1.32×10^{13}	0''20	11	50,331,648	1'72	
21	2^{21}	5.28×10^{13}	0''10	12	201,326,592	51''5	- CFHTLS
22	2^{22}	2.11×10^{14}	50.3 mas	13	805,306,368	25''8	
23	2^{23}	8.44×10^{14}	25.1 mas	14	3.22×10^9	12''9	- HST ACS
24	2^{24}	3.38×10^{15}	12.6 mas	15	1.29×10^{10}	6''44	
25	2^{25}	1.35×10^{16}	6.30 mas	16	5.15×10^{10}	3''22	



----- Tiles -----

□ e.g.

Spitzer GLIMPSE 360



Aladin v8.1 *** BETA VERSION (based on v8.162) ***

Location Frame ICRS

IRIS colored

14:00 12:00 10:00

-30:00

-30:00

15°

134° x 108.6°

select
pan
zoom
dist
phot
draw
tag
filter

rgb
crop
cont
pixel
prop
del

HST-F606W r3
 IRIS colored
 AIIWISE W1 (3.4u)
 AIIWISE color
 PHAT F475W
 HST-F555W r3

epoch
size
dens.
opac.
zoom

Frame: ICRS

+180 +90 -90 -180

05:38:03.22 -69:14:27.9
134° x 108.6°

Search

0 sel / 0 src 66fps / 288Mb

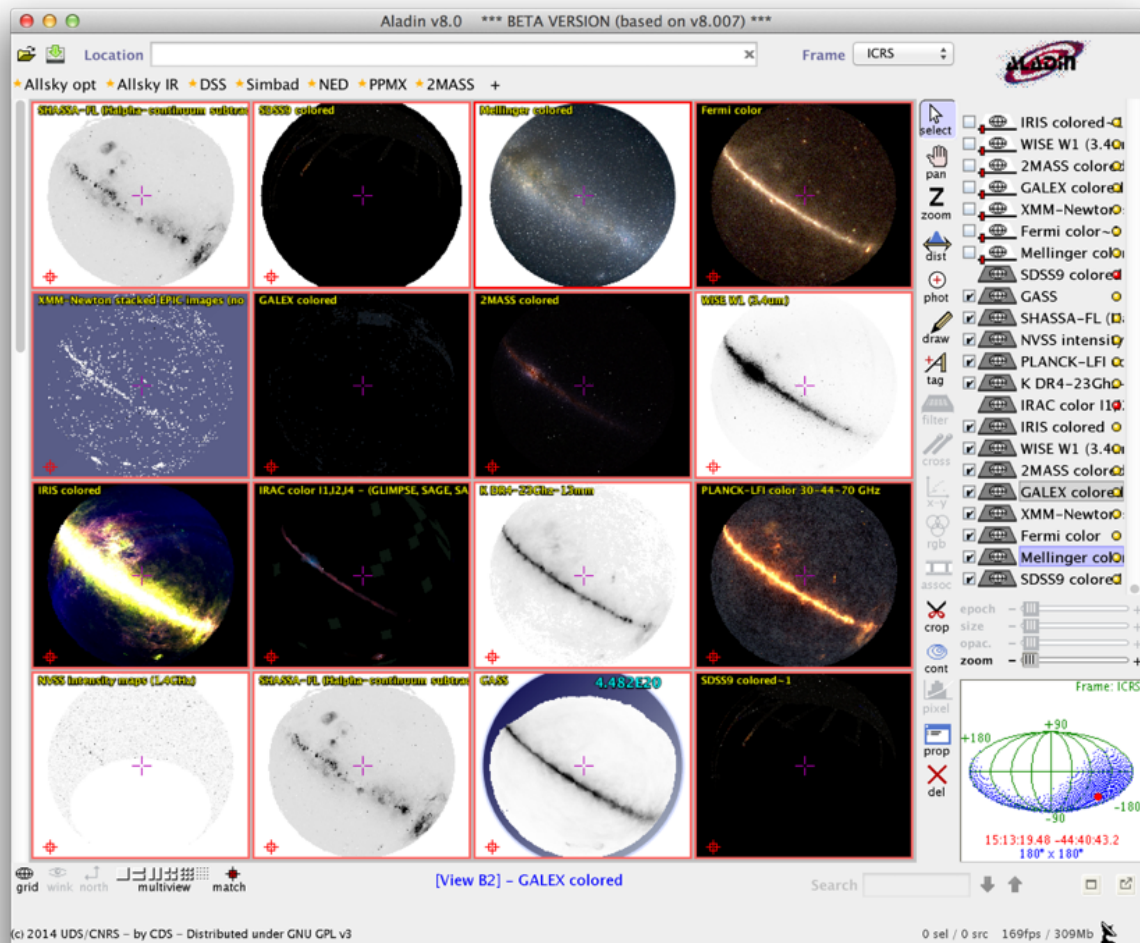
(c) 2015 UDS/CNRS - by CDS - Distributed under GNU GPL v3

Going beyond zoom and pan

- Visualisation and preservation of science data
 - JPEG/PNG and FITS (preserves dynamic range)
 - Link mechanism to original data
 - HEALPix is scientifically robust
- Ease of implementation (generation & publication)
 - No databases or servers, just HTTP
 - Leads to fast take-up and innovation
 - Simple, customisable tools for generating HiPS

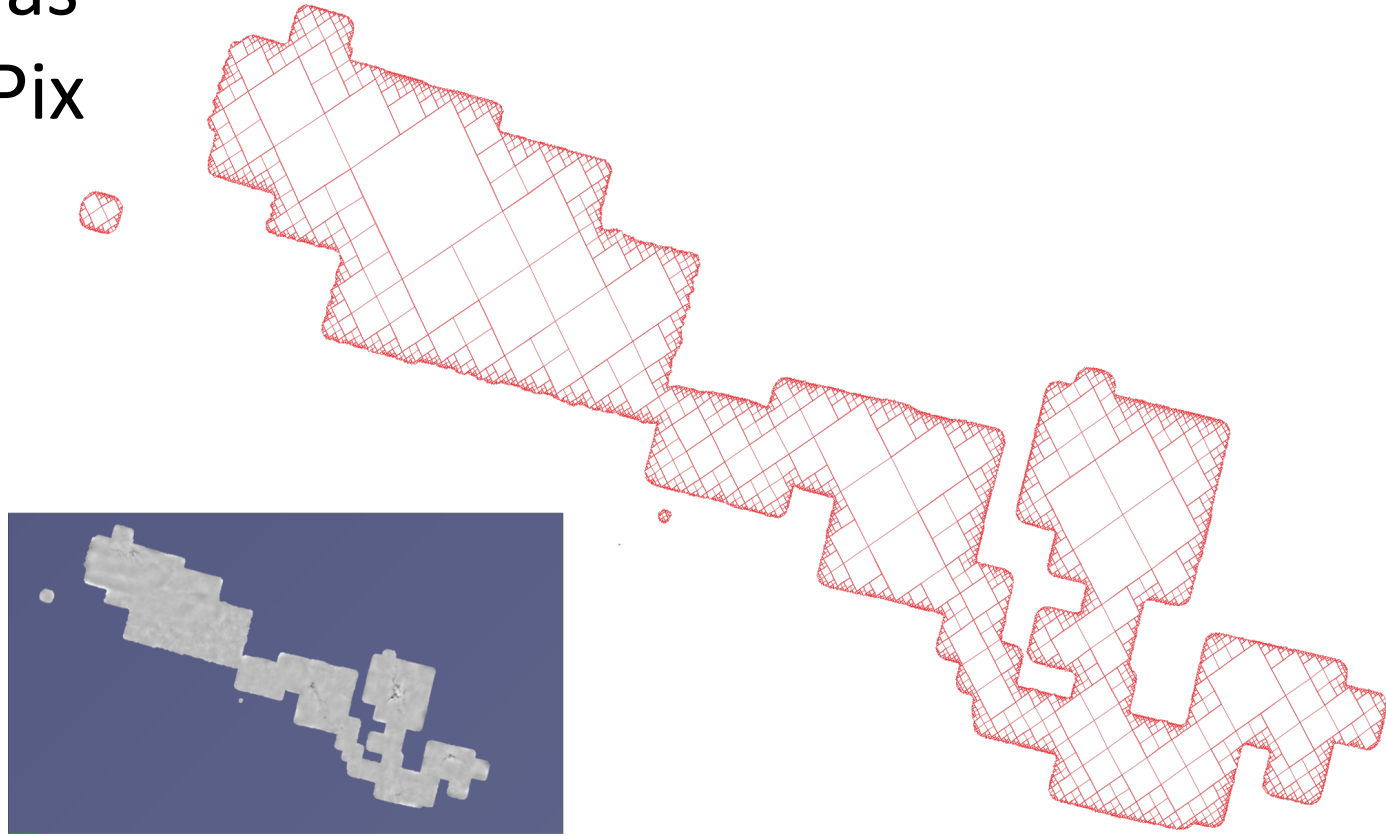
□ HiPS Image Data Sets

- 250 HiPS
- Providers: CDS, CADC, IPAC, ESAC, IRAP/CADE, IAS
- 3 clients: Aladin, Aladin-lite, Mizar (CNES)



☐ Multi-Order Coverage (MOC*) Maps

- Sky regions as list of HEALPix indices
- Multi-order
- Unique
- Fast



* IVOA Recommendation: MOC 1.0



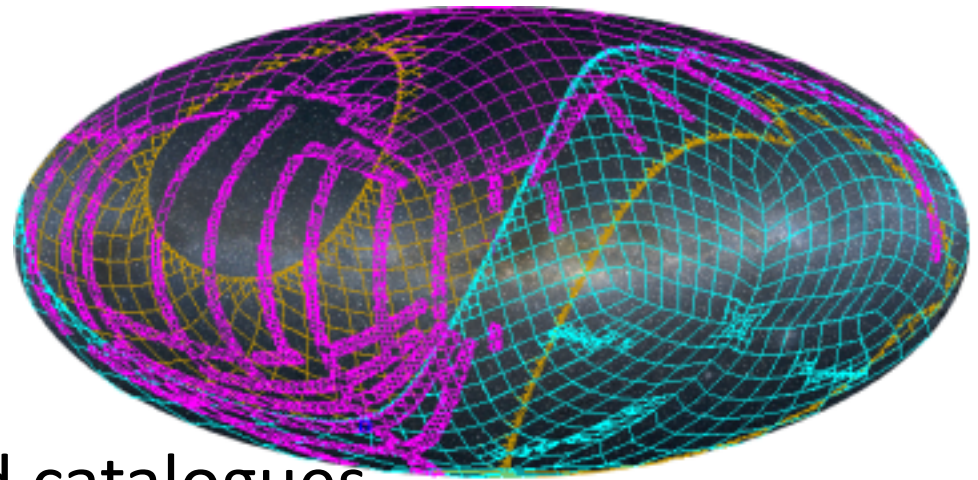
□ HiPS and MOC

Comparing coverage of 100s of large data sets

Intersection/union/complement trivial

Analogue coverage

1000s (Vizier)

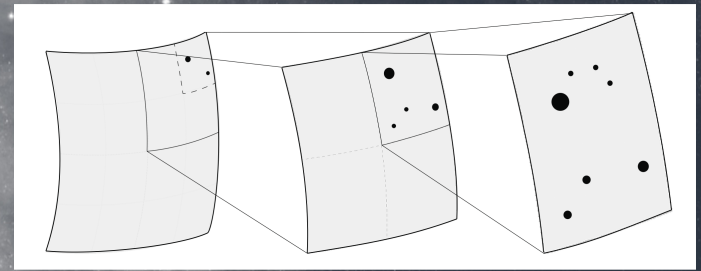


Queries based on coverage and catalogues

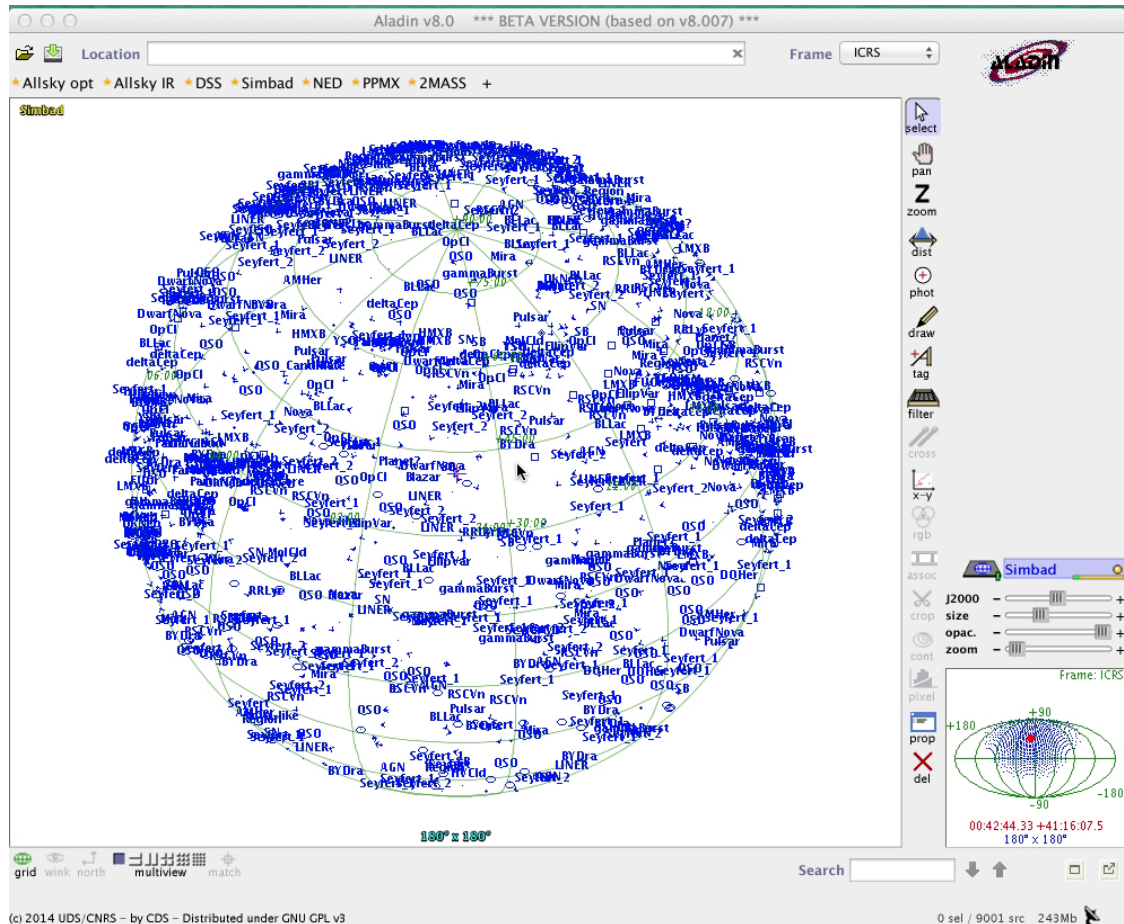
e.g. Veron quasars in HST, XMM and SDSS images



□ HiPS Catalogues



- HiPS ‘container’
- Catalogue rows instead of image pixels
- Hierarchy defined a sort key



HiPS 3-dimensional cubes

Aladin v8.1 *** BETA VERSION (based on v8.162) ***

Location Frame ICRS

DSS SDSS 2MASS WISE GALEX PLANCK AKARI XMM Fermi Simbad NED +

-40026679 m/s

Imagine your eye looking through a stack of planes.

Each plane contains its own data set: image, catalog, graphical overlays...

You see the combination of them.

Use File->Open for discovering all other data, or clic & drag your own files.

Frame: ICRS

01:53:59.77 +23:33:10.9
180° x 150.5°

grid wink north multiview match

Search

(c) 2015 UDS/CNRS - by CDS - Distributed under GNU GPL v3

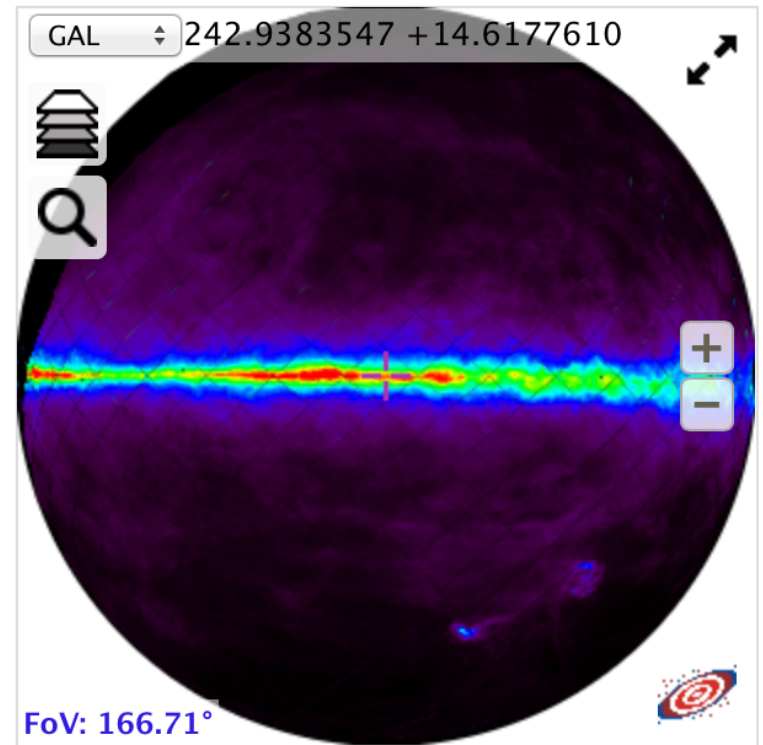
0 sel / 0 src 250fps / 5739Mb

HiPS for Big Data

- All sky at 0.8" (HEALPix level 18)
 - HiPS: 2TB (*16 bit FITS*), 100 GB (*JPEG*)
- All sky at 3" resolution, 1000 bands
 - HiPS: 375 TB (*16 bit FITS*), 19 TB (*JPEG*)
- 18000 deg², 0.3" resolution, each 3 days for 3 years
 - HiPS: ~5 PB (*16 bit FITS*), 256 TB (*JPEG*)
- HiPS access would be feasible today

Aladin Lite

- Aladin running in browser
 - Javascript embeddable widget
 - Customisable
 - Open source GPL3
 - Very easy to implement
- Examples (outside CDS)
 - JAXA
 - Akari Explore Tool
 - GLIMPSE360
 - CADE (*Centre d'Analyse de Données Etendues*)
 - ADS *All-sky Survey*



Summary

- HiPS & MOC solutions for today's Big Data
- Enables new interoperability between images, catalogues and cubes
- Basis for CDS approach to big data
- Simplicity is key for success
 - enables innovation
 - forms a community of HiPS & MOC service providers

Links

- Hierarchical Progressive Surveys
 - *Fernique et al. 2015, A&A 578, 114*
 - HiPS on CDS web pages: <http://aladin.u-strasbg.fr/hips>
- Aladin <http://aladin.u-strasbg.fr/AladinDesktop/>
- Aladin Lite <http://aladin.u-strasbg.fr/AladinLite/>
- HiPS generation tools: <http://aladin.u-strasbg.fr/hips/#tools>

- Fermi
- EGRET
- XMM
- INTEGRAL
- RASS
- GALEX
- DSS
- SDSS
- CFHTLS
- HST
- 2MASS
- ULTRAVISTA
- WISE
- ALLWISE
- DIRBE
- IRIS
- GLIMPSE360
- SPITZER
- AKARI
- SCUBA
- BGPS
- WMAP
- PLANCK
- NVSS
- CHIPASS
- SUMSS
- DWINGELOO
- HASLAM408
- VLSSR
- WENSS
- SHS
- SHASSA
- VTSS
- GASS
- CGPS
- GALFA
- CALIFA
- MUSE
- HARP/JCMT
- Gaia GUMS
- SIMBAD
- GOODS
- PHAT