



HAL
open science

SLICES-FR : l'infrastructure de recherche nationale pour l'expérimentation Cloud et Réseaux du futur

Guillaume Schreiner, Pierre Neyron

► To cite this version:

Guillaume Schreiner, Pierre Neyron. SLICES-FR : l'infrastructure de recherche nationale pour l'expérimentation Cloud et Réseaux du futur. JRES (Journées réseaux de l'enseignement et de la recherche) 2024, Renater, Dec 2024, Rennes, France. hal-04893845

HAL Id: hal-04893845

<https://hal.science/hal-04893845v1>

Submitted on 17 Jan 2025

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 - NonCommercial 4.0 International License

SLICES-FR : l'infrastructure de recherche nationale pour l'expérimentation Cloud et Réseaux du futur

Guillaume Schreiner

ICube - UMR CNRS 7357

300 Bd S. Brant
67412 Illkirch Cedex

Pierre Neyron

LIG - UMR 5217

CS 40700
38058 Grenoble Cedex 9

Résumé

En s'appuyant sur l'expérience acquise avec les plateformes Grid'5000 [1](2003) et FIT [2](2011), l'infrastructure de recherche SLICES-FR [3] démarrée en 2022 regroupe une quinzaine d'établissements de l'ESR. SLICES-FR, financée par les PEPR Cloud et Réseaux du Futur [4][5][6], a pour ambition d'unifier l'accès aux services d'expérimentation dans les domaines des protocoles réseaux, des technologies radio, des services et de la collecte de données ainsi que du calcul parallèle et distribué. Au-delà du paysage français, SLICES-FR s'interface à l'échelle européenne avec SLICES-RI [7] qui est la première infrastructure labellisée ESFRI [8] dédiée à la recherche en informatique, pour une durée d'exploitation prévue jusqu'en 2040.

SLICES-FR exploite une infrastructure à grande échelle répartie sur plus d'une dizaine de sites en France permettant l'accès contrôlé et uniforme à des technologies de pointe hétérogènes et variées. Elle répond à la nécessité de soutenir la recherche expérimentale dans ces domaines (Cloud, HPC, Edge, IoT, SDR, Beyond 5G) où l'accès à de tels instruments est essentiel. En plus de fournir un accès distant aux ressources matérielles, SLICES-FR offre une maîtrise profonde du matériel expérimental, grâce à des services de reconfiguration jusqu'au plus bas niveau (baremetal) et tout une suite d'outils de traitement des résultats, notamment la consommation énergétique, pour permettre l'évaluation complète d'une application numérique de bout en bout.

A travers notre article, nous présenterons l'architecture envisagée pour construire cette infrastructure distribuée en France et en Europe, ainsi que les solutions réseaux pour interconnecter les ressources entre-elles.

Mots-clefs

Infrastructure, Recherche, Expérimentation, Europe, Cloud, HPC, Réseaux, IoT, 5G

1 Introduction

1.1 L'infrastructure européenne SLICES-RI

Depuis l'avènement du concept de l'informatique dans le nuage, les domaines de recherche en informatique des réseaux et des systèmes distribués ont convergé pour créer de nouveaux usages numériques. La forte démocratisation des terminaux mobiles ainsi que la montée en puissance des canaux de communication ont contribué à rendre possible l'accès à la donnée depuis n'importe où et instantanément. Cette donnée se concentre principalement dans les datacenters qui vont permettre son stockage, son traitement et sa diffusion à grande échelle. Le but de SLICES-RI est de fournir aux chercheurs l'infrastructure numérique permettant de réaliser une expérimentation à large échelle de bout en bout mêlant à la fois des équipements dédiés aux systèmes et aux réseaux afin de tester les performances de nouveaux protocoles de communication, d'algorithmes de traitement et de placement de la donnée.

Jusqu'à présent, les grandes infrastructures européennes d'expérimentation du numérique étaient au service d'autres domaines de recherche comme les sciences fondamentales ou humaines. SLICES-RI est la première infrastructure de recherche européenne labellisée ESFRI dédiée à la recherche en informatique dans les domaines des réseaux et systèmes distribués.

A ce titre, SLICES-RI doit remplir de nombreux critères scientifiques, techniques et organisationnels pour se structurer selon les critères ESFRI en un ERIC. La structuration de l'ERIC prévue pour 2025 comporte différents conseils pour le pilotage et la direction, autour desquels gravitent différents comités techniques, opérationnels et exécutifs. SLICES-RI se compose au total de 16 pays ou nœuds hébergeants 25 partenaires (instituts, centre de recherche, universités) distribués sur différents sites géographiques. A l'image d'un grand radio télescope constitué de milliers d'antennes réparties sur plusieurs km², SLICES-RI vise à devenir l'instrument de référence à l'échelle européenne. Chaque nœud partage avec l'ensemble des partenaires des équipements ou ressources en les exposant via des services et en les connectant à travers les réseaux européens de recherche (GEANT [9], NREN). Un nœud central (Central Hub) interconnecte techniquement tous les pays membres (Nodes). Dans le cas particulier de SLICES-RI, le Central Hub sera hébergé en France. Du point de vue utilisateur, le Central Hub est le point d'entrée privilégié pour s'enregistrer, accéder aux ressources et au support. Enfin, une charte d'utilisation commune à tous les nœuds garantit une politique d'accès cohérente et homogène.

1.2 Le nœud français SLICES-FR

1.2.1 Les utilisateurs

Nous distinguons trois grandes catégories d'utilisateurs qui accèdent à nos infrastructures :

- Les chercheurs issus du monde académique (chercheurs, enseignant-chercheurs, ingénieurs de recherche, doctorants, etc.) ;
- Les étudiants en formation initiale ;
- Les industriels (startup, PME, multi-nationale, etc.).

Nous envisageons également différents modèles d'accès :

- Excellence Driven : un chercheur dépose un dossier de candidature pour obtenir un quota de ressources ;
- Market Driven : sur un modèle de prestation, un industriel paye pour accéder à un quota de ressources ;
- Wide Access ou Best effort : les utilisateurs académiques accèdent gratuitement aux ressources non utilisées à un instant donné.

1.2.2 Historique des plateformes précurseuses du domaines (Grid'5000, Equipex FIT)

Lancée en 2003, Grid'5000 est une infrastructure de recherche dédiée aux expériences en informatique distribuée avec un accent particulier sur le calcul haute performance (HPC), le cloud computing, et les systèmes distribués. Grid'5000 est répartie sur plusieurs sites en France dans le but d'autoriser des expérimentations à grande échelle. Les utilisateurs peuvent réserver des nœuds de calcul via un système de gestion de tâches et configurer entièrement le logiciel sur ces nœuds (OS, bibliothèques, applications). Grid'5000 offre un contrôle complet aux chercheurs pour leurs expérimentations et permet de recréer des environnements spécifiques afin d'expérimenter des scénarios complexes dans des conditions maîtrisées. Ainsi, l'infrastructure entièrement isolée (in vitro) autorise des conditions reproductibles avec des mesures de performances exemptes d'interférences non interprétables. Grid'5000 est un instrument essentiel pour la recherche en informatique, permettant de réaliser des expériences réalistes et à grande échelle dans des conditions contrôlées, notamment en fournissant des serveurs et des réseaux dédiés par expérimentation.

FIT est un projet français démarré en 2011 dans le cadre du programme Equipex (Équipements d'excellence). FIT vise à fournir une infrastructure expérimentale dédiée à la recherche sur les réseaux sans fil, les objets connectés (IoT) et les réseaux cellulaires. Ses composantes principales sont FIT IoT-LAB [10] (Internet des objets), FIT Wireless R2LAB [11] (réseaux cellulaires) et CorteXlab [12] (radio cognitive), FIT Cloud (OpenStack). FIT est déployé dans plusieurs villes en France, offrant une infrastructure géographiquement distribuée pour tester des expérimentations multi-sites et des topologies de déploiement radio complémentaires. Les types d'expérimentations menés sont multiples : déploiement de nouveaux protocoles réseaux et tests de leur performance ; expérimentations d'objets connectés dans des scénarios réalistes (smart cities, smart homes, agriculture, etc.) ; développement et évaluation de nouveaux services et applications pour l'IoT et le Cloud.

1.2.3 De SILECS à SLICES-FR

Dans un but scientifique et technique, les deux infrastructures Grid'5000 et FIT ont fusionné en 2018 pour créer l'Infrastructure de Recherche nationale SILECS [13]. Malheureusement, la période Covid de 2020 a considérablement freiné ce rapprochement. Le fonctionnement des deux infrastructures précurseuses est resté en silo, ce qui nécessite toujours la création de comptes distincts. De plus, la séparation au niveau réseau entre les infrastructures complique énormément les communications de bout en bout entre tous les équipements. Finalement, le projet européen SLICES-RI lancé en 2021 a réorienté les objectifs de notre infrastructure nationale SILECS. En 2022, SILECS se transforme en SLICES-FR, le nœud français de SLICES-RI, pour directement se

structurer à l'échelle européenne. En tant que nœud de SLICES-RI, SLICES-FR n'est pas une infrastructure nationale complètement autonome, elle doit s'aligner sur les standards ESFRI qui définissent les grandes infrastructures européennes distribuées. Dans la suite de l'article, nous allons présenter les objectifs et les travaux en cours pour construire cette nouvelle infrastructure nationale et européenne.

2 Architecture SLICES-FR

Dans cette section, nous allons présenter les ressources expérimentales hébergées ainsi que les services d'infrastructures qui les pilotent. La Figure 1 résume l'architecture générale de SLICES-RI distribuée sur plusieurs sites et plusieurs pays (node). Pour chaque type d'équipement proposé, un service d'infrastructure pilote sa configuration et autorise sa prise en main à distance.

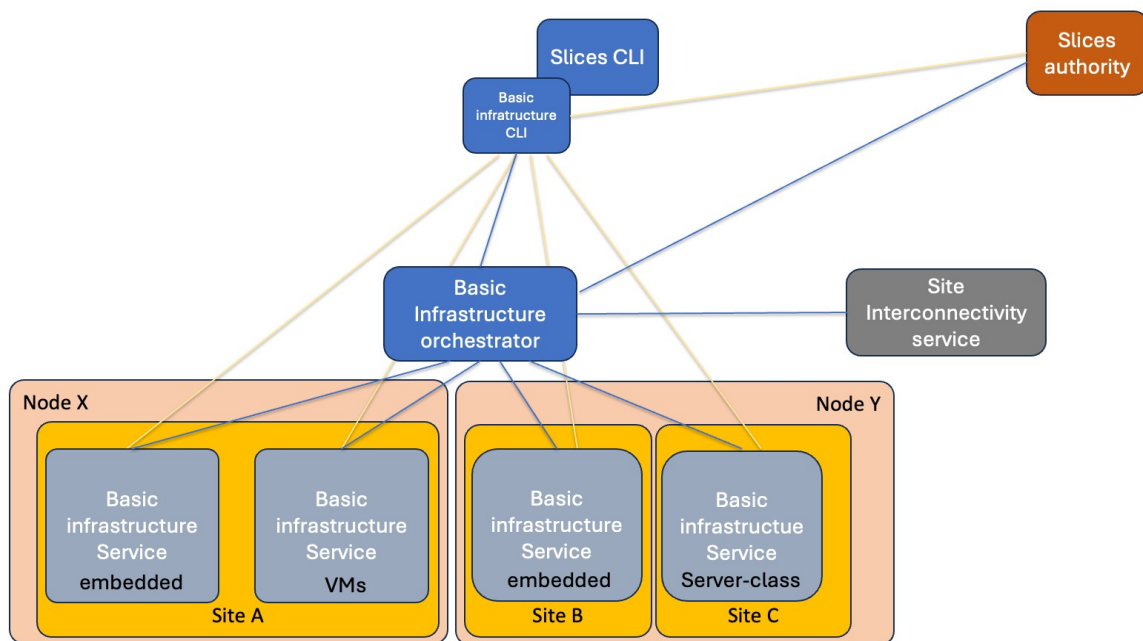


Figure 1: Architecture distribuée SLICES-RI

2.1 Équipements matériels réservables

L'un des objectifs de SLICES-FR est de proposer des équipements réservables et utilisables à distance. Dans le domaine de l'expérimentation des réseaux et systèmes distribués, SLICES-FR a identifié 7 types d'équipements matériels répartis dans 3 catégories distinctes. Ce découpage a plusieurs buts :

- Rendre les équipements matériels agnostiques aux domaines de recherche ;
- Donner une cartographie précise des équipements aux utilisateurs ;
- Rationaliser au maximum les implémentations des services liés aux équipements du même type ou de la même catégorie ;
- Identifier le coût d'intégration des types d'équipements par l'équipe technique ;

- Comprendre le(s) type(s) d'équipement(s) nécessaire(s) afin de mener une expérimentation de bout en bout pour un domaine de recherche.

2.1.1 Systèmes ouverts

La catégorie **open systems** offre à l'utilisateur la possibilité d'installer le système d'exploitation de son choix sur un équipement. Au sein de cette catégorie, on distingue quatre types d'équipements.

- **Server-class machine** : il s'agit d'un serveur type rack hébergé dans une salle machine ou un datacenter. Par défaut, on considère que le serveur possède une interface de prise en main à distance facilitant l'installation de l'OS par nos services d'infrastructures. Ce sont des machines puissantes (CPU, RAM) offrant une variété d'architectures (x86-64, ARM, Power, RISC-V) et d'extensions (GPU, stockage, réseau). Exemples : Dell (PowerEdge), Hewlett Packard Enterprise (ProLiant servers), IBM, SuperMicro, etc.
- **Small form factor machines** : ce type de machine correspond à un mini-ordinateur capable de faire tourner un système d'exploitation type GNU/Linux et suffisamment puissant pour faire un traitement local pour des applications légères d'intelligence artificielle ou machine learning. Son déploiement peut se faire en dehors des salles machines. Exemples : Nvidia Jetson, Intel NUC, Raspberry Pi, BeagleBone, etc.
- **Embedded devices** : cet équipement correspond à une carte de développement composée généralement d'un microcontrôleur, d'une puce communicante et de divers capteurs/actionneurs. Ses performances matérielles limitées contraignent l'utilisation de systèmes embarqués (Arduino, FreeRTOS, Zephyr, RIOT). Exemples : cartes Arduino, STM32, ESP32, Raspberry Pi Pico, etc.
- **Network Devices** : il s'agit d'un équipement réseau type commutateur/routeur ou point d'accès Wi-Fi disposant de multiples interfaces réseaux. L'OS de cet équipement est entièrement reprogrammable (SONiC, OpenWRT). Exemples : commutateurs WhiteBox, commutateurs compatibles P4, etc.

2.1.2 Systèmes configurables

Contrairement à la catégorie open systems, la catégorie **configurable systems** propose des équipements avec un OS non programmable, souvent propriétaire. Toutefois, l'utilisateur peut modifier une partie ou la totalité de la configuration de l'équipement afin d'adapter les paramètres systèmes, réseaux ou applicatifs aux objectifs visés de l'expérimentation. Au sein de cette catégorie, nous avons identifiés deux types d'équipements.

- **Network appliances** : ces équipements sont spécialisés pour une fonction réseau particulière (routage, commutation, réseau sans fil) et proposent un OS non reprogrammable avec des paramètres par défaut que l'utilisateur peut modifier selon son cas d'usage. Exemples : routeurs Cisco ou Juniper, HPC fabric, LoRaWAN Gateways, etc.
- **Software-Defined Radio (SDR) devices** : ces équipements sont spécialisés dans les communications sans fil et permettent de construire logiciellement la couche physique afin d'émuler un standard radio particulier. Exemples : USRP, cartes compatibles GNU Radio, etc.

2.1.3 Systèmes fermés

Cette dernière catégorie propose d'ajouter à l'infrastructure des équipements non programmables et limités au niveau de leur configuration. Nous avons identifié un type d'équipement :

- **Closed systems** : il s'agit d'équipements propriétaires dont la reprogrammation est impossible et dont le fabricant limite fortement ou totalement sa configuration. L'intérêt est par exemple d'évaluer les performances d'un système propriétaire au sein de l'infrastructure ou de jouer des tests d'interopérabilité. Exemples : objets connectés du commerce.

2.2 Équipements virtuels réservables

Au delà des équipements matériels, nos utilisateurs ont exprimés le besoin d'accéder à des ressources virtuels de types machines virtuels (IAAS) et conteneurs (PAAS). En effet, dans certains cas d'usage (NFV, SDN, 5G Core), l'utilisation d'un serveur dédié n'est pas forcément optimale pour déployer rapidement une fonction, ou un service qui sera une des briques de l'expérimentation. Pour répondre aux exigences de reproductibilité et accéder au monitoring de ces infrastructures cloud dédiées, SLICES-FR doit être en mesure d'opérer en interne ces ressources virtuelles.

2.3 Services d'expérimentations

SLICES-FR vise à moyen terme l'unification technique des différentes infrastructures nationales du domaine. Cette structuration technique est orientée essentiellement par le cahier des charges de SLICES-RI qui décrit les services de bases nécessaires pour gérer l'infrastructure de recherche européenne. En tant que nœud de SLICES-RI, SLICES-FR s'aligne naturellement sur les services européens. La Figure 2 présente la vision modulaire et verticale de l'implémentation de chaque service. Un pilote gère la ressource locale à bas niveau et expose un service accessible via une API. Pour chaque service, une librairie cliente est adossée jusqu'à une interface web haut niveau. Cette section présente 4 catégories découpées en 12 services qui pilotent les différents types d'équipements de façon homogène et transparente pour l'utilisateur.

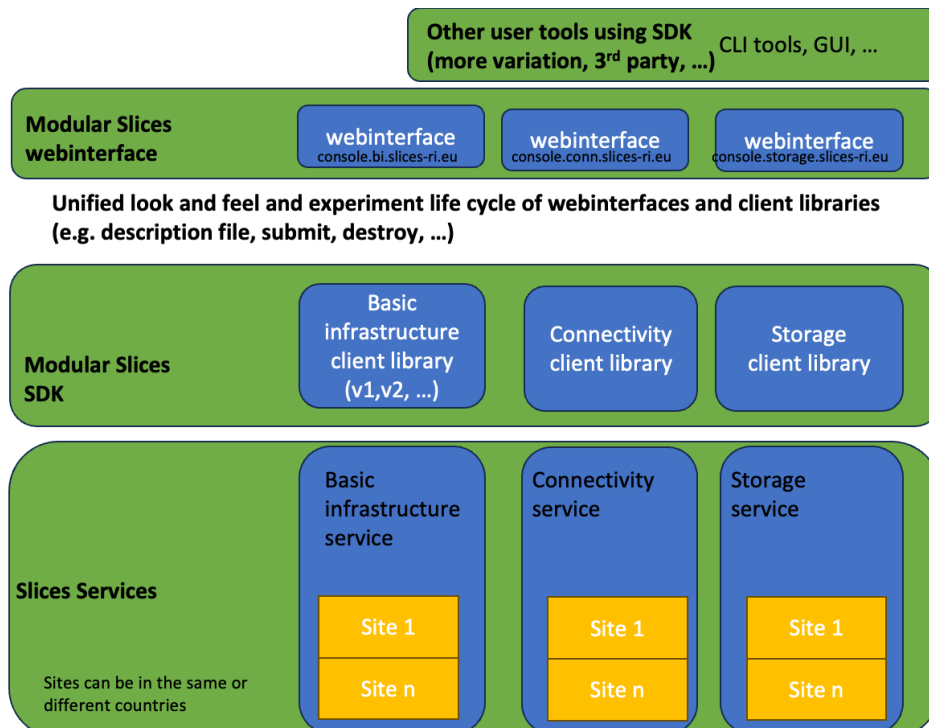


Figure 2: Architecture modulaire des services SLICES-RI

- **User and platform management services :** cette catégorie englobe la gestion des utilisateurs au sens large (création, référentiel), la documentation orientée utilisateur, le support aux utilisateurs et les statistiques d’usage. Le système d’authentification se base sur la fédération d’identité Edugain [14] mais souhaite garder la possibilité d’enregistrer des comptes externes non académiques (ex : industriels). Le référentiel européen de SLICES-RI est hébergé sur le Central Hub et permet l’accès aux ressources partagées de tous les pays pour tous les utilisateurs enregistrés.
- **Resources management services :** les services de cette catégories sont directement liés au pilotage des équipements. Dans un premier temps, les utilisateurs consultent un service de découverte pour lister les équipements de différents types disponibles. Un service de réservation est ensuite appelé afin d’associer les équipements choisis à une expérimentation. Au lancement de l’expérimentation, un service de configuration va déployer les environnements ou les paramètres sur les équipements cibles. Enfin, un service de monitoring récolte pendant toute la durée de l’expérimentation les métriques de performance liées aux équipements.
- **Data oriented services :** cette catégorie vise à couvrir la problématique des services de stockage, de transfert de données (interne entre sites ou externe en dehors de l’infrastructure) et de métadonnées associées aux résultats des expérimentations. Une exigence particulière est l’application des principes FAIR ainsi qu’un référencement sur EOSC (European Open Science Cloud). En dernier lieu, un service d’analyse permet un traitement automatisé des données pour corréler les résultats obtenus entre différentes expérimentations.

- **Experiment management services** : des services de management et d'orchestration d'expérimentations permettront de scripter des scénarios évolués liés aux ressources matérielles. Un service de tableau de bord utilisateur permet de suivre le bon déroulé en temps réel et de consulter les historiques des expérimentations.

2.4 Services managés

Historiquement, les ressources mises à disposition dans les infrastructures du domaine proposent par défaut des expérimentations d'une durée limitée à quelques heures en journée afin de privilégier un mode interactif propice au prototypage. Pour les campagnes d'expérimentations bien rodées, les utilisateurs automatisent les lancements pendant la nuit ou les week-ends en dehors des horaires de bureau. Certaines expérimentations nécessitent un service tiers qui permet l'enregistrement d'équipements sur le réseau. Ces services peuvent être compliqués à configurer et à provisionner à chaque lancement. Pour les utilisateurs se concentrant uniquement sur un aspect particulier de la chaîne de communication, SLICES-FR souhaite proposer des services persistants qui facilitent et accélèrent le déploiement d'une infrastructure réseau complète (ex : LoRaWAN Network Server, 5G Core).

2.5 Connectivité réseau

Le **connectivity service** est en charge de relier tous les équipements matériels et virtuels. Hiérarchiquement, l'architecture réseau de SLICES-RI se découpe en 3 niveaux :

- le **niveau local** qui correspond à un site géographique hébergeant des ressources ;
- le **niveau national** ou nœud (ex. SLICES-FR) au sein duquel les différents sites géographiques nationaux sont interconnectés par un même NREN (ex. RENATER [15]) ;
- le **niveau européen** qui assure l'interconnexion entre les nœuds, les sites géographiques européens sont interconnectés nationalement par leur NRENs respectifs et internationalement via GEANT.

2.5.1 Connectivité L2

En fonction du cas d'usage expérimenté, les prérequis de connectivité sont plus ou moins contraignants selon les équipements qui communiquent avec des protocoles réseaux au niveau L2 de la couche OSI (Ethernet) et au niveau L3 de la couche OSI (Routage). Afin de pouvoir isoler une expérimentation pour des services communicant au niveau L2, SLICES-FR recourt à l'utilisation de VLANs au niveau local. Au niveau national, SLICES-FR se repose sur les services L2 VPN de RENATER pour interconnecter les sites au niveau L2. Dans ce cas, le site géographique doit être en mesure de se connecter directement au routeur RENATER, éventuellement par l'intermédiaire de l'opérateur local du campus. Au niveau européen, nous utilisons les services L2 VPN de GEANT qui opère les liens entre les NRENs. Dans un premier temps, SLICES-RI proposera des liens L2 permanents ou statiques entre les différents sites européens.

Pour passer à l'échelle en nombre de liens L2 dynamiquement créés entre les sites, deux pistes sont encore étudiées. La première solution repose sur le protocole QinQ permettant d'encapsuler des VLANs entre deux sites mais son déploiement risque d'être contraignant car les équipements

réseaux intermédiaires doivent implémenter son support. La solution alternative basée sur VxLAN (L2 over L3) semble être plus souple à déployer (transport au dessus de IP, support sur de nombreux équipements réseaux ou OS) mais pose d'autres problèmes comme la performance en terme de débits ou de latence.

2.5.2 Connectivité L3

SLICES-FR propose une connectivité L3 compatible IPv4 et IPv6. Un plan d'adressage statique est proposé par défaut. Pour les topologies réseaux avancés, un service de déploiement de sous-réseaux IPv4 et IPv6 permettra de dynamiquement configurer un lien L3.

2.5.3 In Vitro VS In Vivo

Dans une démarche d'expérimentation reproductible, l'environnement réseau doit être entièrement contrôlé. Afin de faciliter l'interprétation des résultats liés à la connectivité réseau, SLICES-FR est déployé sur des équipements réseaux dédiés. Au niveau du site local, nous avons un contrôle total du réseau. Concernant l'interconnexion des ressources distribuées sur plusieurs sites, nous déployons entre certains sites des liens dédiés au trafic expérimental. Par défaut, l'interconnexion des sites est mutualisée avec le trafic du campus hébergeant le site.

D'autres cas d'usage liés à la métrologie sur Internet nécessitent d'exposer des ressources avec une adresse IP publique. Dans ce cas, l'expérimentation risque potentiellement d'être perturbée par des requêtes extérieures aléatoires et non contrôlées.

2.5.4 Débits et latence

SLICES-FR vise à fournir à chaque site un débit maximum de 100 Gbit/s mutualisé via les liens de campus opérés par RENATER. En interne sur un site, les ressources de type server-class machine ou network devices proposent des interfaces réseaux avec une vitesse allant de 10 Gbit/s à 100 Gbit/s. Pour des cas d'usage comme le déploiement d'infrastructure 5G multi-sites, une faible latence est nécessaire pour relier les antennes au coeur de réseau. Dans ce cas, l'usage d'un lien dédié avec débit et latence garanti est privilégié.

3 Premières implémentations

3.1 Blueprint 5G

Une démarche originale du projet européen SLICES-RI consiste à partir de cas d'usages réels des chercheurs et de les implémenter manuellement avec tous les types d'équipements et de services nécessaires. Une procédure est rédigée puis partagée avec d'autres partenaires pour reproduire à l'identique l'expérimentation. Après plusieurs itérations satisfaisantes, le blueprint permet de valider les fonctionnalités à ajouter au cahier des charges de l'infrastructure. Cette approche est complémentaire de la phase de spécifications haut-niveau et permet vu l'échelle de temps de projet de valider l'orientation technique en livrant régulièrement des prototypes intermédiaires qui convergent progressivement vers l'implémentation finale.

Le blueprint 5G¹ vise à mettre œuvre différents équipements de télécoms, routeurs, et appliances tournant sur des serveurs Kubernetes. Différentes solutions d'interconnexion réseau permettent de lancer une expérimentation distribuée sur plusieurs sites partageant ainsi les équipements distants.

Suite à ces premiers retours d'expérience positifs, SLICES-RI propose une version pré-opérationnel (Pre-Op) de l'infrastructure cible avec des services avancés d'expérimentation : portail web, authentification centralisée, différents types d'équipements réservables, stockage des données, gestion des métadonnées.

3.2 Production : Focus sur l'intégration d'équipements de type Edge-Computing

Conjointement au développement des services de la nouvelle infrastructure SLICES-FR, la recherche avance et nécessite l'instrumentation des dernières technologies de pointe pour valider expérimentalement les thèses ou même simplement observer en amont les nouveaux objets technologiques. Ainsi va des travaux de recherche sur le Edge-Computing (identifié en partie par le type small form factor machines décrit plus haut) dont le matériel et les services nécessaires à l'expérimentation manquaient.

De plus, tout comme pour les matériels déjà disponibles dans les plateformes historiques, le besoin des chercheurs est en particulier une maîtrise réelle du matériel et la possibilité d'expériences sur un grand nombre d'équipements. S'ajoute à cela les demandes de rationalisation des coût sur les plateformes, et donc la nécessité de mutualiser des efforts d'ingénierie et d'ouvrir l'accès à une très large communauté. SLICES-FR est donc une cible idéale pour ces objectifs.

3.2.1 Quel type d'équipement, quelle plateforme ?

On peut décrire les équipements Edge-Computing au regard d'une ou plusieurs des caractéristiques suivantes : (a) un positionnement au plus près de la production des données (capteurs) ; (b) une forte autonomie impliquant des composants faiblement consommateur d'énergie ; (c) une puissance de calcul cependant suffisante (typiquement avec un GPU) ; (d) une compatibilité avec les contraintes de l'environnement cibles (sécurisation, esthétique) ; (e) une interconnexion réseau variable. Considérant les caractéristiques (a), (c) et (e) comme émulables, nous nous sommes intéressés au matériel Nvidia Jetson que nous avons vu comme typique du Edge-Computing, tout en cherchant un format matériel compatible avec une instrumentation massive dans une plateforme d'expérimentation.

De fait, SLICES-FR pose ses fondations sur plusieurs plateformes historiques éventuellement capables d'héberger des Jetsons, mais toutes nécessitant un travail conséquent du fait de la singularité de ces derniers à la frontière entre un serveur et un « dispositif embarqué ». Une caractéristique fondamentale étant l'OS Linux, c'est dans Grid'5000 qui apporte le support de configuration de plus complet pour ce type de système d'exploitation, que l'étude puis l'instrumentation effective a été réalisée.

¹ https://doc.slices-sc.eu/blueprint/5g_blueprint.html

3.2.2 Contraintes et choix du matériel instrumenté

Quelle que soit la plateforme d'expérimentation « reconfigurable » cible, mais pour Grid'5000 en particulier, les prérequis pour l'instrumentation d'une machine sont les suivants :

1. Compatibilité de l'appareil avec un système GNU/Linux suffisamment ouvert ;
2. Capacité de piloter par le réseau le choix de l'OS à démarrer (booter) ;
3. Disponibilité d'un contrôle indépendant de l'OS pour le redémarrage matériel ;
4. Accès console « série » au système (indépendamment du réseau) ;
5. Compatibilité avec l'hébergement (a priori en datacenter pour Grid'5000).

S'ajoute également le souhait de pouvoir reconfigurer l'interconnexion réseau et mesurer la consommation énergétique.

Après un premier prototype basé sur les kits de développement et un assemblage « artisanal » en cluster malheureusement non abouti en 2018 (Jetson Tegra X1, pas assez fiables), nous nous sommes tournés vers une solution intégrant directement 12 Jetsons AGX Xavier dans un châssis au format datacenter (rack 1U) et disposant des caractéristiques requises (1 à 5) attendues.

Cette solution à priori clé en main a cependant montré des limites à cause de dysfonctionnements inattendus certes finalement corrigés ou contournés grâce notamment à une collaboration proche avec le fournisseur. Mais malheureusement elle ne permet pas d'accéder à une mesure de consommation de chaque Jetson. Ceci encourage donc à reconsidérer dans l'avenir des solutions matérielles plus « artisanales » ou faites « maison », mais offrant un plus grand potentiel d'instrumentation.

3.2.3 Les challenges de l'implémentation logicielle de l'instrumentation

Somme toute, les Jetsons ne sont pas très loin d'une machine « Linux » classique et même d'un serveur. L'architecture ARM64 n'est pas un soucis. La disponibilité d'un contrôle hors-bande et d'une console série non plus, car même si non disponible d'emblée, l'utilisation par exemple de microcontrôleurs (Arduino) pour piloter des kits de développement est facile. Et finalement notre châssis « clé en main » fournis d'emblée ces services.

Lors de nos premiers essais nous nous sommes par contre confronté à un problème épineux pour le contrôle du boot par le réseau permettant aux chercheurs de choisir/maîtriser jusqu'au noyau du système. Effectivement dans un premier temps les modules ciblés (Jetson dernière génération, AGX Xavier) ne disposaient pas de la fonctionnalité PXE pourtant omniprésente de nos jours. Bon an, mal an, notre désespoir de voir arriver un support fonctionnel de PXE nous conduit à agrémenter la panoplie de méthodes d'instrumentation de kadeploy [16] (logiciel qui gère la reconfiguration d'OS) avec un pilotage par U-boot (bootloader classique dans le monde embarqué) et même par un boot sur clé USB programmable à distance.

Un deuxième challenge a été de rendre compatible la surcouche OS de Nvidia (Linux4Tegra [17]) avec l'environnement Grid'5000. En effet, Grid'5000 repose sur un environnement Debian équipé d'un outil de validation (g5k-check [18]) garant de l'intégrité de la plateforme et donc de sa qualité comme instrument scientifique. Il a donc été impératif de réussir à fusionner cette surcouche Nvidia

avec l'OS Debian spécialisé pour Grid'5000 (en particulier les noyaux Linux provenant de l'une et de l'autre), et ce dans une recette de construction automatique (outil kameleon [19]) afin que le tout soit intégrable dans le process industriel de gestion des OS de l'infrastructure (CI-CD).

Le troisième challenge a été l'intégration dans le service des tests automatisés de Grid'5000. En effet, historiquement, celui-ci part du principe que tous les tests passent sur tous les matériels (tous les types de nœuds) et que tous les variantes (Debian, Ubuntu, RHEL-like) et versions d'OS proposées au catalogue de services peuvent opérer tous les nœuds. C'est malheureusement le schéma inverse qui est nécessaire pour des matériels plus exotiques que ceux gérés jusque là, dont nos Jetsons (le « Linux embedded » est bien moins normalisé que le « Linux server », l'exception est plutôt la règle). Pour notre première intégration de matériel Edge-computing, nous avons donc pour l'instant fait l'impasse sur la résolution de cette question, en désactivant les tests automatiques pour ce matériel.

3.2.4 Un pas en avant vers l'instrumentation d'autres devices

Fort de notre succès pour la mise en production dans Grid'5000 du premier cluster Edge-Computing de SLICES-FR (cluster « estats » à Toulouse), nous sommes en train de faire l'acquisition de nœuds de nouvelle génération (Jetson AGX Orin) pour lesquels nous allons capitaliser sur le savoir-faire acquis et les développements. Nous avons cependant choisi d'instrumenter à nouveau des kits de développement plutôt qu'un châssis tout intégré. Cela demandera de nouveaux développements, mais donnera plus de souplesse notamment pour permettre de fournir la mesure de consommation d'énergie à hautes fréquence et précision. Un travail préalable nous a déjà permis de construire un premier prototype pour cela.

Par ailleurs, le travail sur les Jetsons est également mis à profit pour l'instrumentation d'une autre plateforme souffrant des spécificités de l'hégémonie Nvidia : les nœuds HPC Grace-Hopper (ARM64 + GPU). Le travail est en cours.

Enfin, la capacité d'instrumentation de SLICES-FR avec la technologie Grid'5000 devrait prochainement permettre aux chercheurs de déployer l'OS et les configurations de leurs choix sur des équipements de cœur de réseau programmables avec la technologie P4 (équipement de la catégorie open systems et de type network devices), élargissant encore le spectre couvert par l'offre de services SLICES-FR.

4 Les points de vigilance

4.1 Continuité de service

L'implémentation des services et le déploiement des nouvelles ressources de l'infrastructure SLICES-RI s'étalera sur plusieurs années avant d'être ouvert en production aux utilisateurs. Pour assurer une continuité de services pour nos utilisateurs, nous devons en parallèle maintenir les infrastructures précurseuses existantes tout en dégageant du temps ingénieur pour commencer les nouveaux développements SLICES-RI et SLICES-FR. Au moment où les nouveaux services seront opérationnels, nous pourrons alors basculer les équipements non obsolètes des anciennes infrastructures vers SLICES-FR.

4.2 Intégrer les équipements des différentes catégories

Certains types d'équipements n'intègrent pas nativement des outils de prise en main hors-bande, à distance, et sont donc complexes à instrumenter dans SLICES-FR. Nous devons développer des dispositifs matériels qui viendront se greffer sur l'équipement cible jouant le rôle d'interface de contrôle distante, équivalent d'une BMC pour un serveur. De plus, des bibliothèques spécifiques doivent être embarquées sur ce greffon pour pouvoir programmer l'équipement cible. Les microcontrôleurs (définis par le type `embedded devices`) doivent être programmés par une toolchain spécifique afin de téléverser le micrologiciel sur le système. La gestion de la diversité et de la complexité des ces toolchains sera un challenge.

4.3 Le nouveau monitoring énergétique

Les sondes de monitoring d'énergie sont spécifiques aux types d'équipements qui ont des puissances de consommation électrique très variées allant de quelques milliwatts à plusieurs centaines de watts. Nous devons proposer des solutions matérielles de mesure adaptées aux différentes puissances de courant.

Notre expérience avec les précédentes solutions de monitoring a mis en évidence des limites de performances, en particulier concernant la fréquence d'échantillonnage et la précision trop faible de mesure pour détecter des pics de consommation liés à des événements systèmes ou réseaux. Nous travaillons à l'intégration de nouveaux capteurs de mesure plus performants.

Enfin, le déploiement à large échelle de nos équipements et sondes associées soulève un problème de synchronisation des mesures avec une précision visant la microseconde pour être capable de croiser tous les résultats, notamment les envois et réceptions de messages radios.

4.4 Interactions et complémentarité avec d'autres infrastructures de l'ESR

L'organisation SLICES-FR, elle-même multi-partenaires et multi-tutelles, interagit avec le projet Abaca [20] d'Inria qui co-opère les sites SLICES-FR hébergeant les ressources de type server-class machines. D'autres acteurs des infrastructures numériques de l'ESR comme France Grilles [21] ou MesoNET [22] proposent des services de type centre de calculs ou clouds IAAS et PAAS qui répondent en partie aux nouveaux usages de nos utilisateurs, sous réserve de garantir une reproductibilité pour l'instanciation des ressources et de fournir des dispositifs de mesure permettant l'analyse des expérimentations.

La généralisation des datacenters labellisés à l'échelle nationale nous offre des hébergements de qualité pour nos équipements : haute disponibilité des fluides, connectivité réseau redondante et performante, sécurité d'accès, gestes de proximité. Cependant, ces offres sont limitées à des équipements de types racks pas compatibles avec tous nos types d'équipements. Par ailleurs, l'accès au monitoring et contrôle des PDU's n'est pas toujours autorisé. La connectivité réseau au niveau de l'armoire est parfois strictement gérée par une fabrique locale qui peut être en conflit avec les protocoles réseaux que l'on souhaite amener jusqu'aux équipements.

5 Conclusion

Pour arriver à un niveau opérationnel de production, SLICES-FR doit relever de nombreuses difficultés techniques et organisationnelles, aussi bien au niveau français qu'européen. Ce projet d'infrastructure de recherche s'inscrit sur une durée longue de 20 ans avec une migration progressive vers les nouveaux services qui s'étalera sur plusieurs années.

6 Glossaire

5G Core : coeur de réseau 5G

BMC : Baseboard Management Controller

ERIC : European Research Infrastructure Consortium

HPC : High Performance Computing

IoT : Internet of Things

NFV : Network Functions Virtualization

NREN : National Research and Education Network

OSI : Open Systems Interconnection

SDN : Software Define Network

SDR : Software Define Radio

Bibliographie

- [1] Grid5000 ; <https://www.grid5000.fr> .
- [2] Equipex FIT (Equipement d'excellence Future Internet Testing facility); <https://www.fit-equipex.fr> .
- [3] SLICES-FR (Scientific Large Scale Infrastructure for Computing/Communication Experimental Studies – France) ; <https://slices-fr.eu> .
- [4] PEPR (Programmes et Equipements Prioritaires de Recherche); <https://anr.fr/en/france-2030/programmes-et-equipements-prioritaires-de-recherche/> .
- [5] PEPR Cloud ; <https://pepr-cloud.fr/fr/> .
- [6] PEPR Réseaux du futur ; <https://pepr-futurenetworks.fr> .
- [7] SLICES-RI (Scientific Large Scale Infrastructure for Computing/Communication Experimental Studies – Research Infrastructure) ; <https://www.slices-ri.eu> .
- [8] ESFRI (European Strategy Forum on Research Infrastructures) ; <https://www.esfri.eu/> .
- [9] GEANT ; <https://geant.org> .
- [10] FIT IoT-LAB ; <https://www.iot-lab.info> .
- [11] FIT R2LAB ; <https://r2lab.inria.fr> .
- [12] FIT CorteXlab ; <https://www.cortexlab.fr> .

- [13] SILECS (Super Infrastructure for Large-Scale Experimental Computer Science) ;
<https://silecs.net> .
- [14] Edugain ; <https://edugain.org> .
- [15] RENATER ; <https://www.renater.fr>
- [16] Kadeploy ; <https://kadeploy.gitlabpages.inria.fr/>
- [17] Linux 4 Tegra ;
<https://docs.nvidia.com/jetson/archives/r35.6.0/DeveloperGuide/index.html>
- [18] G5k-check ; <https://grid5000.fr/w/G5k-checks>
- [19] Kameleon ; <https://kameleon.imag.fr/>
- [20] Abaca (Moyens de calcul Inria); <https://www.inria.fr> .
- [21] France Grilles; <https://www.france-grilles.fr> .
- [22] MesoNET; <https://www.mesonet.fr> .