



GRICAD

Grenoble Alpes Recherche –
Infrastructure de Calcul Intensif et
de Données

Violaine Louvet

Mars 2016 - Grenoble

- **Calcul intensif**, traitement, diffusion et préservation des **données scientifiques** au coeur des activités de recherche.
- Problématiques du calcul et du traitement intensif de plus en plus **indissociables** de l'exploitation et la valorisation des grandes masses de données.
- **Contexte contraint** sur le plan énergétique, financier, humain.

Nécessité de repenser l'**urbanisation des infrastructures de calcul et de données** ainsi que l'**organisation des moyens humains** en synergie avec les laboratoires de recherche et les services communs.

Contexte et historique du site Grenoblois

- **Projet CIMENT** autour du calcul intensif avec un modèle organisationnel distribué mixte très pertinent.
- **Projet SUMMER** co-construit par des personnels de laboratoires et de services centraux pour la mutualisation d'une plateforme de stockage de données.
- **Engagement politique** des membres des établissements qui s'est traduit par la volonté de créer des **datacentres mutualisés**.



- Partage de ressources de calcul intensif sur le site grenoblois depuis **1998**.
- **13 plateformes de calcul** :
 - une puissance de **132 Tflops/s**
 - de 96 à 3216 coeurs, pour un total de **7068 coeurs**
 - **1,4 Pbytes de stockage**
 - accessibles en **mode grille** (CIGRI).
- Une puissance de calcul **accessible à tous les personnels** des établissements de recherche académiques du site grenoblois.
- Des **ressources dédiées à des usages locaux** mais :
 - qui peuvent être utilisées par la communauté (mode best-effort)
 - dont l'administration est mutualisée
- Un **partage des ressources**, pas de demande d'allocation d'heures

- **Froggy** (calculs parallèles), **Luke** (calculs séquentiels, traitement massif): les deux machines principales communes.
- **CIGRI** : Ressources accessibles en **mode grille**
- Stockage **IRODS** permettant l'accès aux données réparties sur différents sites et sur des supports hétérogènes de façon transparente pour l'utilisateur.



- **Organisation en pôles** : un responsable technique et un responsable scientifique représentant les communautés de recherche.
- **Modèle humain** : un réseau d'ingénieurs, impliqués dans des équipes de recherche, deux ingénieurs dédiés.

SUMMER – Stockage Unifié Mutualisé Massif Evolutif et Réparti

- Etude en **2012** autour des besoins en stockage des laboratoires, composantes et services communs.
- Mise en place en 2013 d'une **infrastructure de stockage mutualisée**, pérenne et évolutive couvrant la majorité des besoins de stockage en s'adaptant le plus possible aux différents niveaux de service requis.
- Squelette de l'infrastructure financé par **l'UJF**, pour permettre ensuite aux entités de pouvoir acquérir de la volumétrie et un niveau de service associé.



SUMMER – Etat actuel

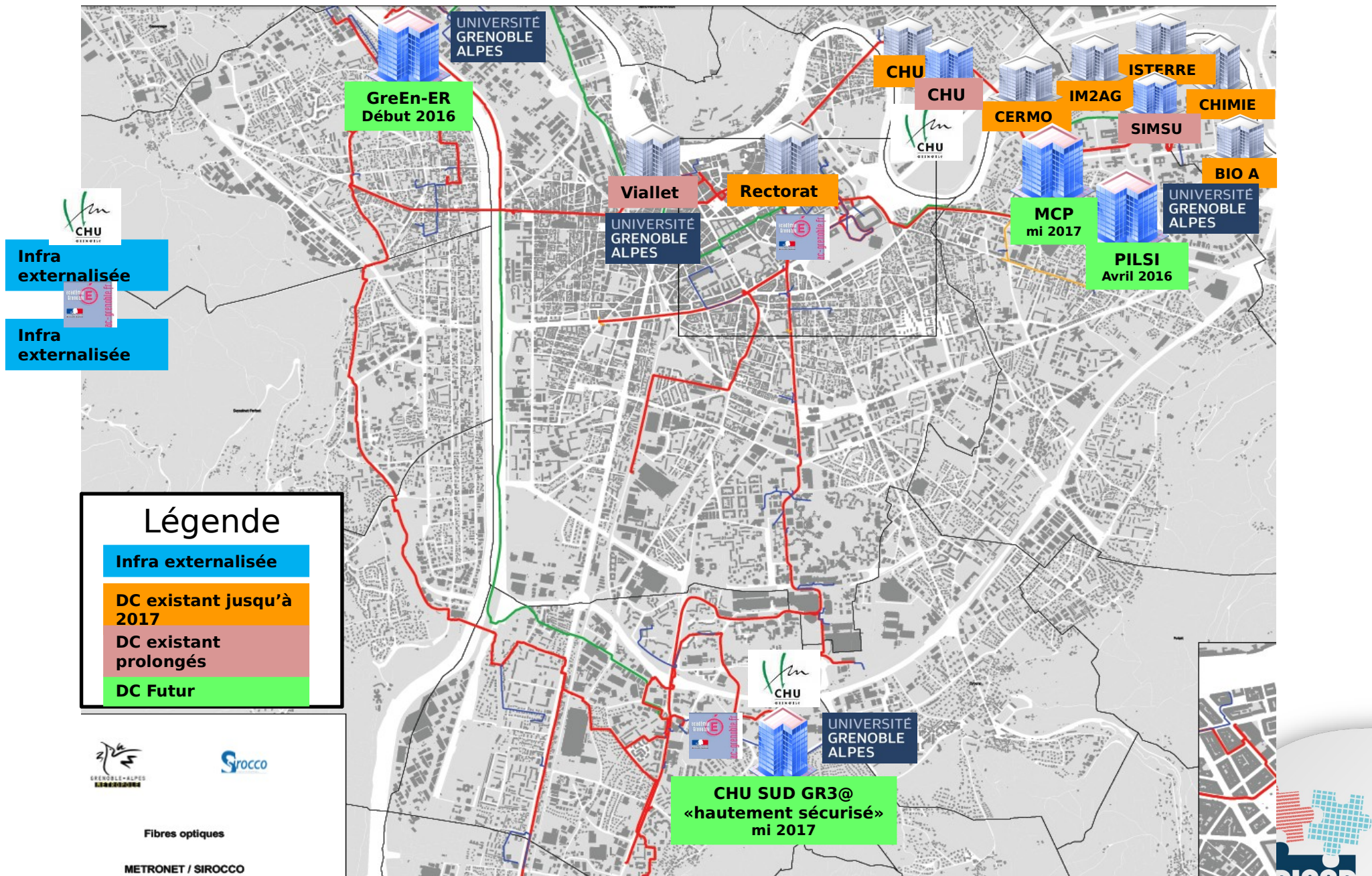
- **Sites**
 - 2 nœuds de stockage (salles BIO A / SIMSU)
 - 1 nœud de sauvegarde (salle IM2AG)
- **Organisation**
 - Comité de Pilotage (moitié recherche, moitié services centraux)
 - Comité Technique Stockage opérationnel : informaticiens (laboratoires, composantes, services communs) détachés d'une partie de leur temps pour exploiter l'infrastructure de stockage.
- **Modèle de coût**
 - calcul au Téra-Octet et par an, fonction de la classe de service du stockage (simple, avec sauvegarde, avec réplication)
- **Indicateurs**
 - 15 Laboratoires (~350To)
 - 4 composantes (~15To)
 - 2 services centraux (~250To)

Datacentres

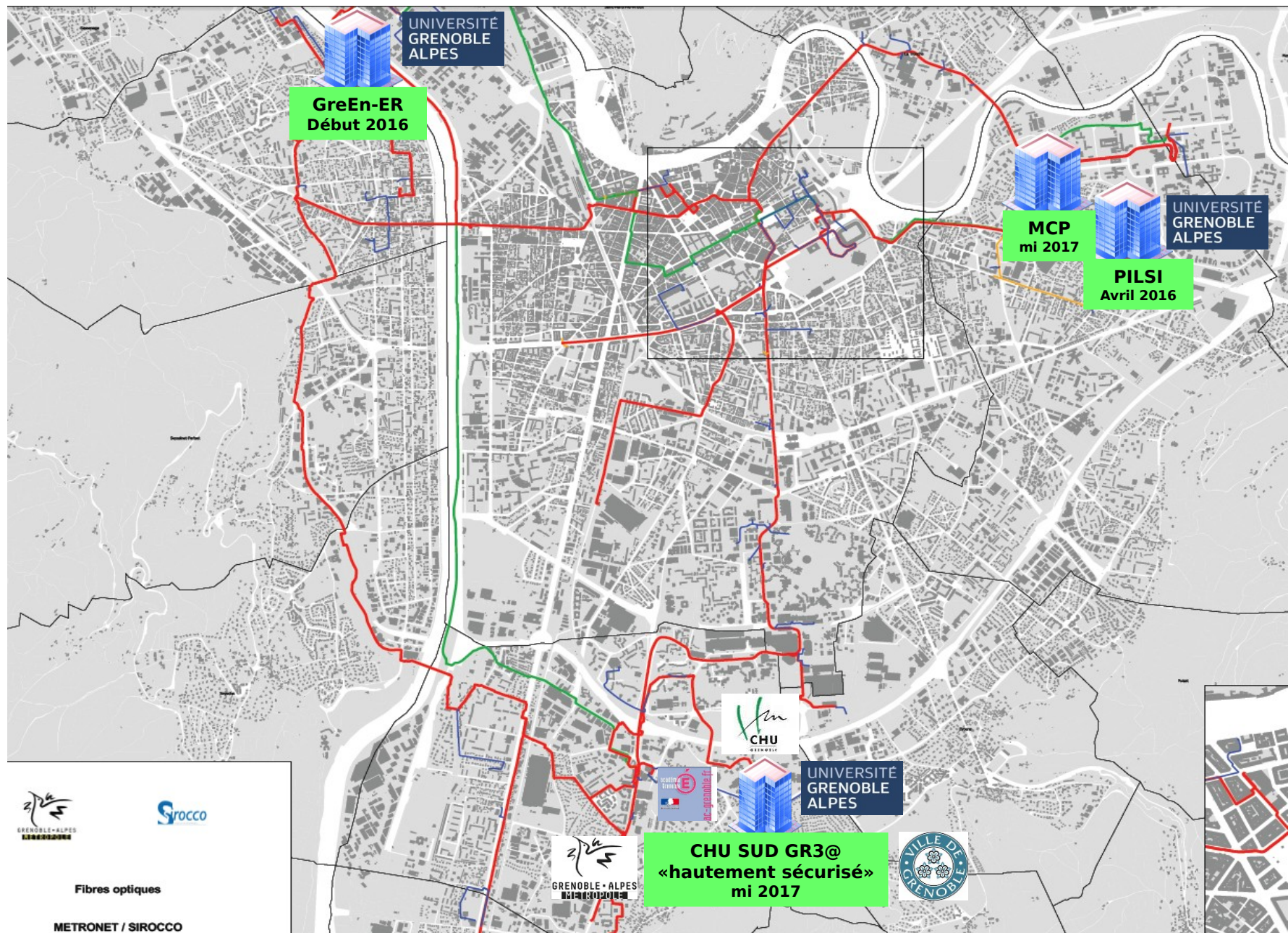
- **Projet d'établissement (UJF) en 2011** : infrastructure mutualisée éco-efficiente et redondante.
- **Etat des lieux**
 - 6 salles mutualisées
 - Une quarantaine de petites à moyennes salles (UJF)
- **Perspectives : projet de site COMUE 2016-2020**
 - 4 datacenters en cours d'aménagement ou de construction (à échéance 2016 à 2018)



Datacentres actuels et futurs



Datacentres cible 2020



Unité Mixte de Services GRICAD

Création d'une UMS au 1^{er} janvier 2016, soutenue par le CNRS, Grenoble INP et l'UGA pour la mise à disposition de l'**ensemble des personnels du site grenoblois de moyens mutualisés d'infrastructures informatiques, de calcul intensif et d'exploitation de données.**



Missions Infrastructures

- **Missions autour du calcul intensif**

- Gestion des plateformes de calcul et de traitement intensif et accompagnement des utilisateurs.



- **Missions autour des données**

- Gestion des infrastructures mutualisées de stockage, de diffusion et de sauvegarde.
- Support aux utilisateurs sur tout le cycle de vie de la donnée.



- **Missions d'infrastructures de base en co-pilotage avec les structures partenaires**

- Gestion des salles des datacentres mutualisés du site grenoblois.
- Hébergement physique de serveurs informatiques.
- Hébergement des infrastructures virtualisées mutualisées.
- Réseaux.

- **Missions d'infrastructures et de services**

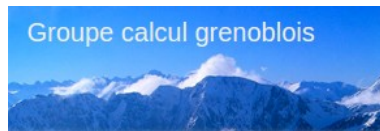
- Exploitation des infrastructures virtualisées
- Administration des services : forge, base de données...



Missions expertise et réseautage

- **Missions d'expertises et de réseautage**

- Réseau de compétences, diffusion d'expertises sur les méthodologies et les outils du calcul intensif et la problématique des données.
- Veille technologique et expérimentation.
- Lieu d'échanges et de développement interdisciplinaire, favorisant les synergies entre les communautés.

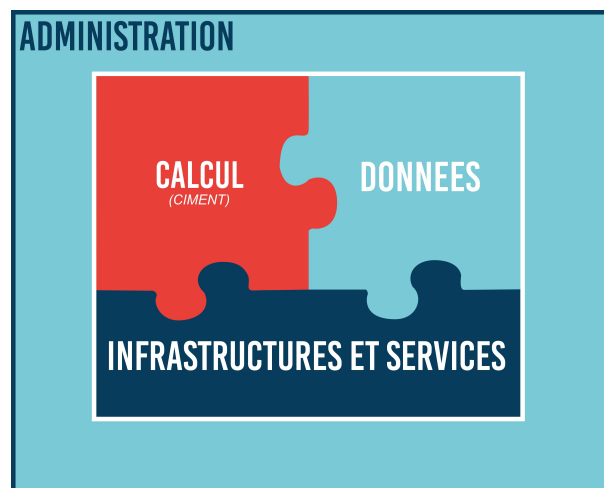


- **Missions de communication et soutien**

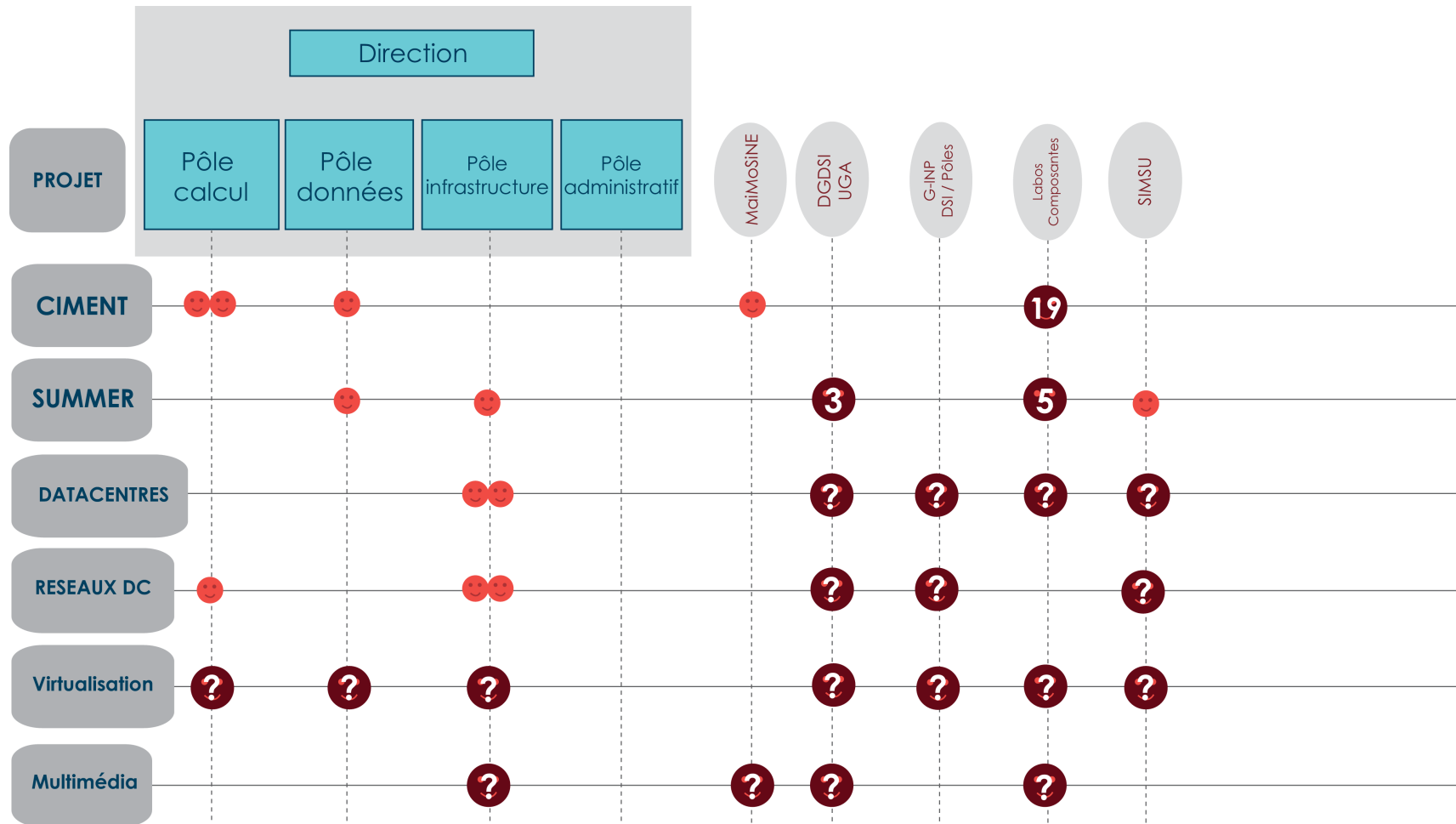
- Communication, évangelisation à tous les niveaux (personnels, DU, ...).
- Aide et support au montage de projets nationaux ou européens sur la partie technique.
- Facilitateur pour le regroupement des infrastructures de serveurs issues des laboratoires et des services communs au sein des datacentres mutualisés (aide au montage financier et au choix techniques).

Modèle organisationnel

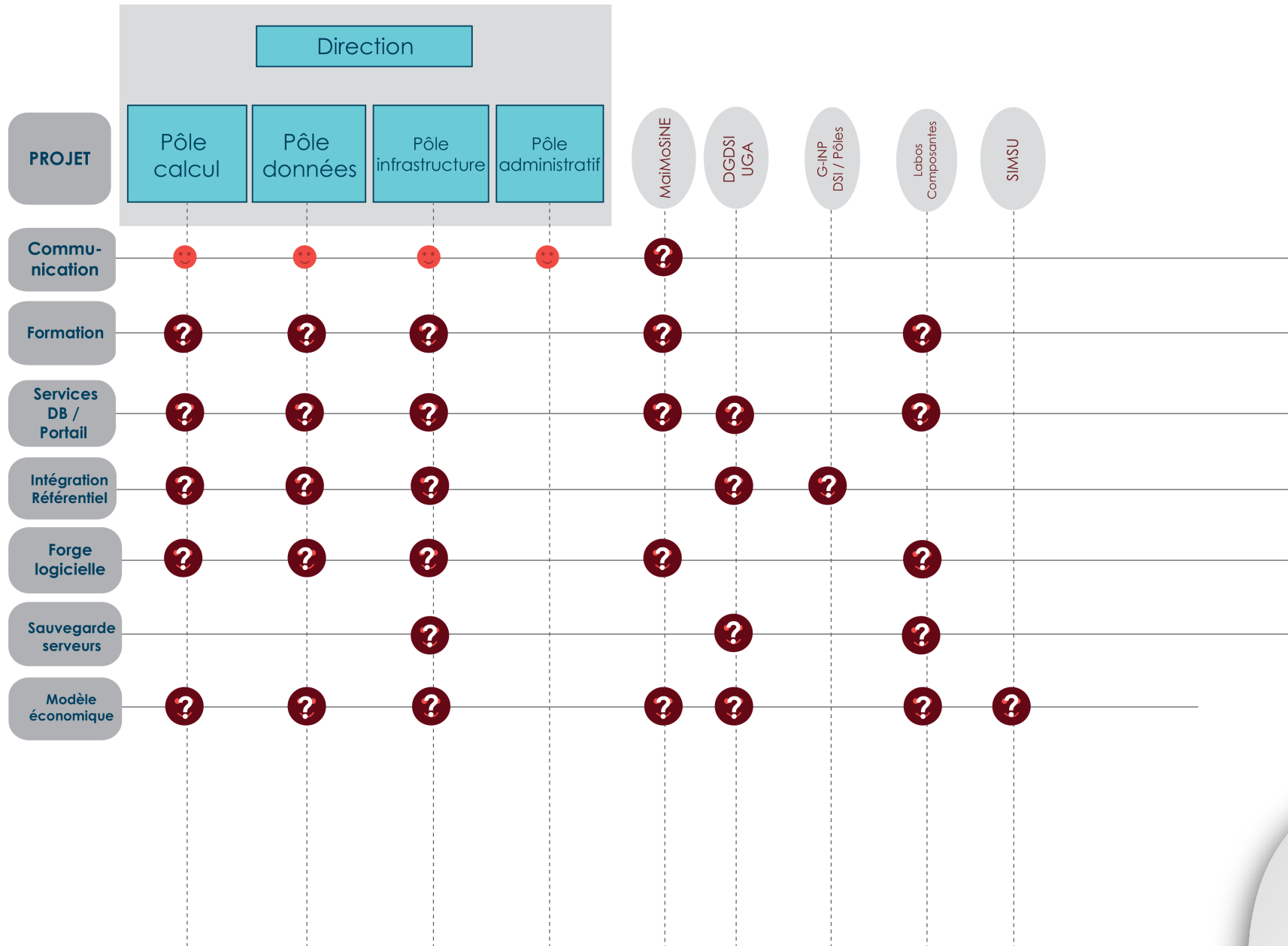
- **Organisation de GRICAD en pôle, avec un fonctionnement par projet**
- **Modèle organisationnel distribué mixte** : personnels attachés à la structure à plein temps et personnels des unités de recherche, des composantes et services communs investis à temps partiel dans les projet portés par l'UMS.
 - Modèle développé dans CIMENT puis dans SUMMER qui a fait ses **preuves d'efficacité et d'intérêt** pour les personnes impliquées.
- Favorise la **proximité avec les équipes de recherche et le réseautage**.
- Facteur de **synergie** et de **montée en compétences** de l'ensemble des personnes concernées.
- Importance **d'officialiser** l'implication des personnels des laboratoires par des lettres de mission.



Organisation par projets (en nombre de personnes impliquées)



Organisation par projets (en nombre de personnes impliquées)



Groupes opérationnels existants / en cours de construction

CIMENT

Bruno Bzeznik
Emmanuel Chaljub

- 4 GRICAD
- 1 MaiMoSiNE
- 20 labos
- 2 INRIA
- 1 CEA

SUMMER

Fabien Drago-Rajon

- 1+? GRICAD
- 5 labos
- 3 DGDSI UGA
- 1 SIMSU

Datacentres

Gabrielle Feltin
Gaëtan Enderlé

- 4 GRICAD
- 4 DGDSI UGA
- ? G-INP
- ? SIMSU

Réseaux DC

Julien Tomassone
Nicolas Gibelin

- 2 GRICAD
- 4 DGDSI UGA
- 4 labos
- ? G-INP
- ? SIMSU

Virtualisation opérationnelle

Guenael Sanchez
Nicolas Gibelin

- 3 GRICAD
- 6 DGDSI UGA
- ? labos
- 1 G-INP
- 1 SIMSU

Virtualisation prospective

Gaëtan Enderlé
Gérard Forestier
Bernard Boutherin

- 3 GRICAD
- 3 DGDSI UGA
- 5 labos
- 3 composantes
- 1 INRIA
- ? G-INP
- ? SIMSU

Formation

Laurence Viry

- 3 GRICAD
- ? labos

Communication

Violaine Louvet

- 4 GRICAD
- 1 MaiMoSiNE

Forge logicielle

Jean-Luc Archimbaud

- 2 GRICAD
- ? labos

Modèle économique

Violaine Louvet

- CD GRICAD
- Tutelles
- Services juridiques
- Services financiers
- DG DSI UGA
- SIMSU
- DSI G-INP
- Labos

Multimédia

Comité d'orientation

1 réunion / an

Représentant par pôle COMUE
Représentant SI UGA, G-INP, CNRS, COMUE
Représentant Scientifique UGA, G-INP, CNRS, COMUE
Représentant MaiMoSiNE, INRIA, CEA
Représentant DSI, composantes, formation
Comité de direction GRICAD
Experts extérieurs : calcul intensif, données, infrastructures

- Apporte une réflexion de prospective à moyen et long terme sur le développement des services proposés par l'UMS.
- Fixe les grandes orientations stratégiques, scientifiques et budgétaires de GRICAD.

Comité de pilotage

Réunions au fil de l'eau en fonction des besoins

Représentant UGA, G-INP, CNRS, COMUE (Recherche et SI)
Représentant DSI
Comité de direction GRICAD
Invitation de personnes extérieures en fonction des dossiers traités

- Examine et émet un avis sur toute question relative à l'organisation générale et au fonctionnement de GRICAD.
- Arbitre les gros projets et les choix scientifiques et techniques.
- Examine et valide les demandes de moyens (humains, financiers).

Comité exécutif

1 réunion / mois

Comité de direction GRICAD
Responsables des groupes projet

- Accompagne la direction dans la gestion des ressources et la définition des orientations.
- Suivi de l'ensemble des dossiers et gestion des dossiers transverses aux groupes opérationnels.
- Circulation des informations vers l'ensemble des personnes investies dans GRICAD.

Comité de direction

1 réunion / semaine

Direction
Responsables des pôles

- Suivi opérationnel au quotidien de l'activité de l'UMS.

Période transitoire en 2016

- Mise en place progressive de **l'organisation**.
- Constitution des nouveaux **groupes projets**.
- Mise en place de **la gouvernance**.
- Gestion de la mise en route du **datacentre PILSI**, suivi des autres chantiers.
- **Communication** autour des services proposés par l'UMS
 - Au près des DU des différents pôles de recherche de la COMUE
 - Au près des scientifiques, chercheurs et ingénieurs :
 - Organisation de rencontres scientifiques en collaboration avec MaiMoSiNE
 - Au sein des réseaux métiers locaux : SARI, Groupe Calcul Grenoblois

Actualités

- Séminaire sur l'ingénierie du big-data le 26 avril en collaboration avec SARI
- Démarrage des groupes virtualisation

Projet de Virtualisation

Communauté Grenoble Alpes





Les atouts de Grenoble

- Un établissement fusionné de grande taille et une COMUE établie
- Mise en place de l'UMS GRICAD = la structure pour héberger / gérer de telles infrastructures
- Des réalisations « modèles » : CIMENT, SUMMER
- Une communauté d'informaticiens dynamique, qui se connaît, qui sait travailler ensemble

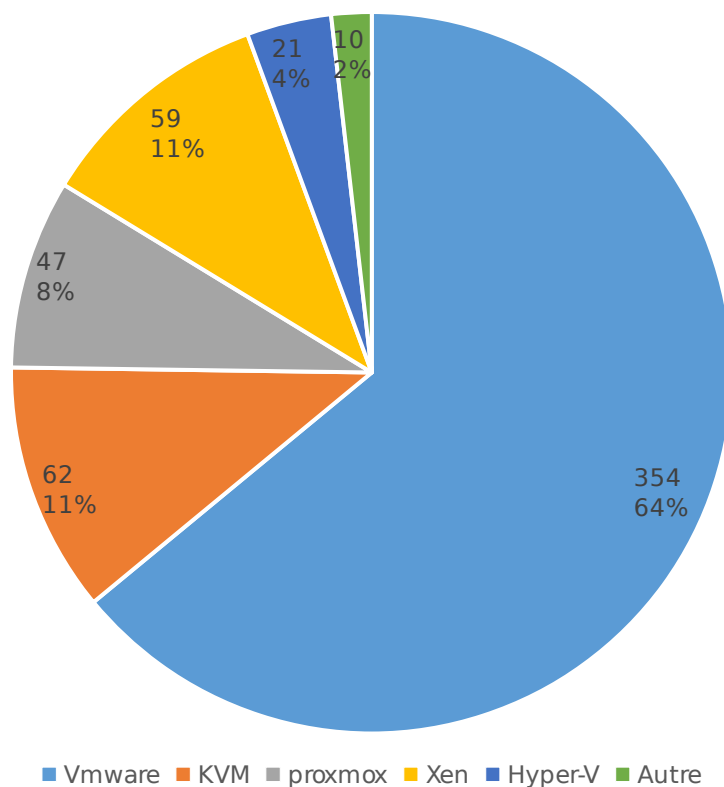
Synthèse de l'enquête

Projet d'Infrastructure d'hébergement mutualisée des ressources informatiques

- Enquête préparée dans le cadre du CPIS (Comité de pilotage des Infrastructures Scientifiques) de l'UJF
- Analyse réalisée par Gabrielle Feltin, Bruno Bzeznik, Françoise Roch, Anthony Defize, Nicolas Gibelin

Virtualisation : usages actuels

Nombre de serveurs par type de virtualisation



Synthèse de l'enquête

Projet d'Infrastructure d'hébergement mutualisée des ressources informatiques

- Dans les réponses reçues, 35% souhaitent un service mutualisé de virtualisation
- Pour susciter une adhésion plus large, accompagnement de la communauté par :
 - une communication plus ciblée auprès des équipes techniques
 - aide et conseil dans le recensement et l'expression des besoins de leurs entités

==> rôle du groupe prospective

Le projet de virtualisation

Contexte

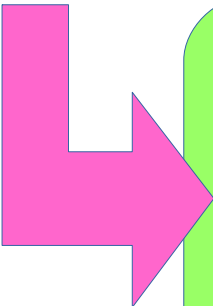
- Chaque ex-DSI, chaque laboratoire, chaque composante ainsi que le SIMSU, ont leur infrastructure de virtualisation propre.
- 2000 VMs déjà existantes à migrer (634 à la DGD SI + 500 au SIMSU) /
Cible au moins le double

Une infrastructure unique de virtualisation, sur le modèle de SUMMER, répartie sur les datacentres du schéma directeur, gérée par GRICAD, pour les besoins de la recherche, de l'enseignement et de l'administration.

Dimensionnée pour tous les besoins (recherche, formation, administration)
Résiliente, haute disponibilité, sécurisée
Evolutive de façon linéaire, hyper-convergente
Adaptée aux besoins de projets Big Data et de traitements cartographiques
Création et gestion des VM déléguée aux utilisateurs (informaticiens)
Facturation en fonction de la consommation

Les enjeux / objectifs

- **Rationalisation des infrastructures pour diminuer les coûts globaux au niveau du site.**
- Augmentation du niveau de service et de la délégation vers les utilisateurs.
- Gain en sécurité / fiabilité
- Fédération de la communauté des informaticiens autour d'un gros projet. Evolution des compétences des informaticiens



Cercle vertueux : Grâce à cette solution, candidature possible à des appels à projets (nationaux, européens, internationaux), qui permettront de financer la suite du projet



Proposition d'accord avec Dell

- Dell a sollicité l'UGA
- Dans le cadre du marché Mat Info 3
- Participation au projet avec don de matériels à hauteur de 100k€ environ, sur 2016 (rédaction d'un contrat)
- Communication commune (partenariat UGA/Dell, maîtrise de la haute technologie sur des projets complexes de grande ampleur, etc)



Fonctionnement du projet virtualisation

- 2 groupes de travail :
 - Virtualisation opérationnelle
 - Virtualisation prospective
- Réunion plénière : 1/mois
- Éléments de planning :
 - mi-2016 : mise en prod du starter pour les VM traditionnelles
 - Fin 2016 : rendu de l'étude sur l'état de l'art des technologies

Groupe Opérationnel

- **Objectif** : Consolidation des infrastructures
 - DSI Ex U1/2/3
 - Zimbra UGA
 - MI2s
- **Fonctionnement** :
 - 1/2 journée par semaine en présentiel
- **Calendrier Prévisionnel**
 - Printemps 2016 : Design & Maquettage
 - Été 2016 : Installation & Déploiement
 - Rentrée 2016 : Mise en production

Composition du groupe

- Pilotage : Nicolas Gibelin (CNRS/GRICAD) & Guenaël Sanchez (UGA/DGDSI)
- Membres actuels :
 - Gilles Bruno (UGA/DGDSI)
 - Anthony Defize (UGA/GRICAD)
 - Gabrielle Feltin (CNRS/GRICAD)
 - Yoann Mitaine (UGA/DGDSI)
 - Thomas Montfort (UGA/SIMSU)
 - Pascal Praly (UGA/DGDSI)
 - Patrick Petit (INPG/DSI)
 - Benjamin Rocton (UGA/DGDSI)
 - Loïc Thillier (UGA/DGDSI)
- + Représentant du groupe « Prospective »



Groupe Prospective

- **Cible : Mise en place d'une infrastructure de virtualisation unifiée et mutualisée**
- **Objectif :**
 - Prise en charge des besoins de l'ensemble de l'UGA (Recherche, Enseignement, Big data et Calcul)
 - Prise en compte des paramètres économique, performance, sécurité et du développement durable.
- **Fonctionnement :**
 - 1 plénière par mois
- **Calendrier Prévisionnel**
 - Printemps/Été 2016 : Etats des lieux des besoins et des usages
 - Printemps/Automne 2016 : veille technologique, rencontre des différents acteurs du marché

Composition du groupe :

- Pilotage : Bernard Bouterin (CNRS/LPSC), Gaëtan Enderlé (UGA/DGDSI) et Gérard Forestier (UGA/IM2AG)
- Membres actuels :
 - Cyrille Bonamy (CNRS/LEGI)
 - Romain Cavagna (CIMENT/GRICAD)
 - Marc Coiffier (UGA/IM2AG)
 - Fabien Drago Rajon (UGA/DGDSI)
 - Jacques Garnier (CNRS/OSUG)
 - Anne Farnier (UGA/IUT)
 - Gabrielle Feltin (CNRS/ GRICAD)
 - Patrick Fulconis (UGA/Verimag)
 - Christian Lenne (UGA/GRICAD)
 - Patrice Navarro (CNRS/LJK)
 - Pascal Praly (UGA/DGDSI)
 - Jean-François Scariot (INRIA/Montbonnot)
 - + Représentant du groupe « Opérationnel Virtu »

