La synergie grenobloise

Frédéric Desprez (Grid'5000), Violaine Louvet (GRICAD)

Le contexte grenoblois

- Projet CIMENT depuis 1998
- Fusion des universités, IDEX
- Engagement fort des établissements vers la mutualisation des ressources : datacentre, plateformes (stockage, cloud)
- Unité de lieux des laboratoires du numérique
- Communauté soudée qui a l'habitude de décloisonner et de travailler ensemble sur des projets communs

Communauté
UNIVERSITÉ Grenoble Alpes

Grid’5000 : instrument scientifique national

- « A **large-scale** and **versatile** testbed for **experiment-driven research** in all areas of computer science, with a focus on parallel and distributed computing including **Cloud, HPC and Big Data** »

- Plate-forme **nationale** et **distribuée** structurée en Groupement d'Intéret Scientifique (**GIS**)
  - CNRS, Inria, CEA, CPU, CDEFI, IMT, Renater
  - Implication des **scientifiques sur chaque site**
  - Une équipe technique : **5-8 ingénieurs répartis sur les sites**

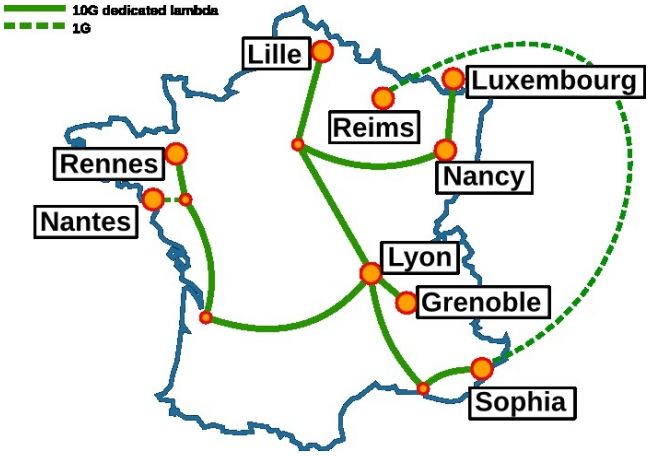
  → **LIG acteur majeur de la plate-forme** (pilotage, design, logiciels, utilisateurs)

- Une communauté d'utilisateurs **nationale et internationale**
  - Ouvert à **tous les partenaires académiques** depuis 2006
  - Depuis fin 2013 : programme « **open access** »
  - 2 825 utilisateurs au total depuis 2005, 1 099 comptes actifs sur les 3 dernières années, 532 comptes actifs sur les 12 derniers mois
  - > 1 401 publications, environ 30 HDRs et thèses par an
  - 7 473 153 réservations de ressources, pour 32 100 coeur.années, 549 016 réservations ; 4 988 coeur.années sur les 12 derniers mois

- Une **renommée internationale**
  - USA : projets NSF inspirés de Grid'5000 (Chameleon cloud)
  - Ainsi qu’au Luxembourg, Brésil, Colombie, Portugal, …

Grid’5000 : une plateforme expérimentale

- Infrastructure de **grande échelle** et **mutualisée**



- **Une plate-forme nationale**
  - 8 sites avec interconnexion dédiée (10Gbps)
  - 8000 cœurs (1000 nœuds)
  - 500+ utilisateurs actifs chaque année
  - Des clusters thématiques : HPC, Cloud, BigData

- **Reconfigurabilité et contrôle**
  - L'utilisateur déploie sa propre plate-forme ! (**Hardware-aaS**)
  - Expérience sur toutes les couches logicielles
  - Isolation/reconfiguration réseau
  - Monitoring (performance, énergie)
  - Traçabilité de la plate-forme
  - Orchestration d'expérience

→ **L'utilisateur déploie sa propre pile HPC, Cloud ou Big-Data !**

Grid’5000 – travaux en cours et futurs

- **Calcul à hautes performances**
  - Relations avec des fonctionnalités hardware particulières et combinaisons CPU/réseau
  - Nouveaux stockages (SSD)
  - nouvelles fonctionnalités matérielles (manycore, hyperthreading, burst mode, …)
- **Clouds et virtualisation**
  - Virtualisation du réseau (Impact des SDNs)
  - Virtualisation du stockage
  - Piles logicielles Clouds disponibles pour les développeurs d'applications
  - Injection de traces réelles (workloads, fautes)
- **Gestion de l'énergie**
  - Monitoring de l'énergie des infrastructures virtualisées
  - Besoin d'avoir des architectures hétérogènes (GPU, ARM)
- **Gestion de grandes masses de données**
  - MapReduce, Workflows
  - Problématiques de stockage liées au HPC
- **Recherche reproductible**
  - Description et vérification de l'environnement expérimental
  - Améliorer le contrôle et la description des expérimentations

GRICAD

- **Unité Mixte de Service**, création au **1er janvier 2016**
- **Tutelles** : CNRS, Université Grenoble Alpes, Grenoble-INP.
  - Liens avec l'INRIA et le CEA
- **Partenariats forts** avec tous les acteurs du site : DSI UGA et G-INP, SIMSU, laboratoires, composantes d'enseignements
- **Porteur des projets d'infrastructures numériques mutualisées** du site : coordination et pilotage en mutualisant les compétences et les expertises présentes

- **Modèle organisationnel mixte**
  - Personnels rattachés à la structure et personnels des unités, services communs et composantes investis à temps partiel dans les projets.
  - Favorise la proximité avec les équipes de recherche, facteur de montée en compétence.

GRICAD : Missions

- **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.

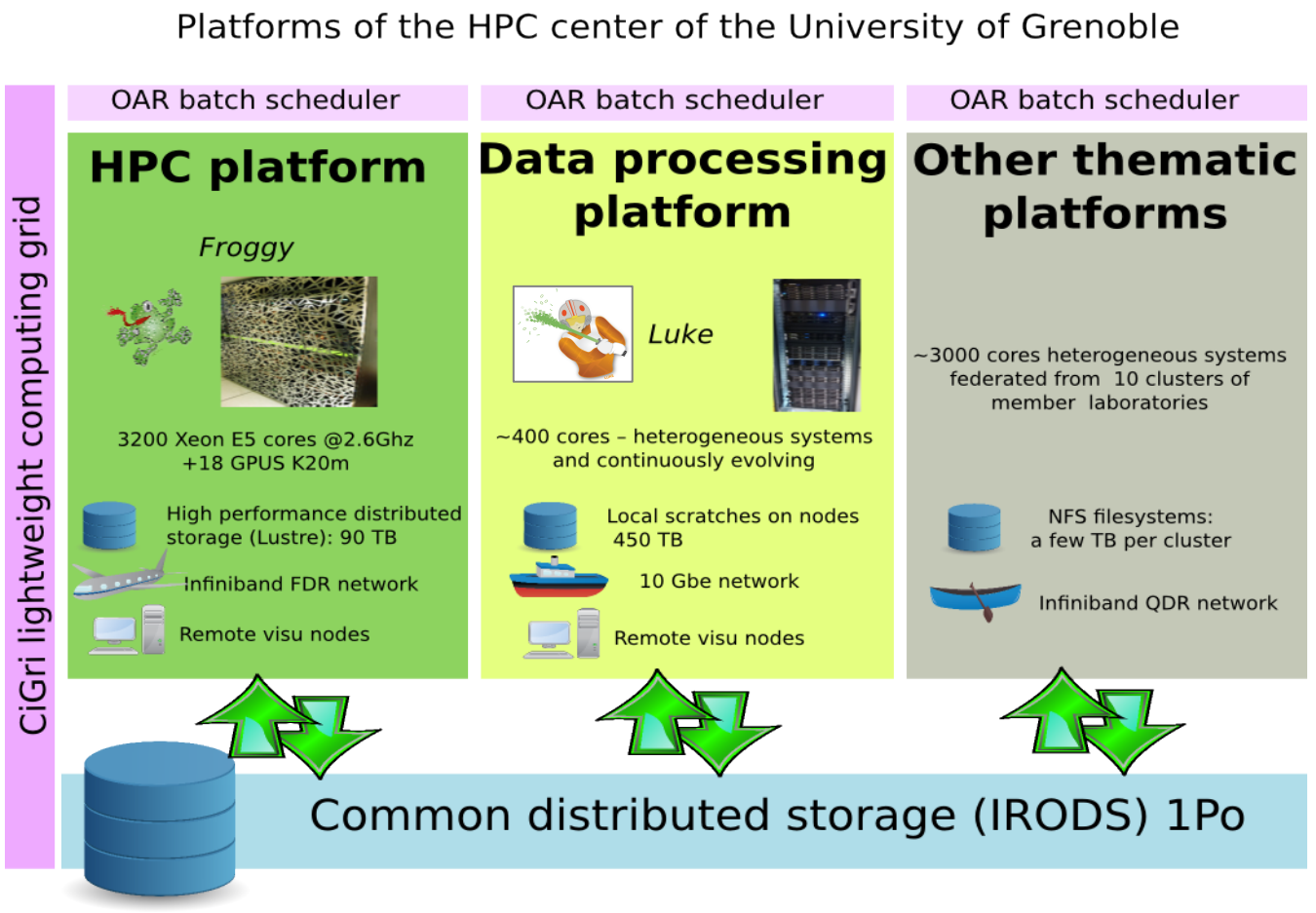
- **Missions d'infrastructures de base en co-pilotage avec les structures partenaires**
  - Gestion des salles des datacentres mutualisés du site grenoblois, hébergement, réseaux

- **Missions d'infrastructures et de services**
  - Administration des services à valeur ajoutée pour la recherche : forge logicielle,bases de données…

GRICAD
GRENOBLE ALPES RECHERCHE
INFRASTRUCTURE DE CALCUL INTENSIF ET DE DONNÉES

- **Missions d'expertises, de réseautage et de soutien**
  - Réseau de compétences.
  - Veille technologique et expérimentation.
  - Support au montage de projets nationaux ou européens sur la partie technique.

CIMENT : plateforme de calcul grenobloise



7068 cores
132 Tflops
1.4 PBytes

Pilotage par les chercheurs

Organisation en **pôles** par domaine scientifique

Correspondants dans les labos

Groupe Opérationnel mettant à **contribution** les ingénieurs des labos

Accès souple :

Ouverture de comptes utilisateurs et création de projets **au fil de l'eau**

Partage des ressources pragmatique

Accès via un gestionnaire de tache (batch scheduler)

Intégration de ressources dédiées (labo) avec **partage en mode grille / best-effort**

Administration **mutualisée**

Communauté et formations

Partenariat avec **MaiMoSINE**

Journée annuelle des utilisateurs, séminaires labo

Formation école doctorale, groupe calcul grenoblois

2 plateformes complémentaires

	Grid'5000	GRICAD CIMENT
Objectif général	Contribution à la recherche informatique *«L'objectif est la méthode »*	Calcul Scientifique *«L'objectif est le résultat »*
Communauté utilisateurs	Recherche informatique « et + si affinité » Plate-forme nationale	Toutes les disciplines scientifiques Plate-forme régionale
Domaines d'experimentation	Multiples : HPC, Cloud, Big Data... Informatique distribuée au sens large	Calcul Intensif, traitement intensif (cas d'usage classiques)
Particularité de l'Infrastructure	« Expérimentation » Contrôle, traçabilité, reconfigurabilité, interactivité	« Production » Optimisée pour la puissance de calcul, traitement par lot

Challenges Grid'5000 VS Mésocentre

- **Expérimentation VS Production**

  - **Finalité:** objectif méthode VS résultat
    - pas besoin de Walltime aussi long pour l'expérimentation

  - **Interactivité:** traitement batch difficile pour l'expérimentation (réservation à l'avance)
    - Difficulté de scripter des environnements complexes et d'avant garde/instables
    - Gros effort pour l'instrumentation des expériences (reproductibilité), mais plus complexe qu'un script HPC classique

  - **Traçabilité de la plateforme:** pour la reproductibilité de l'expérimentation, il est important de garder les logs des modifications de l'infrastructure VS production vise surtout la justesse du résultat
    - **Evolution :** production scientifique et non industrielle -> plus proche de l'expérimentation

  - **Communauté:** informaticien VS tout domaine scientifique
    - culture différente: l'informaticien veut des fonctionnalités en avance VS le physicien veut son résultat !

  - **Environnement applicatif:** la plateforme d'expérimentation est « nue » VS l'utilisateur HPC production s'attend à avoir un environnement application clé en main (modules & co)

Challenges Grid'5000 VS Mésocentre

- **Expérimentation Cloud&co VS Expérimentation HPC**

  - Déploiement *bare metal* nécessaire pour des expériences bas niveau sur le système
    - bcp plus commun pour la communauté cloud que la communauté HPC
    - le Cloud met en avant les "couches systèmes" et "Linux", alors que le HPC essaie historiquement plutôt de cacher la couche système (Appli <-> hardware)

  - Machines Virtuelles essentielles pour le cloud VS superflues pour le HPC
    - mais les containers arrivent dans le monde du HPC (l'utilisateur vient avec son environnement d'exécution)

  - Modification du réseau (kavlan) inutile pour le HPC ?
    - cela arrive maintenant aussi pour le HPC (SC16: SDN Infiniband, modification du routage/job)

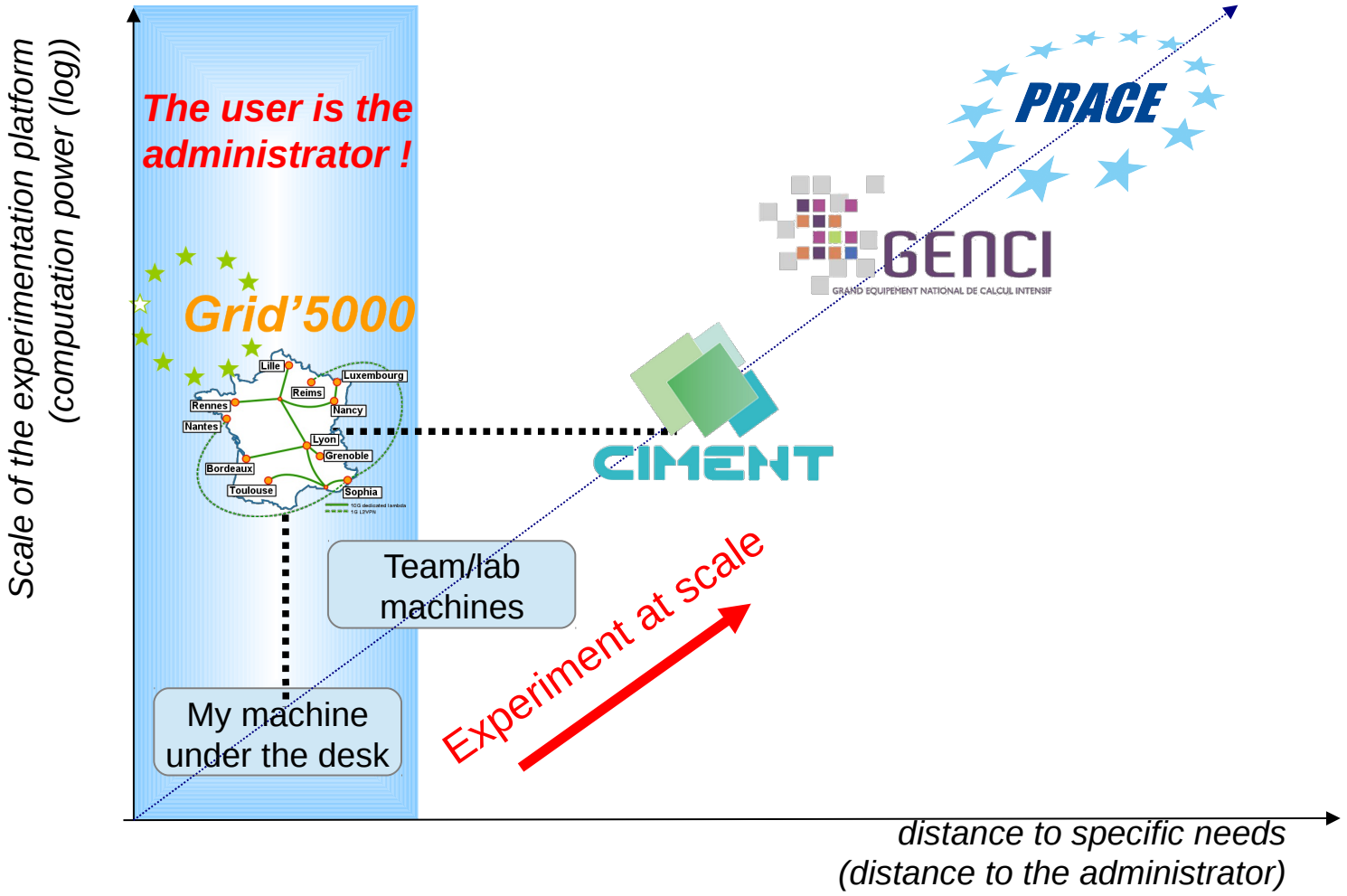
  - Description de la plateforme interrogeable de manière programmatique
    - nécessaire pour un datacentre expérimental cloud (introspection, etc), moins pour HPC: on joue moins avec la plate-forme, on part du principe qu'un cluster est installé à l'achat et évolue peu.

  - Environnement applicatif: pile BigData/etc vs pile HPC (modules), mais ça converge !

Synergies : Dev & Admin

- **Développements communs :**
  - Contribution des 2 plateformes au développement de OAR (REST API, module hulot/energy saving)
  - CIGRI (développement commun de la version 3)
  - COLMET : profiling système des jobs
- **Mutualisation des compétences administrateurs**
  - Expertise croisées sur les matériels/achats
  - Gestion d’infrastructure (services)
  - Bonnes pratiques pour la gestion des utilisateurs
  - Expertise HPC

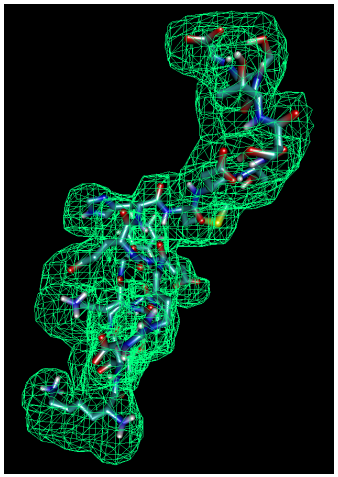
Synergies : Expérimentation HPC

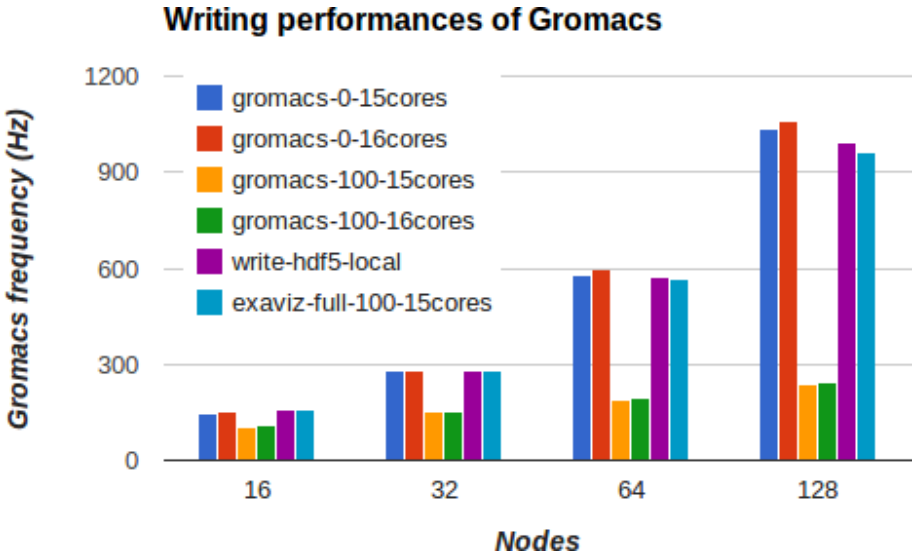


Passage à l'échelle : complémentarité des plateformes

Un exemple de la complémentarité

- Travaux B. Raffin - **HPC: In Situ Analytics**
- **But:** mieux organiser les phases de simulation et d'analyse de données
- Simuler sur un cluster; déplacer les données; analyse post-mortem
  - Inadapté à l'Exascale (volume des données, temps)
- **Solution:** analyser sur les nœuds, pendant la simulation
  - Entre ou pendant les phases de simulation ? cœur ou nœud dédié ?
- **Grid'5000** utilisé pour le développement et les tests, car contrôle:
  - De l'environnement logiciel (piles MPI)
  - Des paramètres CPU (Hyperthreading)
  - Des paramètres réseaux (QoS Infiniband)
- Puis évaluation du passage à l'échelle sur le supercalculateur Froggy (**CIMENT**)





Matthieu DREHER et Bruno RAFFIN. “**A Flexible Framework for Asynchronous in Situ and in Transit Analytics for Scientific Simulations**”. In : CCGrid. 2014.

Perspectives

- Convergence des plateformes
  - un pas de plus pour la mutualisation des financements matériels et des efforts humains ...
  - pas nouveau + nouveaux challenges pour aller + loin ! (notamment quels arbitrages entre les usages différents "Expérimentation informatique " VS "prod HPC tout domaine")
- Mésocentre/G5K = plateforme expérimentale HPC ?
- Enfin une machine dimensionnante dans Grid'5000 ?
  - 1 machine de plusieurs milliers de coeurs ?
- Une plateforme BDEC (*Big Data Extreme Computing* http://www.exascale.org/bdec/) ?
  - NVM - Storage Class Memory

Perspectives

- Contributions actives dans des communautés internationales : IRODS, NIX
- Evolution des solutions de traitement de la donnée
- Nouvelle plateforme robuste aux I/O intensifs : convergence avec le besoin de machine big-data/burst buffers de g5k@grenoble
- Utilisateurs demandeurs d'outils de 'cloud léger' pour tout ce qui concerne les applications de traitement intensif (avec la problématique de la gestion des ressources de stockage) : aller plus loin avec CIGRI ?
- Serveur de notebooks, projet UGA 2.0 (recherche) et projets en enseignement

Merci de votre attention

Questions ?

Remerciements à **Bruno Bzeznik** (responsable pôle calcul de GRICAD), **Emmanuel Chaljub** (responsable scientifique CIMENT), **Pierre Neyron** (g5k@Grenoble), **Olivier Richard** (g5k@Grenoble)