



Projet HPCDA
Cluster commun CIMENT/Grid'5000

Pierre Neyron (LIG)



Sommaire

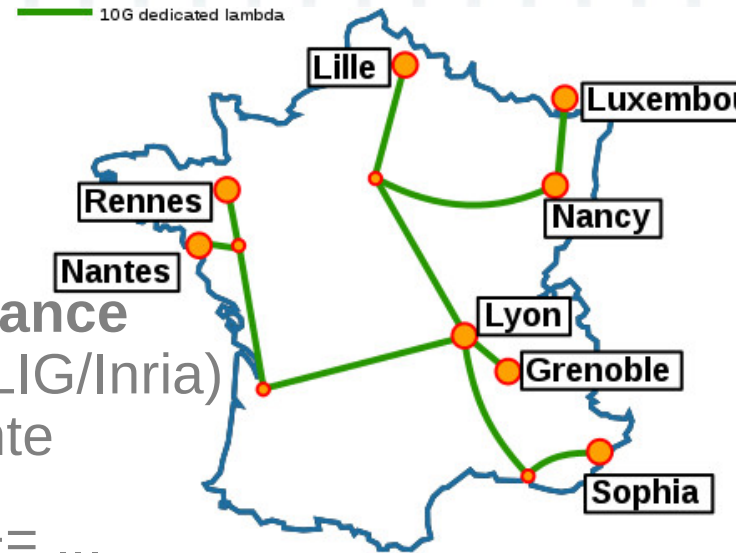
1. Rappels sur Grid'5000

2. Le projet HPCDA

Grid'5000 en bref

Plate-forme nationale 8 sites en France

- Implication très forte de Grenoble (LIG/Inria)
- Ouverte en 2006, évolution constante (matériel, services) :
HPC/Grid += Cloud += Big-Data += ...



- Développement et validation **expérimentale** du **passage à l'échelle des solutions informatiques**
- **Instrument de Recherche** (au même titre que les grands télescopes ou accélérateurs de particules) pour l'**informatique distribuée**

HPC, Cloud Computing, Big Data, Networking

Les mécanismes de Grid'5000



Expérimentation sur toutes les couches de la pile logicielle

→ en particulier : capacité assez unique pour cette taille de plateforme de **changer le système d'exploitation** (ou l'hyperviseur !)

L'utilisateur gagne les **privilèges d'administrateur** !

Configuration par l'utilisateur de **sa propre topologie du réseau** et isolation (création de vlans, réservation de pool d'IP, routage)

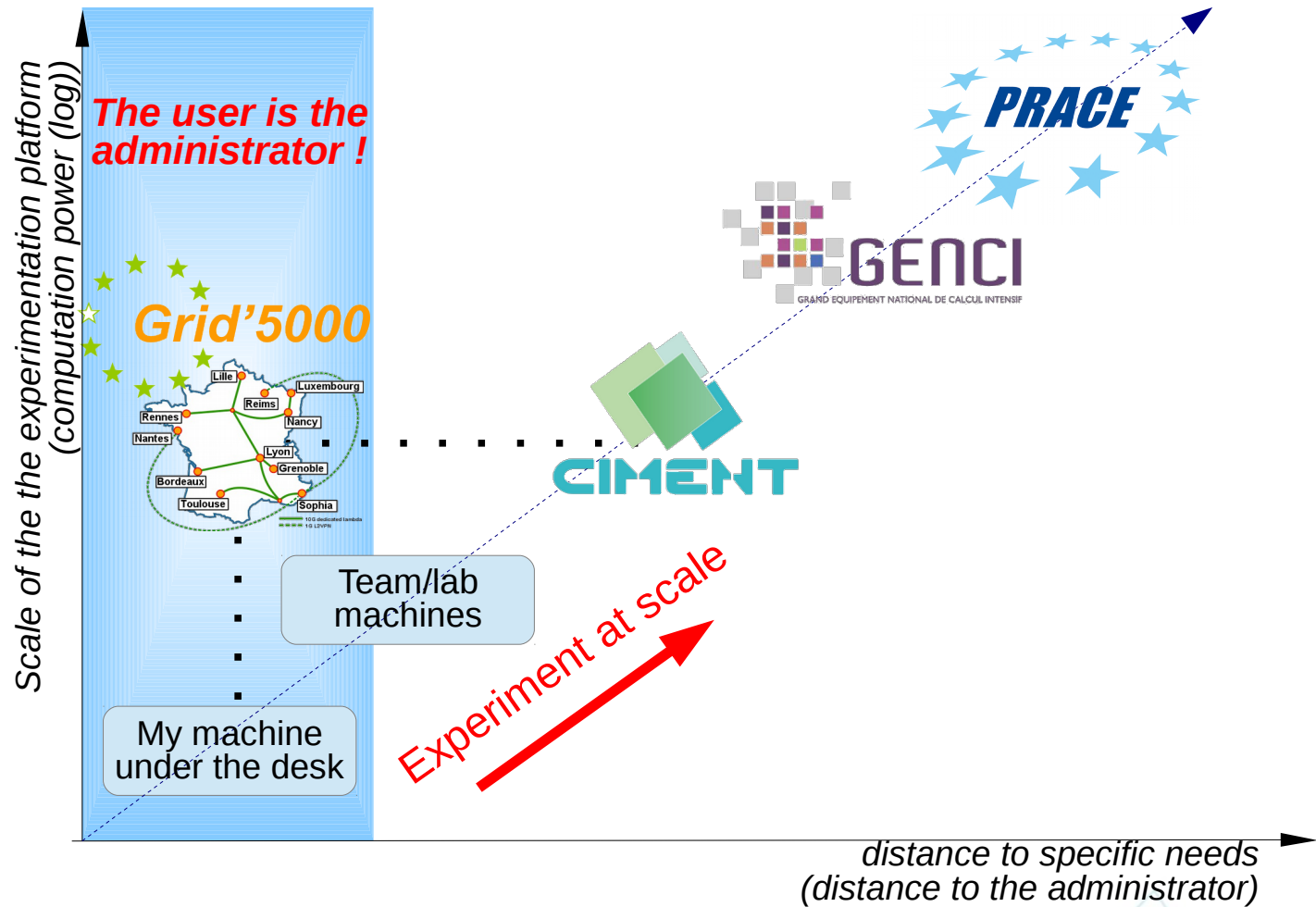
Maitrise de l'environnement d'expérimentation (≠ AWS !):

- Description, vérification et traçabilité de la plateforme / de l'environnement d'expérimentation
- Monitoring: sondes réseau, énergie

Interrogeabilité programmatique (Rest API), instrumentation des expériences → **Reproductibilité**

Une plateforme conçue pour la Recherche en Informatique Distribuée

Maîtrise du passage à l'échelle (cas HPC)



Quelques chiffres

Une plate-forme nationale homogène (administration centralisée)

8 Sites, 30 clusters, 840 noeuds, 8500 coeurs

Réseau d'interconnection dédié entre les sites: **10Gbps**

Variété matérielle: processeurs, réseaux rapides, accélérateurs, stockage

Stats : **550 utilisateurs actifs, 100 publications / an**

Pilotage structuré dans un **Groupement d'Interet Scientifique** (Inria, CNRS, CPU, CDEFI, IMT, CEA, Renater)

Une **équipe technique** nationale et distribuée sur les sites

→ ~10 ingénieurs/techniciens/chercheurs impliqués

Sommaire

1. Rappels sur Grid'5000

2. Le projet HPCDA

HPCDA = double convergence

Convergence HPC – Data Analytics (BigData)

→ *le calcul intensif rencontre les technologies des hyperscalers*

Mais aussi

Convergence **Expérimentation – Production**

→ *cluster commun pour les 2 communautés CIMENT et Grid'5000*

⇒ ***nouvel élan de collaboration technique et scientifique fructueuse !***



CIMENT : Un cluster « expérimental »

- Mécanisme de déploiement système et applicatif
→ *s'affranchir de certaines contraintes des environnements de « production »*
- Ouverture sur les nouvelles technologies matérielles et logicielles : NVRAM, burst buffers, big data
→ *lever certains verrous des technos HPC classiques*
- Maîtrise de la plate-forme → *monitoring énergétique, ...*

Grid'5000 : Un compromis gagnant-gagnant

- Plate-forme dimensionnante → *validation expérimentale*
- Développement des collaborations transversales → *cas d'usage réels*
- Transfert bi-directionnel : *technologies d'avant garde, traces de productions*

Challenge Expérimentation VS Production

		
Objectif général	Contribution à la recherche informatique «L'objectif est la méthode »	Calcul Scientifique «L'objectif est le résultat »
Communauté utilisateurs	Recherche informatique Plate-forme nationale	Toutes les disciplines scientifiques Plate-forme régionale
Domaines d'utilisation	Expérimentation HPC, Cloud, Big Data, ... Informatique distribuée au sens large	Traitement Intensif de Calculs et de Données
Particularité de l'Infrastructure	« Expérimentation » Interactivité, contrôle, reconfigurabilité	« Production » Optimisée pour la puissance de calcul, traitement par lot

Un cluster commun = un beau challenge technique :

- Une plate-forme de production orientée vers l'expérimentation
- Une plate-forme expérimentale production-proof

Comment ?

Groupe de travail G5K/CIMENT pour l'architecture
→ rédaction d'un document technique (hal-01511285)

Plusieurs scénarios envisagés

- Comment réunir les gestions des comptes ?
- Quel partitionnement des ressources, quel arbitrage pour l'exécution des tâches ?
- Quel système d'exploitation ? Système unique ?
- Quelle interconnection avec les ressources externes ?
- Quelle collaboration entre les équipes techniques
- ...

⇒ *Des compromis, du pragmatisme...*

**Scénario n°1 : intégration dans l'infrastructure Grid'5000
+ couche HPC CIMENT**

Status actuel

Financements matériel acquis !

- CPER/Inria : 275K€
- Grenoble-INP : 70K€
- Isterre (Seiscope) : 100K€
- Idex plate-forme : 100K€
→ total 545K€

Demande financement ingénieur CDD Inria en cours

Procédure d'achat en cours :

- ~100 Nœuds bi-CPU → ~ 2000 cœurs
- Réseau rapide 100Gbps Intel Omnipath
- Réseau 10Gbps Ethernet
- Stockage local sur les nœuds SSD/HDD
- Quelques nœuds NVRAM (burst buffers)

Hébergement :

- DC-IMAG, colocalisé avec **BeeGFS**, **Luke**, **Irods**
- Setup réseau 10GE en cours d'étude (**ACI Cisco UGA**)

Prochaine étapes

Juillet :

- *Lancement de l'achat (matinfo4)*
- *Réponse pour le CDD*

Septembre-Octobre :

- *Réception/installation du matériel*
- *Recrutement ingénieur (?)*
- ***Mise en place du scénario 1 :***
 - *Intégration dans l'infrastructure de service Grid'5000*
 - *Ajout de la couche CIMENT*
 - *Passerelle de gestion pour les comptes CIMENT → G5K*
 - *...*

Cible : → Mise en service au moins partielle avant la fin de l'année

Questions

Questions ?