

Journées **Grid'5000** à Lille

30-31 octobre 2006

Nouredine MELAB

Equipe OPAC, CNRS/LIFL, INRIA Futurs
Membre du comité national de pilotage de Grid5000



ministère délégué
à la recherche



INRIA *Renater*



CENTRE NATIONAL
DE LA RECHERCHE
SCIENTIFIQUE



Programme du 30 octobre 2006

09h45 - 10h00 : Accueil des participants – Café

10h00 – 11h00 : Projet Grid5000 : Etat actuel et perspectives

N. Melab, Membre du CP de Grid5000

11h00 – 11h15 : Pause

11h15 – 12h00 : La grappe de Lille : Matériel et outils logiciels

J. Garet, IE Grid5000, INRIA Futurs

12h00 – 12h30 : The KOALA grid scheduler

O.O. Sonmez, Technische Universiteit, Delft, Pays-bas

14h00 - 17h30 : Retours d'expériences – 2 sessions



Programme du 31 octobre 2006 (Formation)

09h00 - 9h30 : Accueil des participants - Café

9h30 - 12h00 : Travaux Pratiques : Initiation à Grid5000

- Découverte des outils logiciels de Grid'5000 (30')
- Réalisation d'une expérience sur une grappe de processeurs (1h)
- Réalisation d'une expérience sur une grille de processeurs (1h)

13h45 - 14h00 : Accueil des participants - Café

14h00 - 18h00 : Travaux pratiques : Utilisation avancée de Grid5000

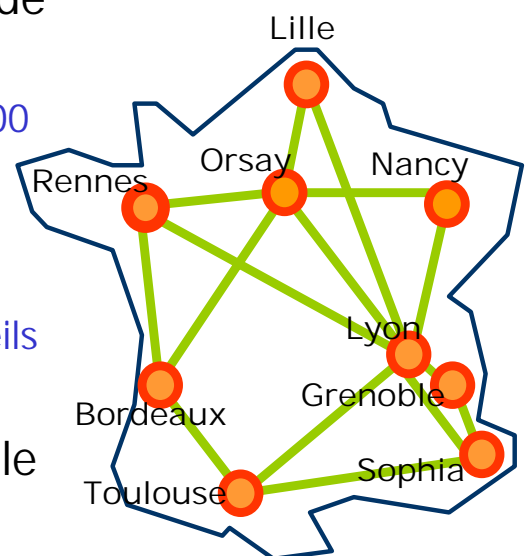
Le projet **Grid'5000**

Nouredine MELAB

Membre du comité national de pilotage de Grid5000

- Motivations et objectifs du projet
- Conception de la grille
- Etat actuel du projet
- Premiers résultats
- Perspectives du projet

- Projet lancé en 2003 par le Ministère de la Recherche
 - ACI GRID – Programme national GRID5000 (M. Cosnard puis T. Priol)
- Supporté par d'autres organismes
 - INRIA, CNRS, Universités, Renater, Conseils régionaux et généraux, ...
- Acquisition et mise en place d'une grille informatique ...
 - ... nationale **expérimentale** à l'horizon 2007
 - ... constituée de 5000 CPU réparties sur 9 sites





Motivations du projet (1)

- Existence en France d'une communauté scientifique dans le domaine des grilles informatiques
 - Projets de l'ACI GRID (slide suivant ...) ...
 - ... puis ACI Masse de données, ACI Sécurité & Informatique
 - ... et maintenant l'ANR « Calcul Intensif et Grilles de Calcul »
- Défis scientifiques et applicatifs nombreux
 - Problématiques de recherche
 - Caractéristiques : large échelle, multi-domaine d'admin., hétérogénéité, volatilité
 - Intergiciel (tolérance aux pannes, ordonnancement, performances, ...), réseau (sécurité, protocoles de communication, QoS, facturation), algorithmique et programmation (passage à l'échelle), ...
 - Applications de grande taille dans plusieurs domaines
 - Bioinformatique, médecine, physique, météo, ...



ACI GRID projects

Thierry Priol



- Peer-to-Peer
 - CGP2P (F. Cappello, LRI/CNRS)
- Application Service Provider
 - ASP (F. Desprez, ENS Lyon/INRIA)
- Algorithms
 - TAG (S. Genaud, LSIIT)
 - ANCG (N. Emad, PRISM)
 - DOC-G (V-D. Cung, UVSQ)
- Compiler techniques
 - Métacompil (G-A. Silbert, ENMP)
- Networks and communication
 - RESAM (C. Pham, ENS Lyon)
 - ALTA (C. Pérez, IRISA/INRIA)
- Visualisation
 - EPSN (O. Coulaud, INRIA)
- Data management
 - PADOUE (A. Doucet, LIP6)
 - MEDIAGRID (C. Collet, IMAG)
- Tools
 - DARTS (S. Frénot, INSA-Lyon)
 - Grid-TLSE (M. Dayde, ENSEEIHT)
- Code coupling
 - RMI (C. Pérez, IRISA)
 - CONCERTO (Y. Maheo, VALORIA)
 - CARAML (G. Hains, LIFO)
- Applications
 - COUMEHY (C. Messenger, LTHE) - Climate
 - GenoGrid (D. Lavenier, IRISA) - Bioinformatics
 - GeoGrid (J-C. Paul, LORIA) - Oil reservoir
 - IDHA (F. Genova, CDAS) - Astronomy
 - Guirlande-fr (L. Romary, LORIA) - Language
 - GriPPS (C. Blanchet, IBCP) -Bioinformatics
 - HydroGrid (M. Kern, INRIA) - Environment
 - Medigrid (J. Montagnat, INSA-Lyon) - Medical
- Grid Testbeds
 - CiGri-CIMENT (L. Desbat, UjF)
 - Mecagrid (H. Guillard, INRIA)
 - GLOP (V. Breton, IN2P3)
 - GRID5000 (F. Cappello, INRIA)
- Support for disseminations
 - ARGE (A. Schaff, LORIA)
 - GRID2 (J-L. Pazat, IRISA/INSA)
 - DataGRAAL (Y. Denneulin, IMAG)



Motivations du projet (2)

- Besoin d'outils de validation par approche ...
 - ... analytique
 - ... expérimentale

- Approche analytique
 - Modèles analytiques simplifiant les principes mis en oeuvre ...
 - ... difficile d'exprimer toute la complexité logicielle et matérielle
 - Trop abstraite pour convaincre les praticiens !



Motivations du projet (3)

- Approche expérimentale
 - Simulation : SimGrid, NS, MicroGrid, Bricks, ...
 - Emulation : GridExplorer, Emulab, WANinLab
 - Systèmes réels
 - PlanetLab – Grille planétaire (340 sites répartis sur 25 pays)
 - TeraGrid – NSF/USA (9 sites)
 - NAtional REsearch Grid Initiative (NAREGI) – Japan (11 sites)
 - Distributed ASCII Supercomputer 2 (DAS2) – Pays Bas (5 sites)
 - ...
 - **Projet français GRID5000**



Objectifs du projet (3)

- Construire une plate-forme **expérimentale** nationale
 - 9 sites distribués en France et connectés par Renater
 - Entre 500 et 1000 CPUs sur chaque site
 - Mise en place d'un système pour des expérimentations sûres et reproductibles
- Utiliser la plate-forme pour des expériences de grille **en condition réelle** afin de lever les verrous scientifiques au niveau ...
 - ... intergiciel : algorithmique et programmation, tolérance aux pannes, ordonnancement, ...
 - ... réseau : protocoles de communication haute performance, QoS
 - ... applicatif : déploiement à grande échelle, gestion des données et résultats, ...



Plan

- Motivations et objectifs du projet
- Conception de la grille
- Etat actuel du projet
- Premiers résultats
- Perspectives du projet



Objectifs conceptuels (1)

- Répondre à des besoins diversifiés
 - Enquête auprès des chercheurs
 - Résultat : ~ 100 expériences avec différentes configurations
 - Expériences à tous les niveaux de la pile logicielle (réseau, intergiciel/système, programmation, application)
 - Outils de mesure et conditions expérimentales différents
 - *Différentes métriques* : performance (temps CPU, débit, ...), extensibilité, tolérance aux pannes (fréquentes, massives)
 - *Injection de conditions expérimentales* : charge (CPU, mémoire, ...), stress (grand nombre de clients/serveurs/tâches/données), ...

Grille de processeurs hautement reconfigurable

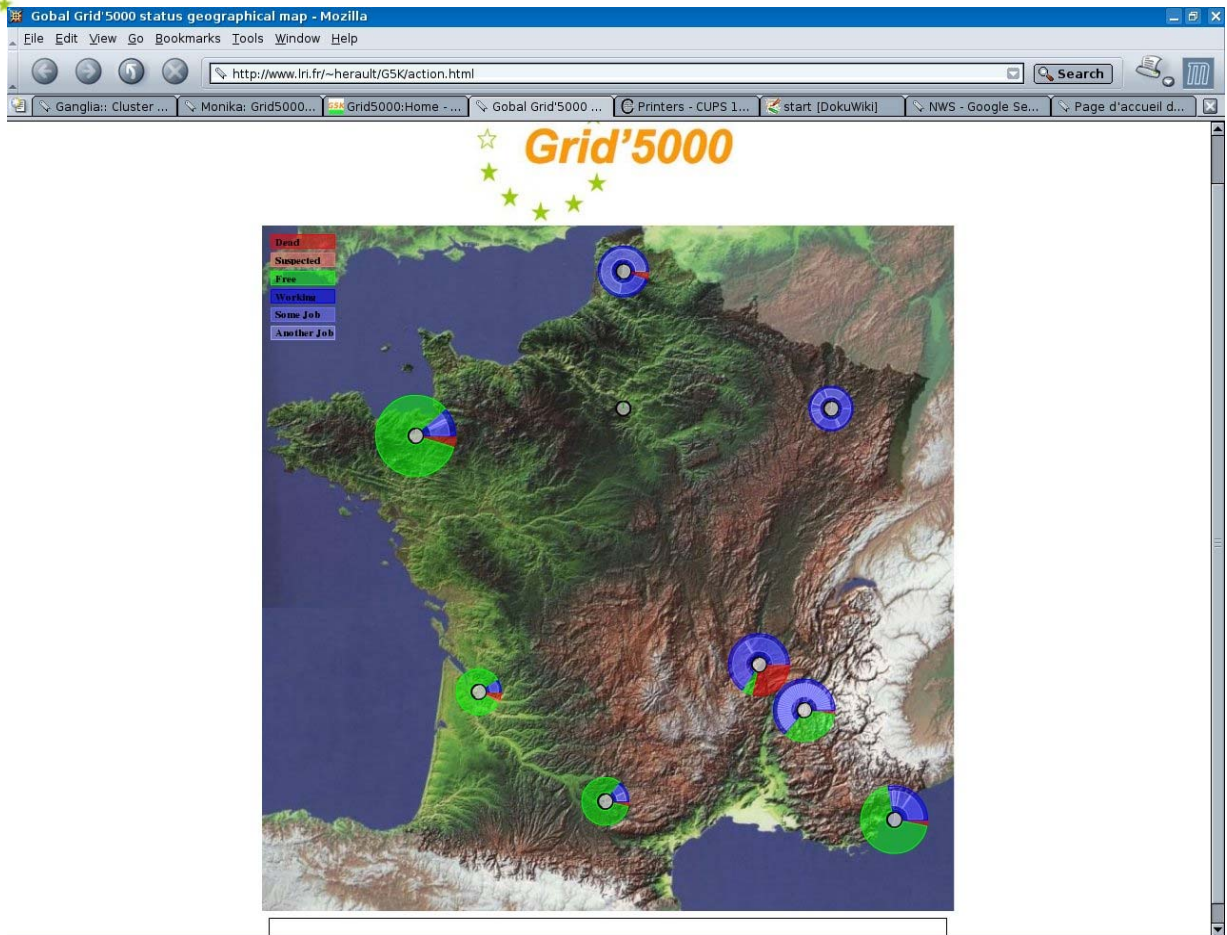


Objectifs conceptuels (2)

- Protéger les ressources des utilisateurs de la grille
 - ... malgré les possibilités de reconfiguration
 - ... environnement confiné
- Faciliter l'exploitation de la grille en mettant à la disposition des utilisateurs des outils de ...
 - connexion à la grille à partir de n'importe quel site
 - réservation des ressources
 - reconfiguration
 - monitoring

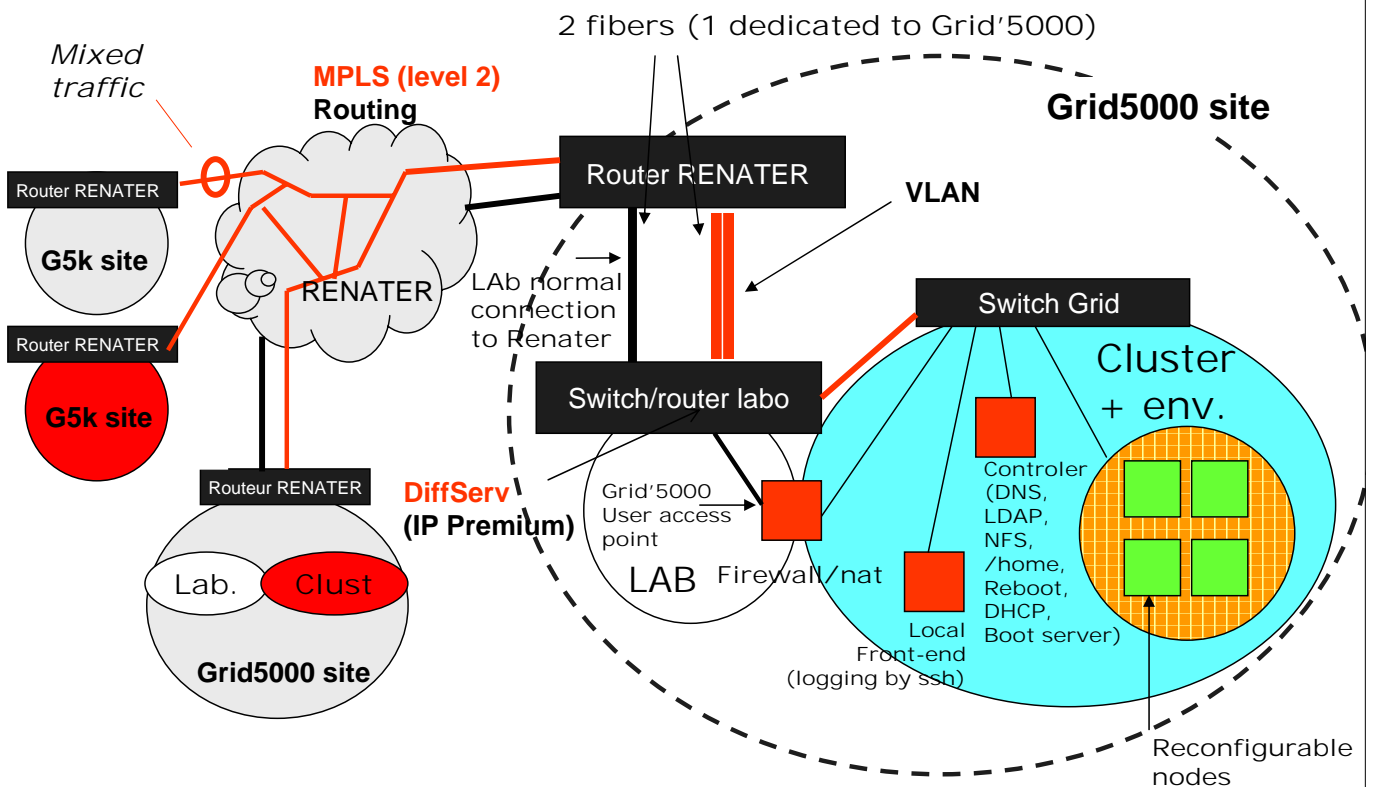
Grille hautement sécurisée et facilement exploitable

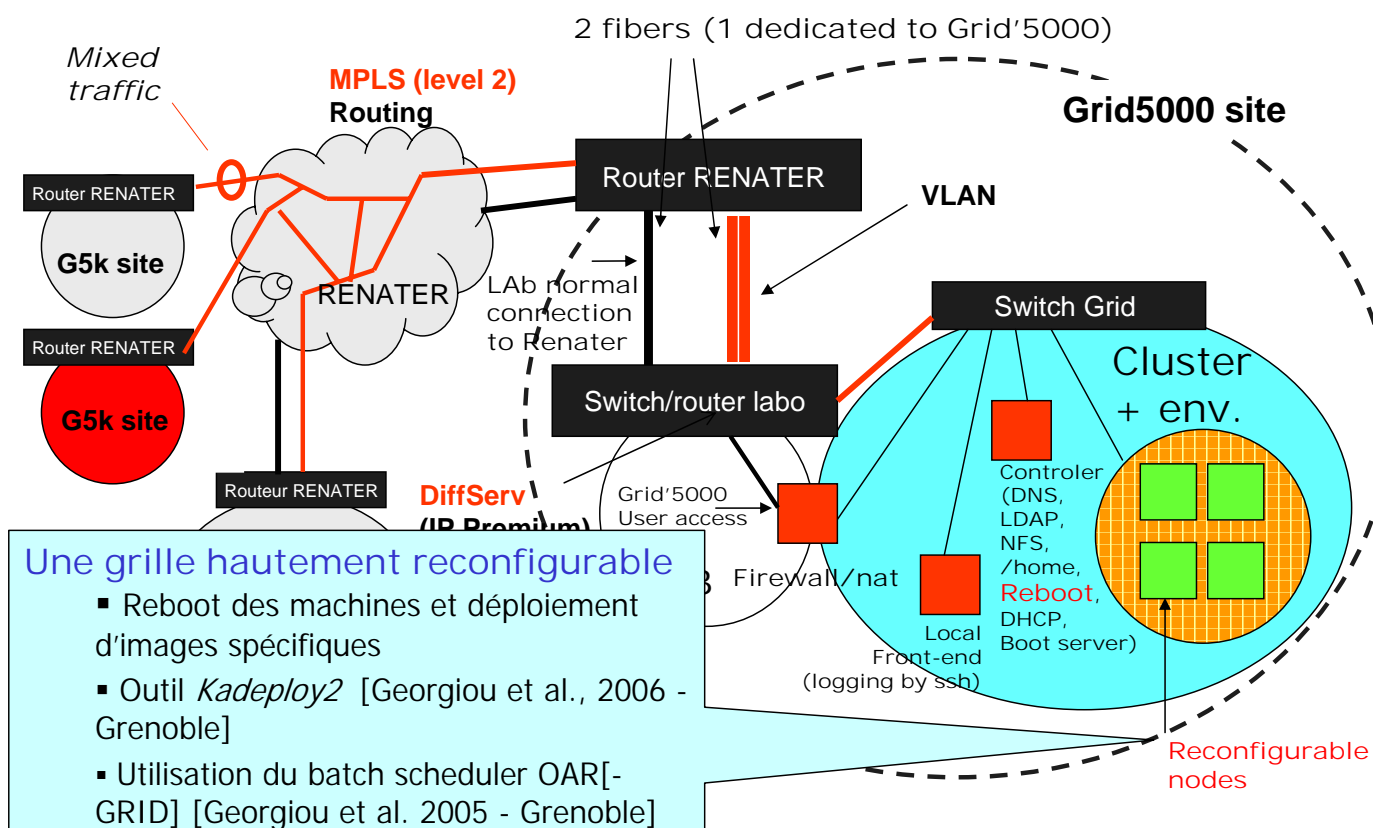
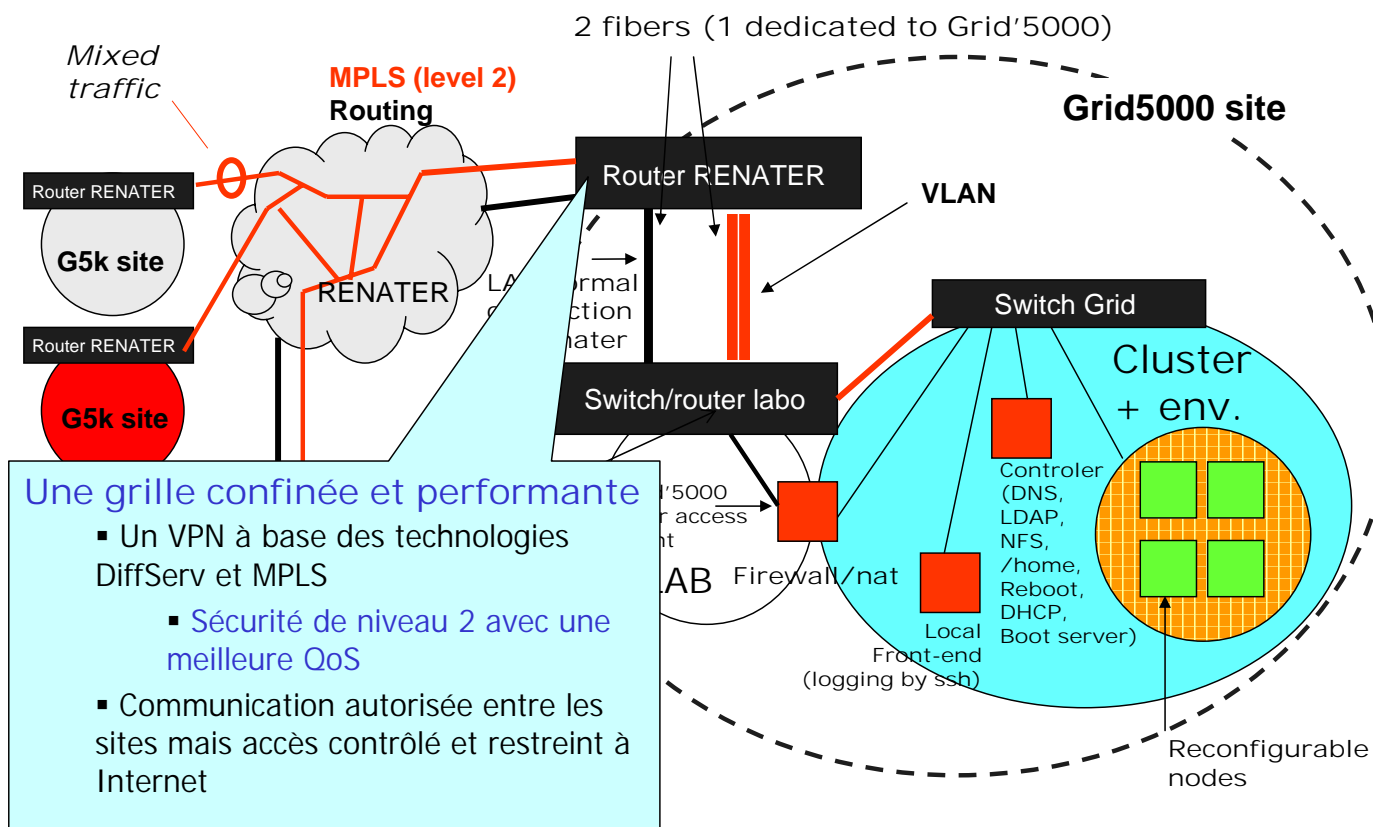
Grid'5000 Vue avec Google Earth ☺

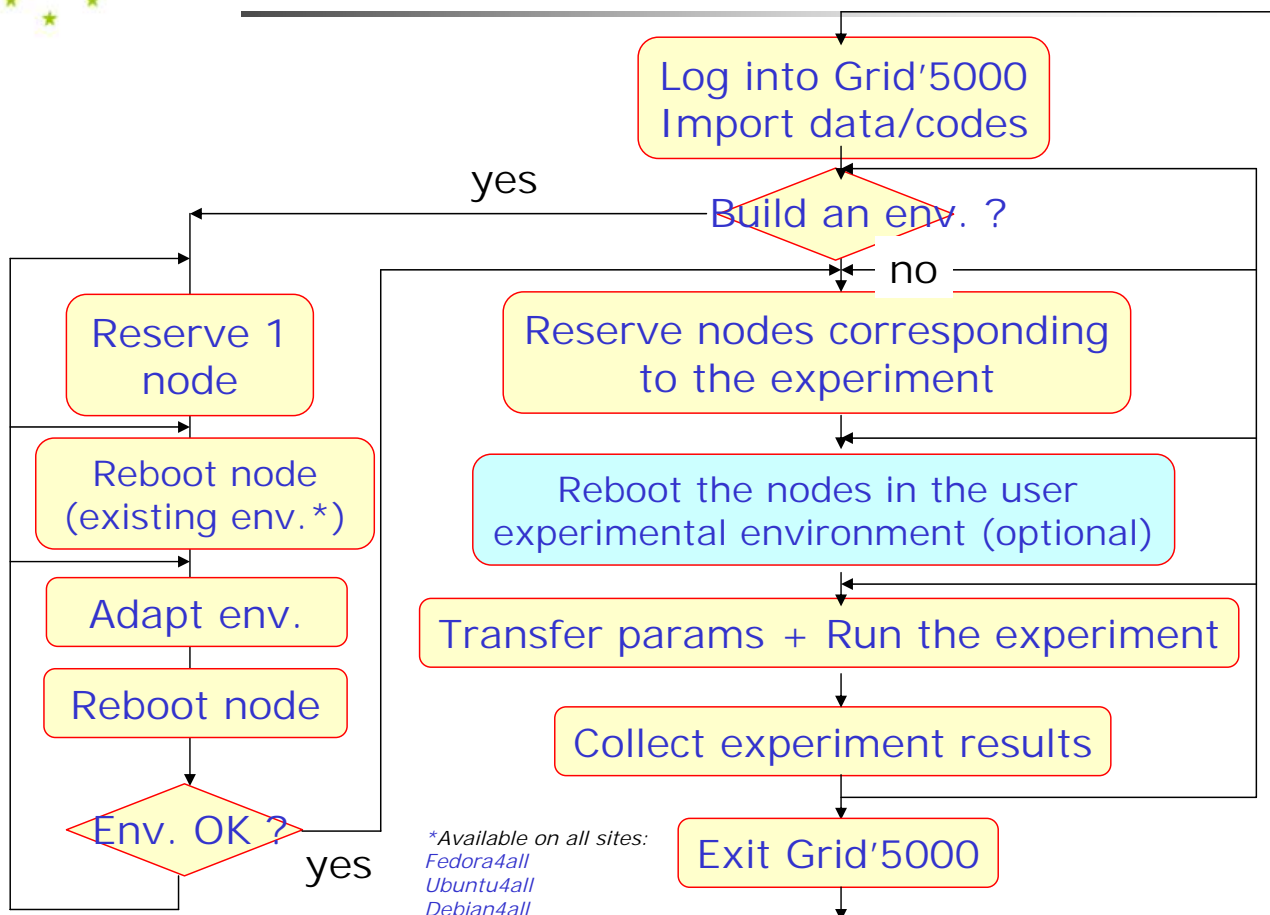
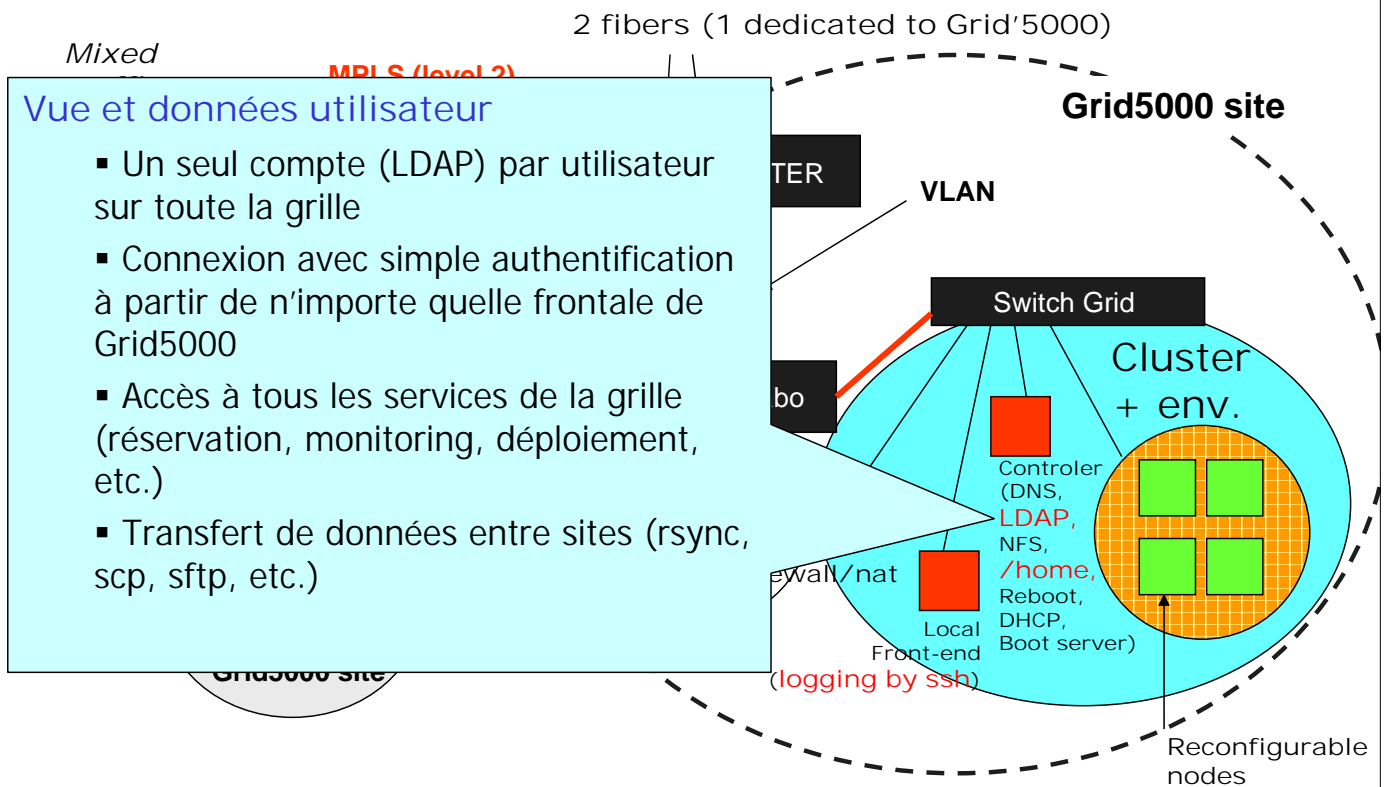


Grid'5000 Architecture de Grid5000 (1)

Franck Cappello



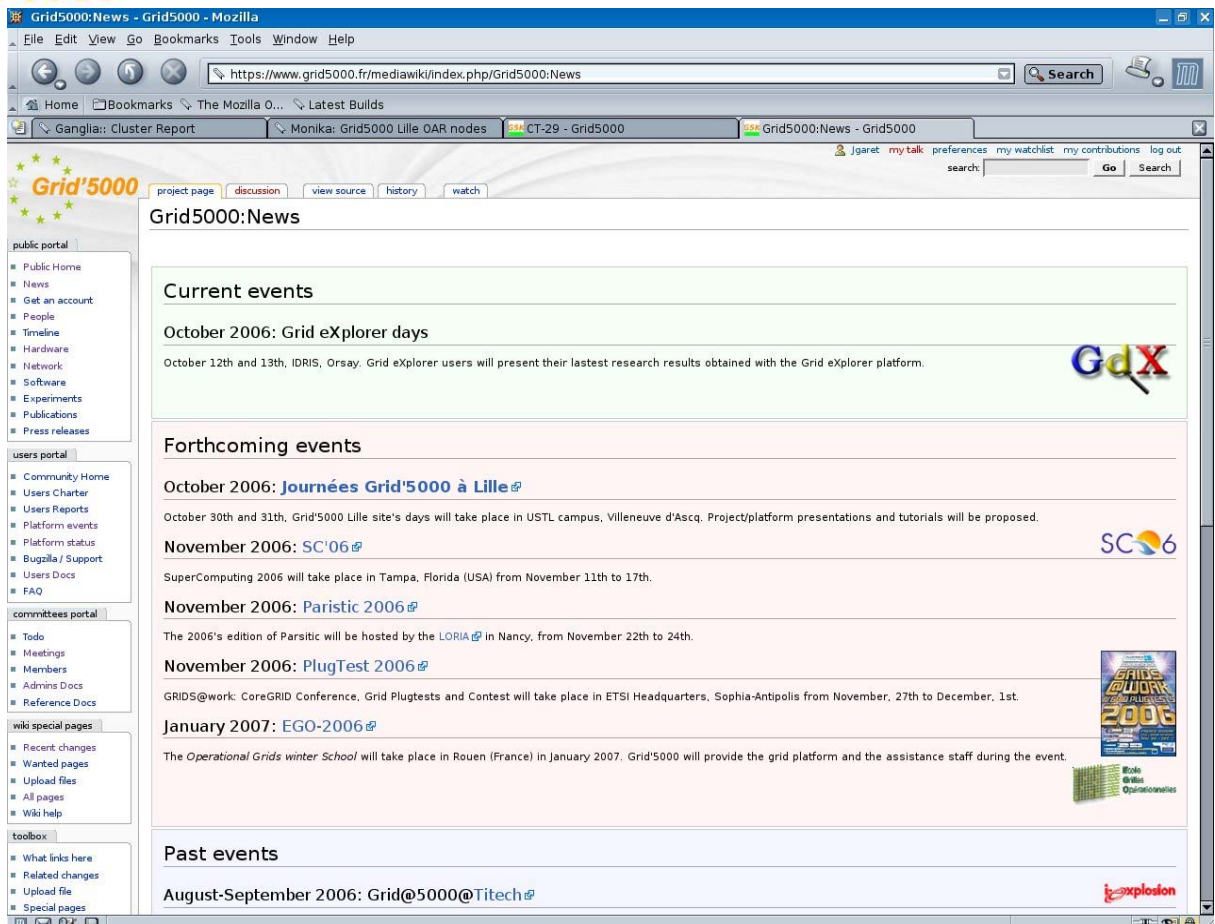




- Motivations et objectifs du projet
- Conception de la grille
- Etat actuel du projet
- Premiers résultats
- Perspectives du projet



Un site Wiki (<https://www.grid5000.fr>)



Grid5000:News

public portal

- Public Home
- News
- Get an account
- People
- Timeline
- Hardware
- Network
- Software
- Experiments
- Publications
- Press releases

users portal

- Community Home
- Users Charter
- Users Reports
- Platform events
- Platform status
- Bugzilla / Support
- Users Docs
- FAQ

committees portal

- Todo
- Meetings
- Members
- Admins Docs
- Reference Docs

wiki special pages

- Recent changes
- Wanted pages
- Upload files
- All pages
- Wiki help

toolbox

- What links here
- Related changes
- Upload file
- Special pages

Current events

October 2006: Grid eXplorer days

October 12th and 13th, IDRIS, Orsay. Grid eXplorer users will present their latest research results obtained with the Grid eXplorer platform.

Forthcoming events

October 2006: Journées Grid'5000 à Lille

October 30th and 31th, Grid'5000 Lille site's days will take place in USTL campus, Villeneuve d'Ascq. Project/platform presentations and tutorials will be proposed.

November 2006: SC'06

SuperComputing 2006 will take place in Tampa, Florida (USA) from November 11th to 17th.

November 2006: Paristic 2006

The 2006's edition of Parsitic will be hosted by the LORIA in Nancy, from November 22th to 24th.

November 2006: PlugTest 2006

GRIDS@work: CoreGRID Conference, Grid Plugtests and Contest will take place in ETSI Headquarters, Sophia-Antipolis from November, 27th to December, 1st.

January 2007: EGO-2006

The Operational Grids winter School will take place in Rouen (France) in January 2007. Grid'5000 will provide the grid platform and the assistance staff during the event.

Past events

August-September 2006: Grid@5000@Titech



Ressources matérielles

Hardware Abstract - Grid5000 - Mozilla

https://www.grid5000.fr/mediawiki/index.php/Special:HardwareAbstract

Platform status
Bugzilla / Support
Users Docs
FAQ

committees portal
Todo
Meetings
Members
Admins Docs
Reference Docs

wiki special pages
Recent changes
Wanted pages
Upload files
All pages
Wiki help

toolbox
Upload file
Special pages

Sites	Processors	AMD Opteron 246	AMD Opteron 248	AMD Opteron 250	AMD Opteron 252	AMD Opteron 275	Intel Itanium 2	Intel Xeon IA32	PowerPC	Site total
Bordeaux		96								96
Grenoble							206	64		270
Lille			106		40					146
Lyon	112			140						252
Nancy	94									94
Orsay	432			252						684
Rennes	198		128					128	64	518
Sophia	210					224				434
Toulouse										116
Processors total		1046	446	392	40	224	206	192	64	2610

Processor details

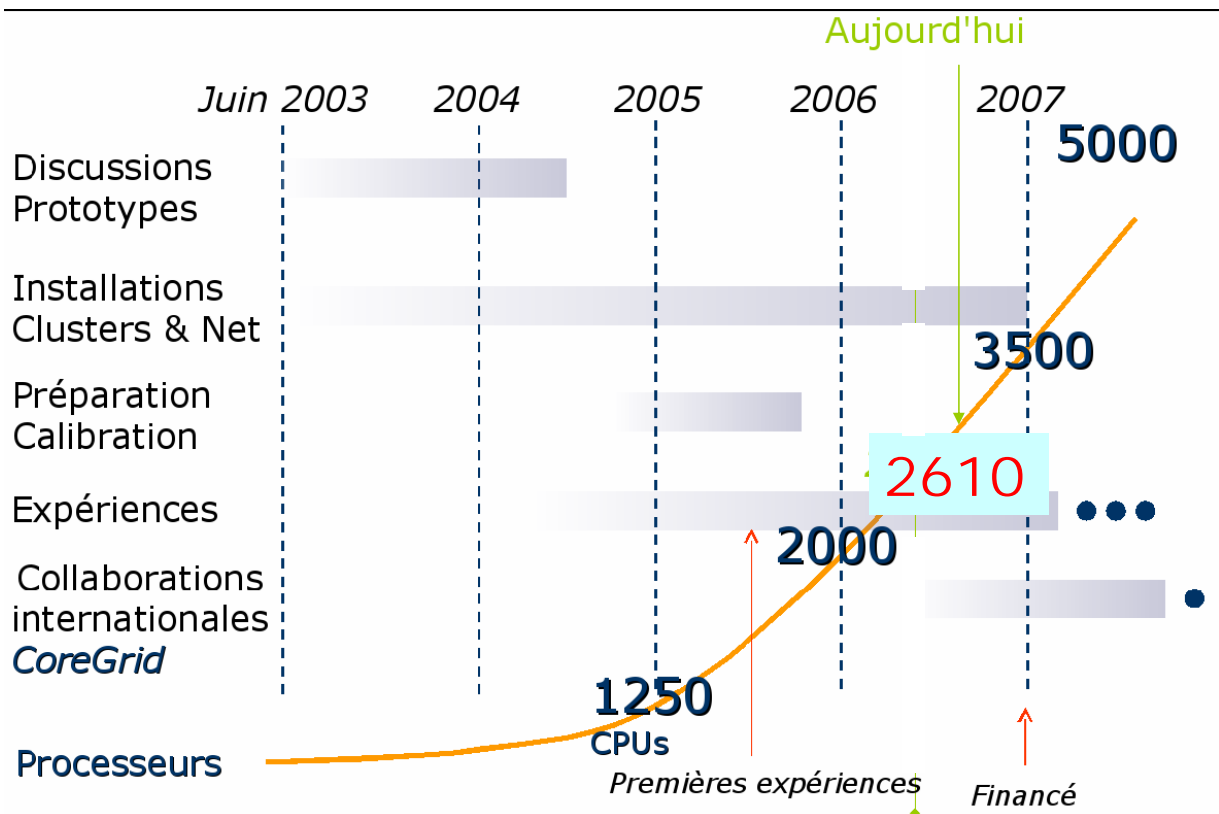
Sites \ Nodes	Apple Xserve	Dell PowerEdge	HP Integrity	HP ProLiant	IBM eServer	Sun Fire	Site total
Bordeaux					48		48
Grenoble		32	103				135
Lille					73		73
Lyon					56	70	126
Nancy					47		47
Orsay					342		342
Rennes	32	64		99		64	259
Sophia					105	56	161
Toulouse						58	58
Nodes total	32	96	103	146	624	248	1249

Node details

Sites \ Nodes	Apple Xserve G5	Dell PowerEdge 1600SC	Dell PowerEdge 1750	HP Integrity RX2600	HP ProLiant DL145G2	IBM eServer 325	IBM eServer 326	IBM eServer 326m	Sun Fire V20z	Sun Fire X4100	Site total
Bordeaux											48
Grenoble											135
Lille											73
Lyon											126
Nancy											47
Orsay											342
Rennes	32	64									259
Sophia											161
Toulouse											58
Nodes total	32	96	103	146	624	248					1249



Planning d'exécution





Grid'5000

Budget de Grid5000 à Lille

146 CPUs, 1 nœud de compilation, 1 nœud frontal, 1 nœud d'admin., 1 NAS 1To, 1 réseau Gbit Eth.

Source de financement	2004 (en K€TTC)	2005 (en K€TTC)	2006 (en K€TTC)	2007 (en K€TTC)	Total / source (en K€TTC)
ACI GRID	60	60	60	20	200
CNRS - STIC		60			60
CNRS - LIFL		15			15
USTL - BQR		50			50
Génopole Lille		25			25
INRIA Futurs			300	240	540
Total / année	60	210	360	260	890

+ 79 K€ TTC – Région Nord-Pas-de-Calais + INRIA Futurs

Financement d'un Ingénieur INRIA Expert Grid5000 en CDD de 2 ans

Total : 969 K€ TTC



Grid'5000

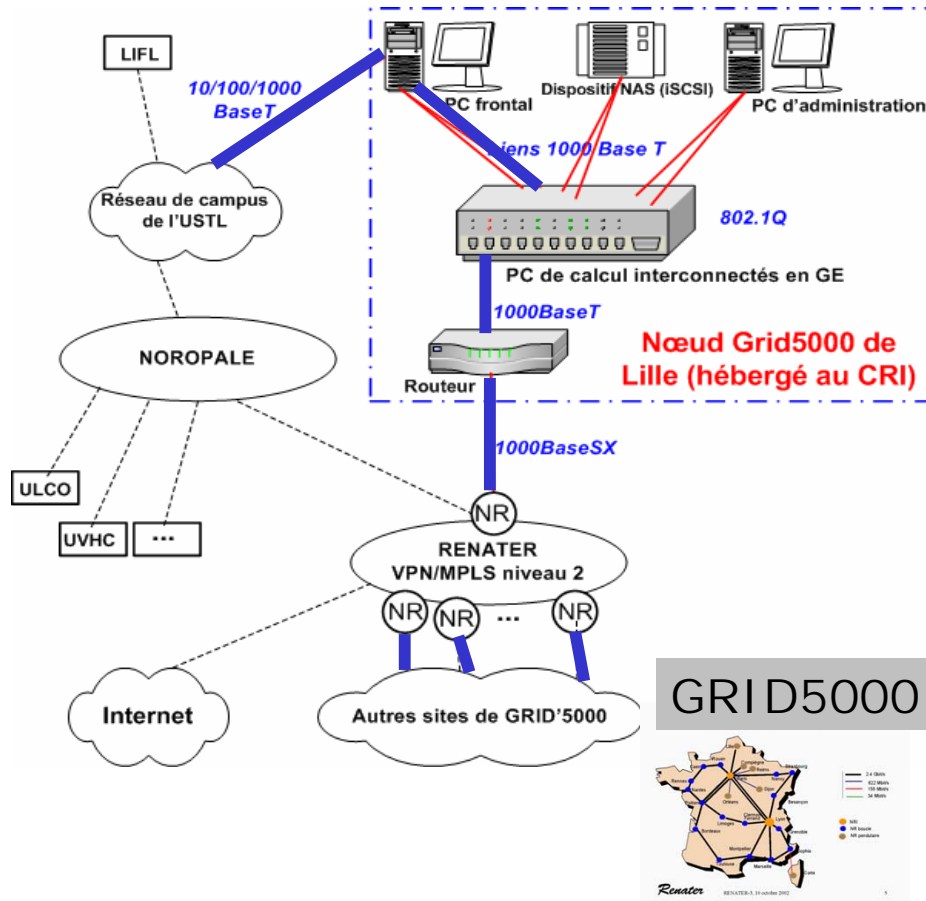
Grappe de Lille

Hébergée au CRI de Lille

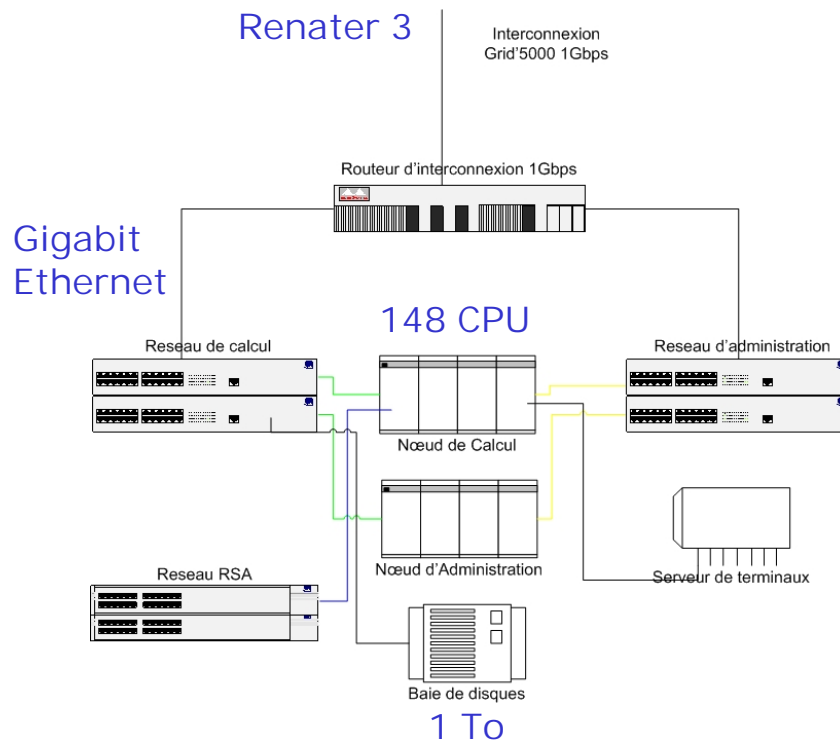




Grappe de Lille connectée à Grid5000

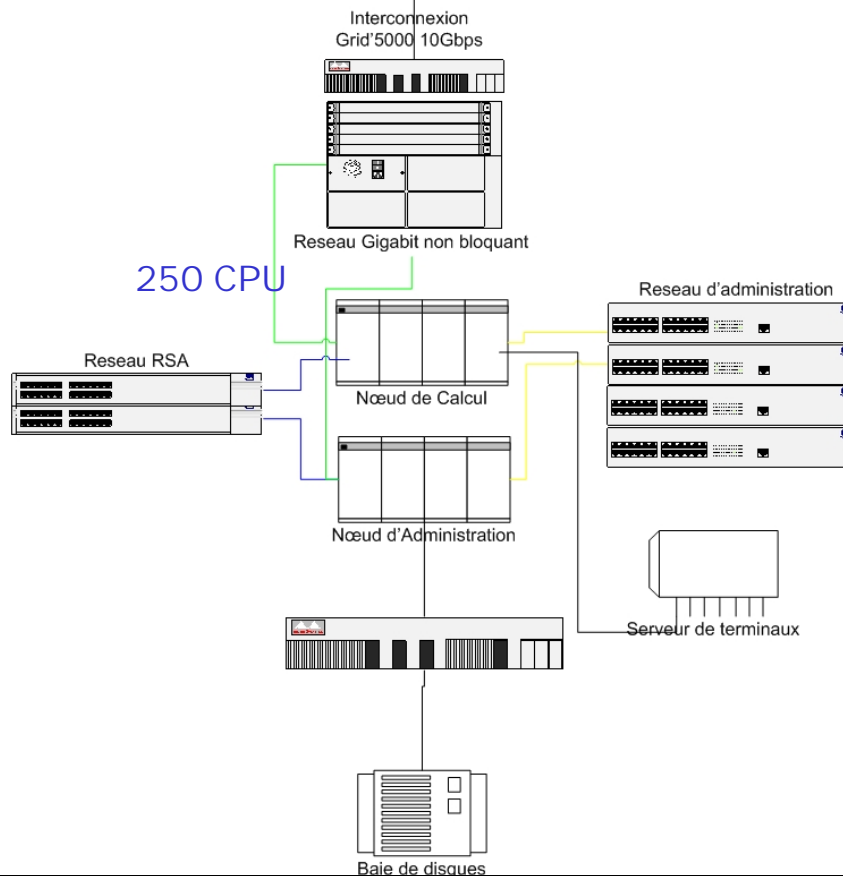


Grappe de Lille (état actuel)



Grappe de Lille (Fin 2006)

Renater 4



Ressources humaines

Grid5000:People - Grid5000 - Mozilla

https://www.grid5000.fr/mediawiki/index.php/Grid5000:People

Grid'5000 project page | discussion | view source | history | watch

Grid5000:People

Director

Firstname	Lastname	Institute
Franck	Cappello	INRIA, PCRI, LRI Grid'5000 (Director and Head of Steering Committee)

Steering Committee

Firstname	Lastname	Institute	In charge of
Brigitte	Plateau	INRIA, ID IMAG	ACI Grid Scientific Committee (Chairwoman)
Thierry	Priol	INRIA, IRISA	ACI Grid (Director)
Dany	Vandromme	RENATER	RENATER (Director)
Franck	Cappello	INRIA, PCRI, LRI	Grid'5000 Orsay
Michel	Dayde	IRIT, ENSEEIHT	Grid'5000 Toulouse
Frédéric	Desprez	INRIA, LIP	Grid'5000 Lyon
Emmanuel	Jeannot	LORIA	Grid'5000 Nancy
Yvon	Jegou	INRIA, IRISA	Grid'5000 Rennes
Stéphane	Lanteri	INRIA Sophia	Grid'5000 Nice-Sophia
Noureddine	Melab	LIFL	Grid'5000 Lille
Raymond	Namyst	INRIA, LABRI	Grid'5000 Bordeaux
Martin	Quinson	LORIA	Grid'5000 Nancy
Olivier	Richard	INRIA, ID IMAG	Grid'5000 Grenoble
Pascale	Vicat-Blanc Primet	INRIA, LIP	Grid'5000 Lyon

Technical Committee

Firstname	Lastname	Institute
Orsay		
Franck	Cappello	INRIA, PCRI, LRI
Pierre	Neyron	INRIA Futurs
Pierre	Sens	LIP6
Grenoble		
Olivier	Richard	ID IMAG, INRIA
Lyon		

- Définir et veiller à l'application de la politique de fonctionnement de la grille
- Assurer la pérennité du projet (demande de moyens financiers, ...)
- Mise en œuvre et maintenance technique de la grille, formation des utilisateurs, ...



... sur Lille

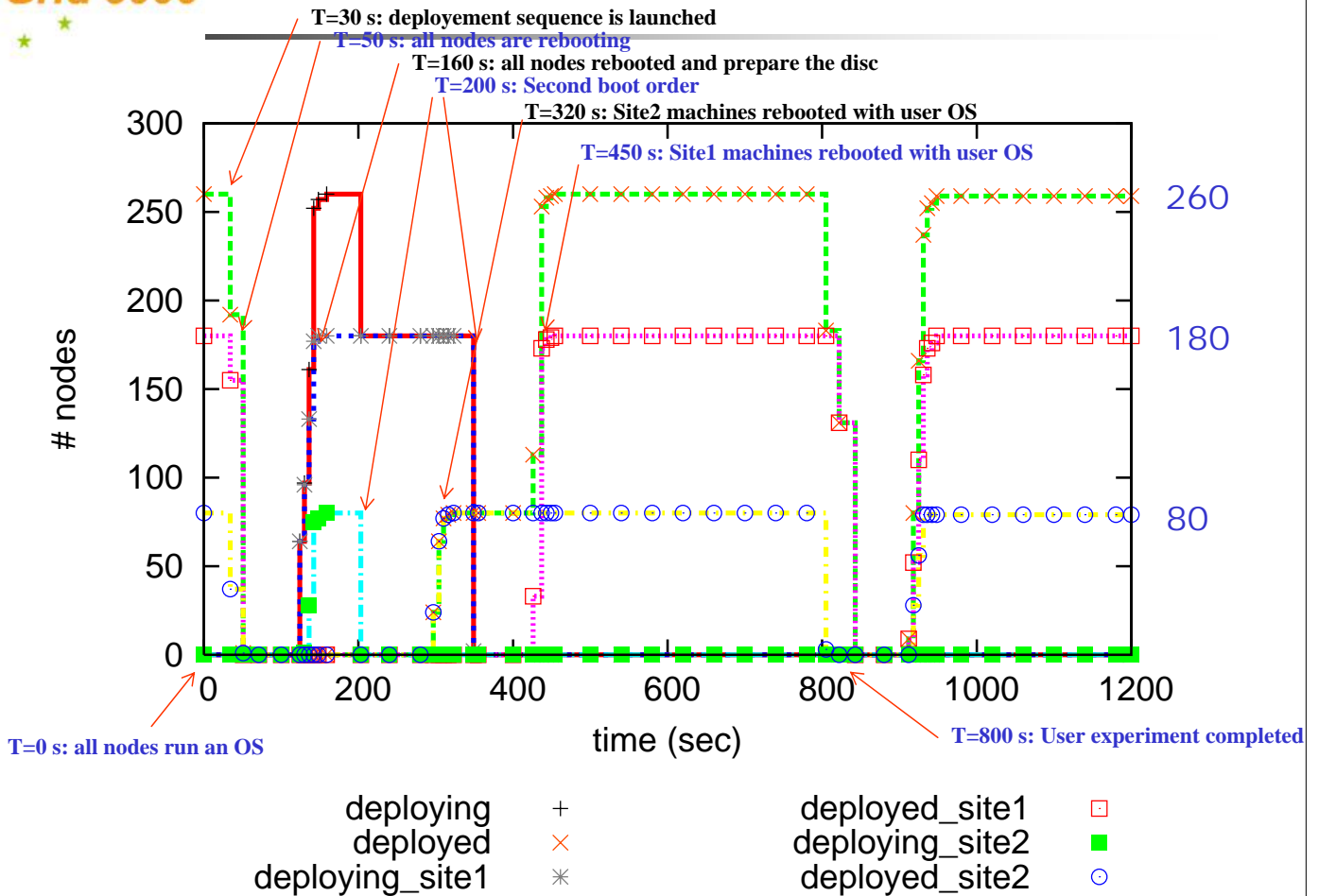
The screenshot shows a Mozilla browser window displaying the Grid'5000 wiki page for 'Lille:People'. The browser's address bar shows the URL: <https://www.grid5000.fr/mediawiki/index.php/Lille:People>. The page content includes:

- Lille:People** [edit]
- Steering committee** [edit]
 - Nordine Melab [edit]
- Technical committee** [edit]
- Head** [edit]
 - Julien Garete [edit]
- Members** [edit]
 - Julien Garete [edit]
 - Nordine Melab [edit]
- Contributors** [edit]
 - Zouhir Hafidi [edit]

The left sidebar contains several navigation menus: 'public portal', 'users portal', 'committees portal', 'wiki special pages', and 'toolbox'.

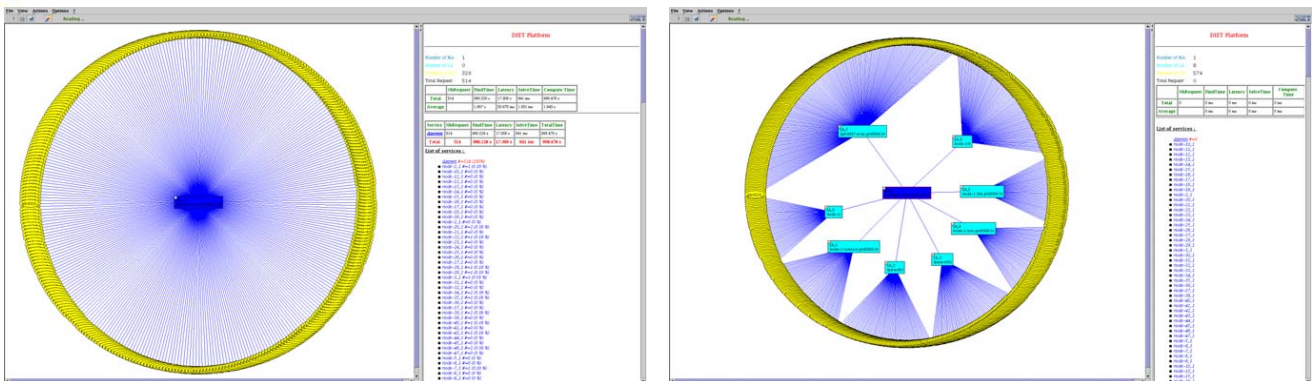
Grid'5000 Plan

- Motivations et objectifs du projet
- Conception de la grille
- Etat actuel du projet
- Premiers résultats
- Perspectives du projet



DIET à grande échelle

Distributed Interactive Engineering Toolbox



- Plus de 45.000 requêtes GridRPC (de type multi. de matrices dgemm) soumises par 1120 clients à des serveurs distants
- La charge de l'ordonnancement des requêtes est distribuée sur l'ensemble des agents DIET

- 7 sites : Lyon, Orsay, Rennes, Lille, Sophia-Antipolis, Toulouse, Bordeaux
- 8 grappes - 585 machines - 1170 CPUs

Objectifs

- Prouver l'extensibilité de DIET
- Tester les fonctionnalités de DIET à une grande échelle



Solving the Challenging Flow-Shop Scheduling Problem on Grid'5000

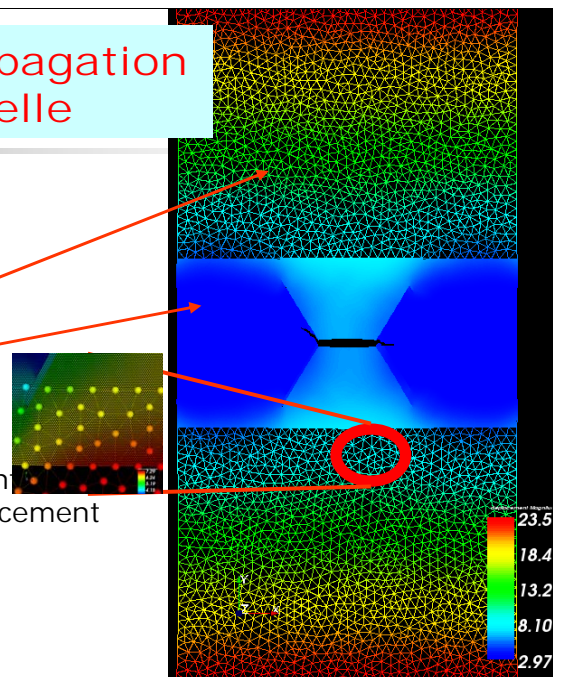
M. Mezmaiz, N. Melab and E-G. Talbi
CNRS/LIFL – INRIA DOLPHIN, Université de Lille1

- **The Flow-Shop scheduling problem**
 - N jobs to be scheduled on M machines
 - The objective is to minimize the total completion time
- **The proposed grid-enabled branch and bound**
 - Optimized coding of the search tree
 - Adaptive load balancing to explore an irregular B&B tree in a heterogeneous context
 - Checkpointing-based fault tolerance to support the dynamic availability of resources
 - Termination detection of asynchronous B&B processes
 - Global information sharing in a large scale environment
- **Experimentation on Grid5000 – a success story**
 - Standard Taillard's benchmark: **Ta056 - 50 jobs on 20 machines**
 - Best known solution: **3681, Ruiz & Stutzle, 2004**
 - **Exact solution: 3679, Mezmaiz, Melab & Talbi, 2006**

Running wall clock time: 25 days 46 min.	CPU time on a single processor: 22 years 185 days 16 hours
Avg. number of exploited processors: 328	Maximum number of exploited processors: 1 195
Parallel efficiency: 97 %	Bordeaux (88), Orsay (360), Sophia (190), Lille (98), Toulouse (112), Rennes (456), Univ. Lille1 (304)

Suivi temps réel avec EPSN d'une propagation de fissure par une méthode multi échelle

- **Modèles**
 - **Modèle macroscopique**
 - Élasticité non linéaire / éléments finis
 - **Modèle microscopique**
 - Dynamique moléculaire
 - **Zone mixte**
 - Pondération de l'énergie dans le recouvrement
 - Contrainte pour recoller les vitesses de déplacement
 - **Couplage de codes**
 - LAMMPS (Dynamique moléculaire –Sandia)
 - Élasticité basée sur LibMesh
- **Simulation**
 - 1/2 million d'atomes, 100 processeurs – Cluster Bordelais
 - 5 000 éléments finis, 20 processeurs – Cluster Rennais
 - Visualisation, 4 processeurs – Cluster Bordelais
- **Pilotage**
 - Monitoring : énergie, température, ...
 - Visualisation en ligne
 - Modification de paramètres : conditions aux limites, fréquence de sauvegarde, ...





Grid'5000 Autres expériences

- Peer-to-Peer for Computational Grids: Mixing Clusters and Desktop Machines – 628 machines, 5+1 sites
- Challenge: use P2P mechanisms to build grid middleware – 580 machines, 6 sites
- P2P experiments (H. Bal, NL) – Up to 960 machines, 6 sites
- Protein folding and docking on the Grid – Up to 520 machines, 4 sites
- Fully Distributed Batch Scheduler – 944 machines, 4 sites (6 clusters)
- Fault tolerant MPI for the Grid – 536 machines, 6 sites
- Fractal Deployment Framework – 1000 nodes
- Grid@work (Octobre 10-14 2005) - 1600 CPUs in total: 1200 provided by Grid'5000 + 50 by the other Grids (EGEE, DEISA, NorduGrid) + 350 CPUs on clusters
- ...



Grid'5000 Plan

- Motivations et objectifs du projet
- Conception de la grille
- Etat actuel du projet
- Premiers résultats
- Perspectives du projet



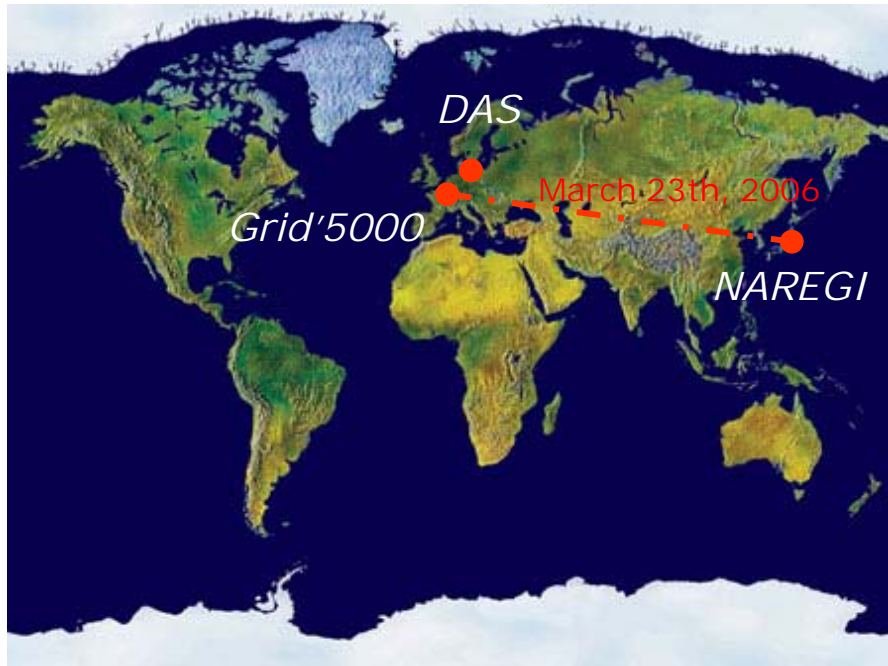
Grid'5000 Conclusions (1)

- Une grille hautement reconfigurable et sécurisée (originale) opérationnelle
 - 2610 processeurs, 2,5 To de mémoire, 150 To de disque, réseaux haut débit, dorsale à 10Gb
 - Différents outils : réservation, déploiement, etc.
- Plusieurs expérimentations à grande échelle ...
 - ... avec plus de 1000 processeurs sur plus de 3 sites
 - ... à différents niveaux : application, algorithmique/programmation, middleware/système, réseau
- Production scientifique
 - Projets ACI (GRID, Masse de données, Sécurité), ANR CIGC
 - Plus de 200 publications
 - Plusieurs thèses (Doctorat et HDR), projets Post-Doc, Master, intership, ...



Grid'5000 Conclusions (2)

- Une très bonne organisation
 - Un comité de pilotage, un comité technique, une communauté d'utilisateurs
 - Réunions régulières des comités
 - Communication et visibilité
 - Organisation de conférences (EXPGRID/HPDC, IPDPS, CCGRID, ...) et écoles (GridUse, Grid@work, Grid'5000, etc.)
 - Stands dans différentes conférences (SuperComputing, ...), ...
 - Présence au Global Grid Forum
 - Direction et présence dans de nombreux projets Européens (CoreGrid)
 - Projets communs avec le Japon (financements INRIA, CNRS-JST, Sakura, etc.)
 - Projets avec les USA (équipes associées INRIA, NSF/INRIA)



- NAtional REsearch Grid Intiative (NAREGI) – Japan (11 sites)
- Distributed ASCII Supercomputer 2 (DAS2) – Pays Bas (5 sites)

