



A nation wide Experimental Grid

Franck Cappello
INRIA

fci@lri.fr

With all participants



Action Concertée Incitative [ACI]
Globalisation des Ressources Informatiques
et des Données [GRID]





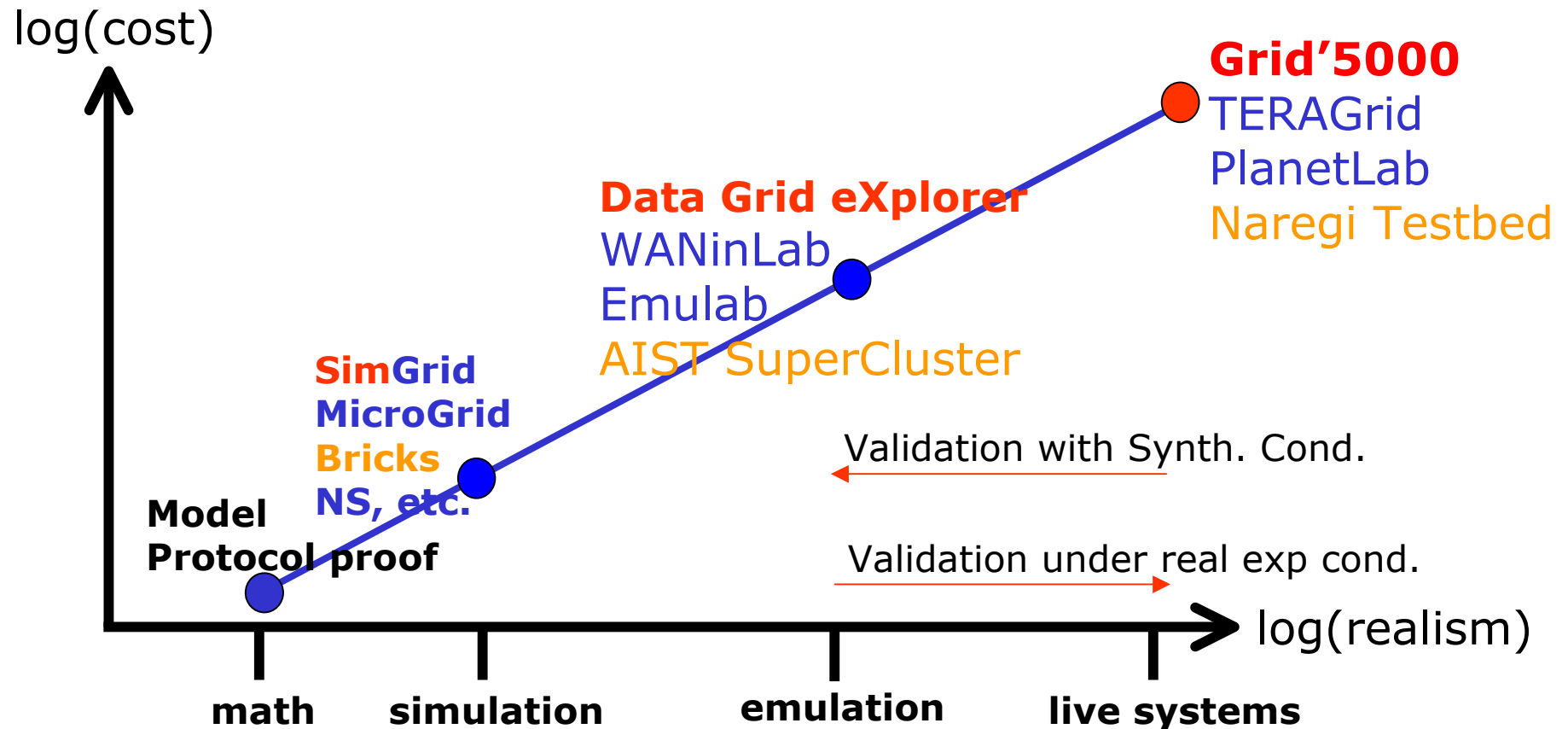
We need Grid experimental tools

According to the current knowledge:

There is no large scale testbed dedicated to Grid experiments

→ Grid'5000 as a real life system

→ Data Grid eXplorer as a large scale emulator





Grid'5000:

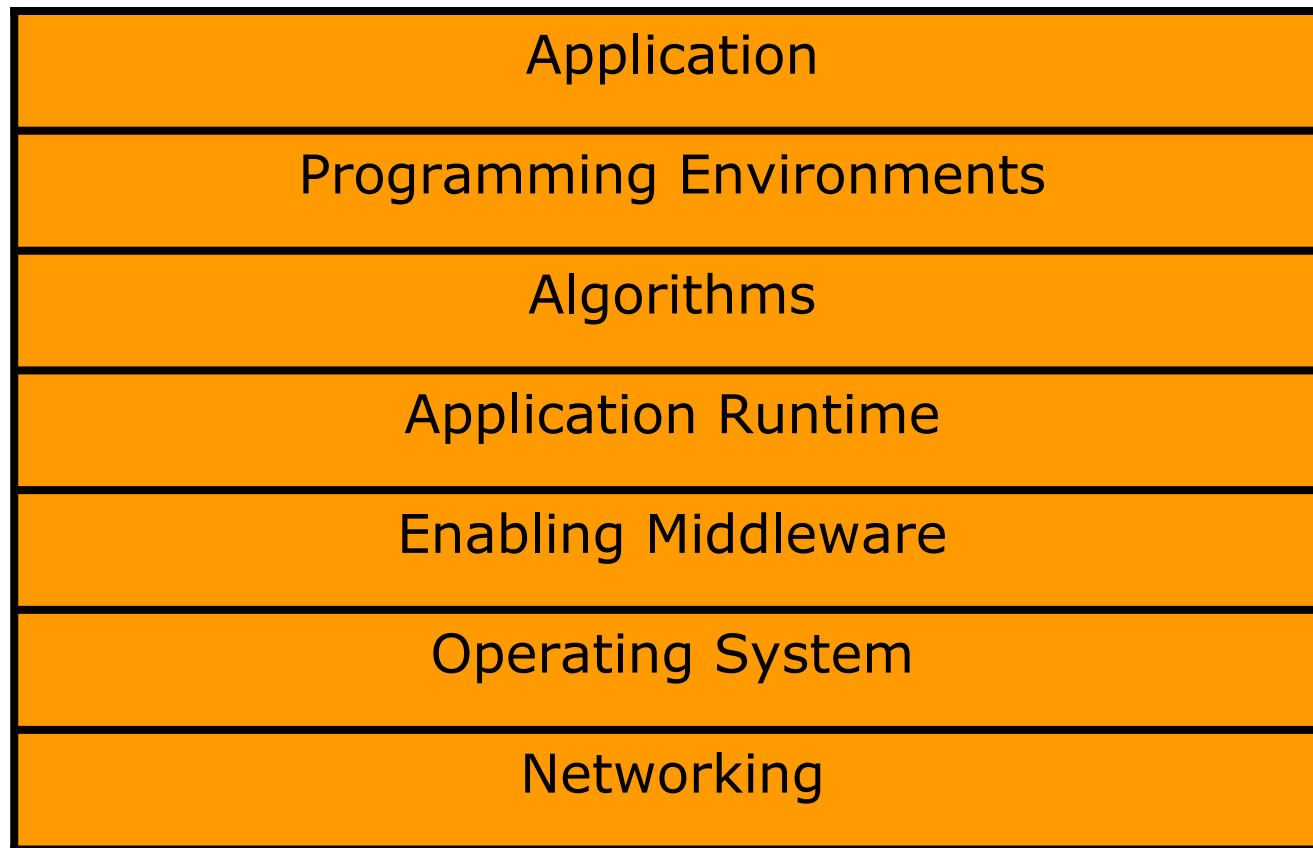
Collection of experiments to be done

- **Networking**
 - End host communication layer (interference with local communications)
 - High performance long distance protocols (improved TCP)
 - High Speed Network Emulation
- **Middleware / OS**
 - Scheduling / data distribution in Grid
 - Fault tolerance in Grid
 - Resource management
 - Grid SSI OS and Grid I/O
 - Desktop Grid/P2P systems
- **Programming**
 - Component programming for the Grid (Java, Corba)
 - GRID-RPC
 - GRID-MPI
 - Code Coupling
- **Applications**
 - Multi-parametric applications (Climate modeling/Functional Genomic)
 - Large scale experimentation of distributed applications (Electromagnetism, multi-material fluid mechanics, parallel optimization algorithms, CFD, astrophysics)
 - Medical images, Collaborating tools in virtual 3D environment



Grid'5000 Design goal:

Experimenting all layers of the Grid software stack





Grid'5000 required features

- 1) Remotely controllable Grid nodes installed in geographically distributed laboratories
- 2) A « Controllable » and « Monitorable » Network between the Grid nodes → (may be unrealistic in some cases)
- 3) A middleware infrastructure connecting allowing users to access the Grid sites
- 4) A playground to prepare experiments
- 5) A toolkit to deploy, run, monitor, control experiments and collect results



Funding & Participants

Funding:

- 1) ACI GRID (Hardware)
- 2) INRIA (Hardware, Engineers)
- 3) CNRS (AS, Engineers, etc.)
- 4) Regional councils (Hardware)

Steering Committee (11) :

- Franck Cappello** (animateur)
- Thierry Priol** (Directeur directeur ACI Grid)
- Brigitte Plateau** (Directrice CS ACI Grid)
- Dany Vandrome** (Renater)
- Frédéric Desprez (Lyon)
- Michel Daydé (Toulouse)
- Yvon Jégou (Rennes)
- Stéphane Lantéri (Sophia)
- Raymond Namyst (Bordeaux)
- Pascale Primet (Lyon)
- Olivier Richard (Grenoble)

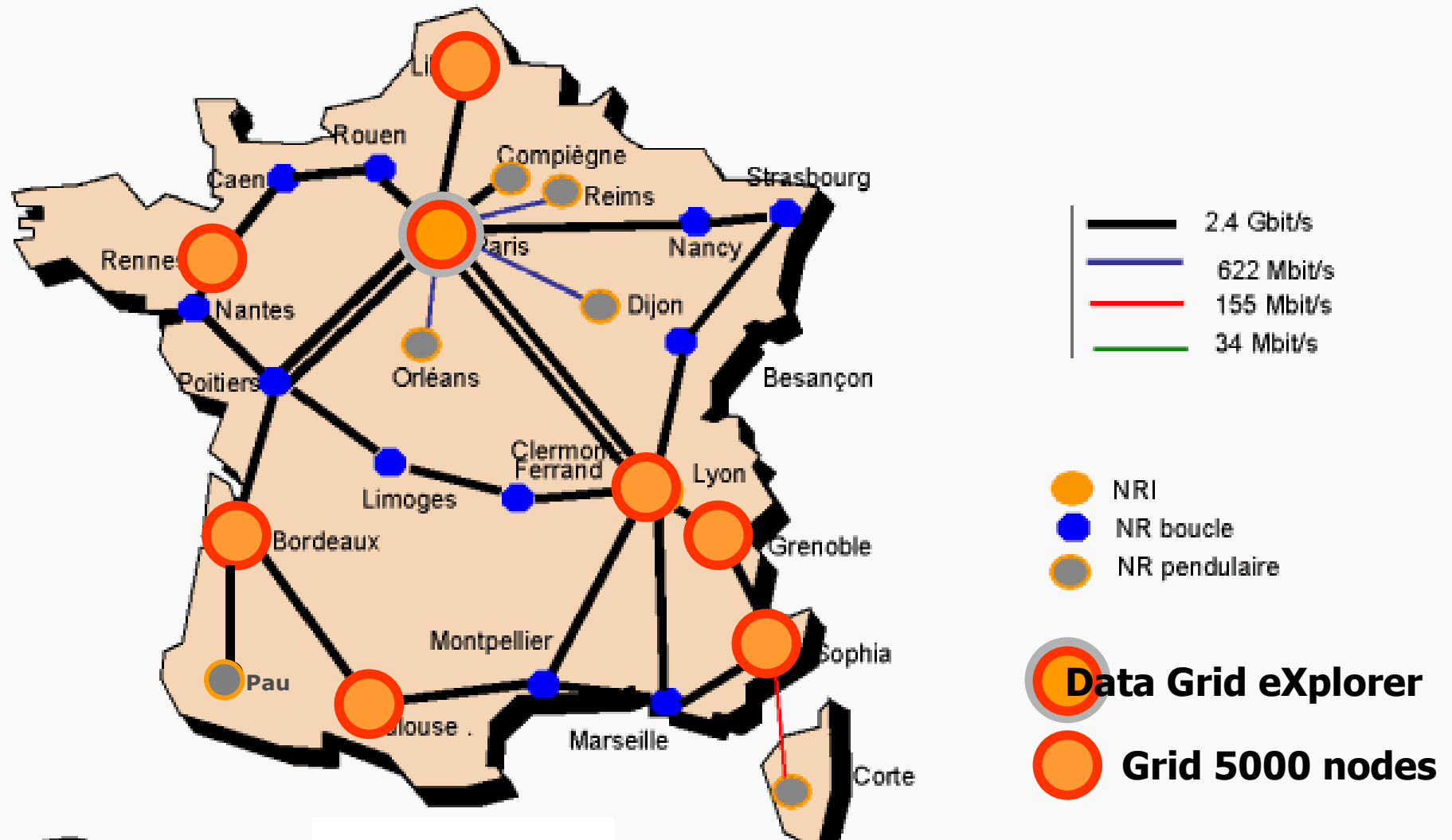
Technical Committee (28) :

Jean-Luc ANTHOINE
Jean-Claude Barbet
Pierrette Barbaresco
Nicolas Capit
Eddy Caron
Christophe Cérin
Olivier Coulaud
Georges Da-Costa
Yves Denneulin
Benjamin Dexheimer
Aurélien Dumez
Gilles Gallot
David Geldreich
Sébastien Georget
Olivier Gluck
Claude Inglebert
Julien Leduc
Cyrille Martin
Jean-Francois Méhaut
Jean-Christophe Mignot
Thierry Monteil
Guillaume Mornet
Alain Naud
Vincent Néri
Gaetan Peaquin
Franck Simon
Sebastien Varrette
Jean-Marc Vincent



Grid'5000

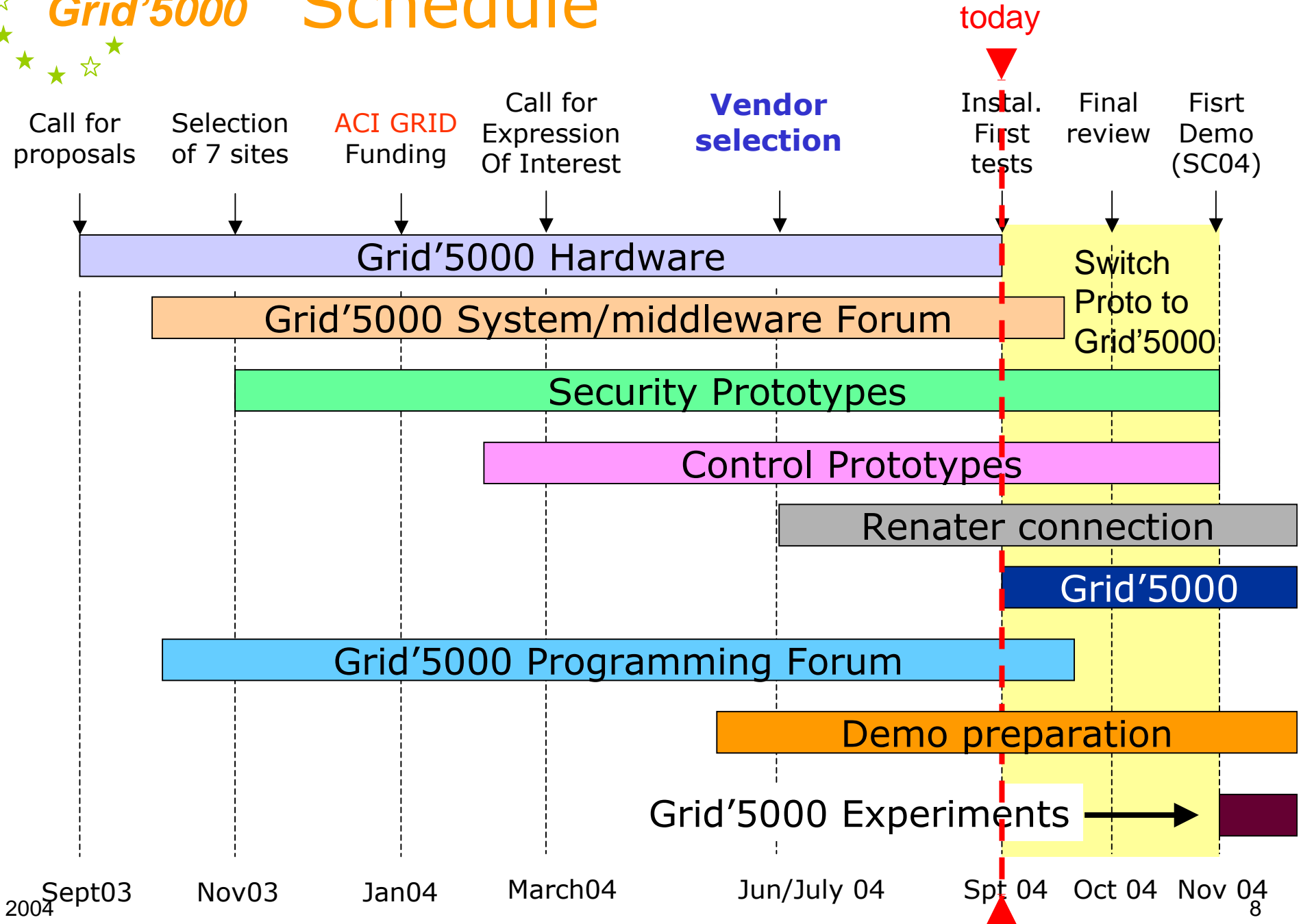
Grid'5000 in December'2004



Renater 3 (soon 4) @ 1 Gb/sec ; may evolve to 10 Gb/s



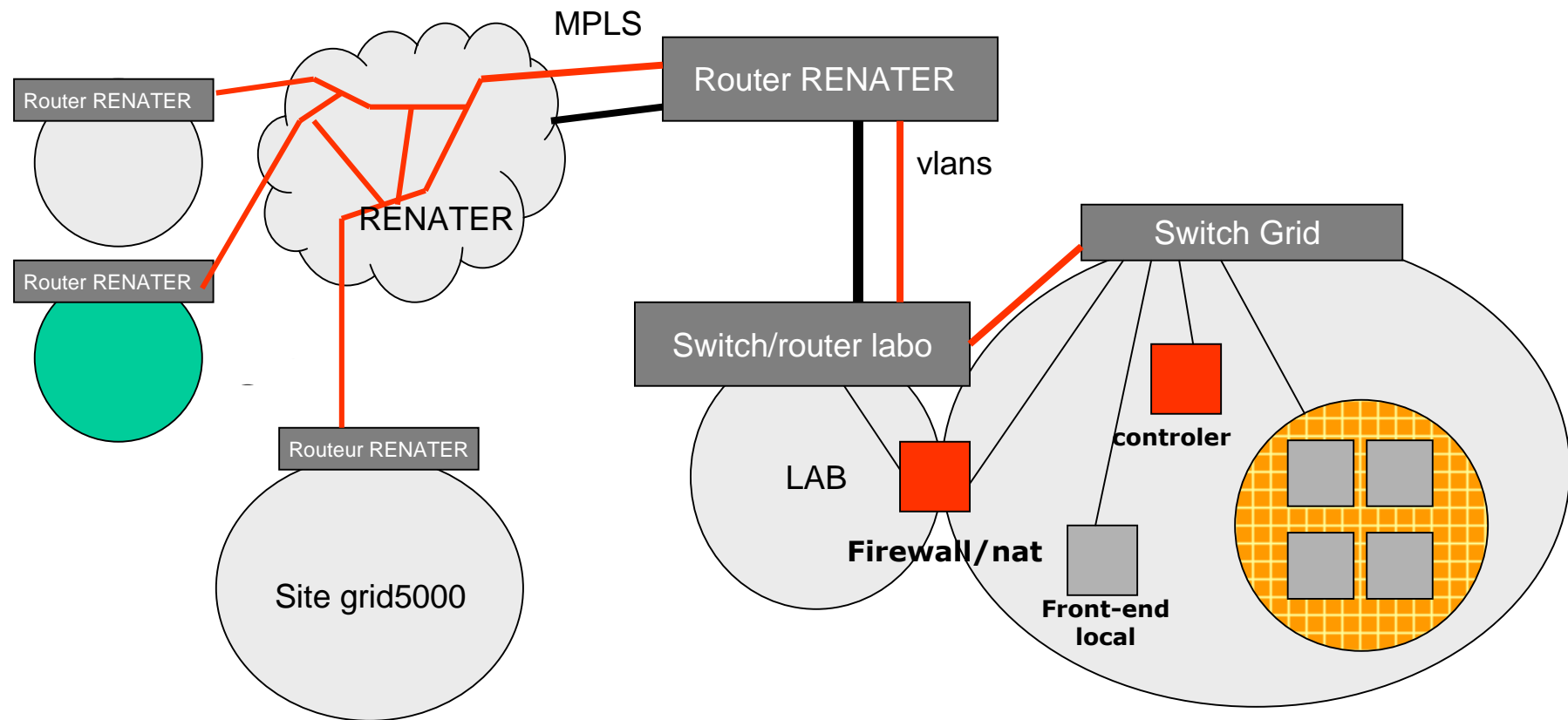
Grid'5000 Schedule





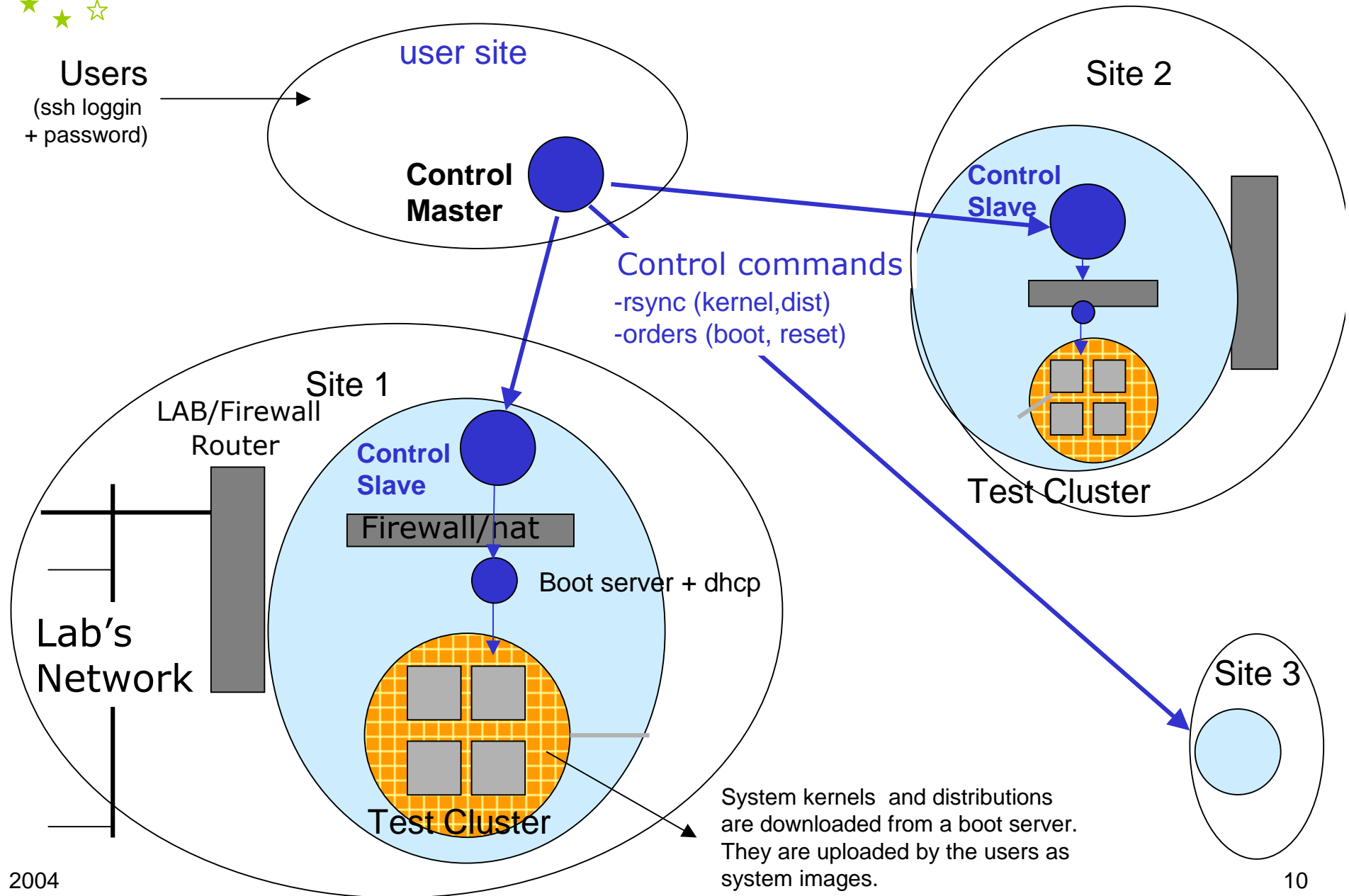
Grid'5000 Security architecture

G5k is a closed system





Grid'5000 Control Architecture





Grid'5000 prototype

Nagios - Microsoft Internet Explorer

Fichier Edition Affichage Favoris Outils ?

Nagios

Supress popups: Update

General

- Home
- Documentation

Monitoring

- Tactical Overview
- Service Detail
- Host Detail
- Status Overview
- Status Summary
- Status Grid
- Status Map
- 3-D Status Map
- Service Problems
- Host Problems
- Network Outages
- Comments
- Downtime
- Process Info
- Performance Info
- Scheduling Queue

Reporting

- Trends
- Availability
- Alert Histogram
- Alert History
- Alert Summary
- Notifications
- Event Log

Configuration

- View Config

```
graph TD; Nagios((Nagios Process)); gw_sophia((gw_sophia Up)); gw_orsay((gw_orsay Up)); gw_nancy((gw_nancy Up)); gw_lyon((gw_lyon Up)); gw_grenoble((gw_grenoble Up)); gw_bordeaux((gw_bordeaux Up)); gw_belfort((gw_belfort Up)); gw_toul((gw_toul Up)); node06_001((node06-001 Up)); node06_002((node06-002 Up)); node06_003((node06-003 Up)); node06_004((node06-004 Up)); node06_005((node06-005 Up)); node06_006((node06-006 Up)); node06_007((node06-007 Up)); node06_008((node06-008 Up)); node90_002((node90-002 Up)); node90_003((node90-003 Up)); node33_002((node33-002 Up)); node38_001((node38-001 Up)); node69_002((node69-002 Up)); node54_002((node54-002 Up)); node54_003((node54-003 Up)); node54_004((node54-004 Up)); node54_005((node54-005 Up)); node54_006((node54-006 Up)); node54_007((node54-007 Up)); node54_00a((node54-00a Up)); node54_00b((node54-00b Up)); node91_253((node91-253 Up)); Nagios --- gw_sophia; Nagios --- gw_orsay; Nagios --- gw_nancy; Nagios --- gw_lyon; Nagios --- gw_grenoble; Nagios --- gw_bordeaux; Nagios --- gw_belfort; Nagios --- gw_toul; Nagios --- node06_001; Nagios --- node06_002; Nagios --- node06_003; Nagios --- node06_004; Nagios --- node06_005; Nagios --- node06_006; Nagios --- node06_007; Nagios --- node06_008; Nagios --- node90_002; Nagios --- node90_003; Nagios --- node33_002; Nagios --- node38_001; Nagios --- node69_002; Nagios --- node54_002; Nagios --- node54_003; Nagios --- node54_004; Nagios --- node54_005; Nagios --- node54_006; Nagios --- node54_007; Nagios --- node54_00a; Nagios --- node54_00b; Nagios --- node91_253;
```



Usage modes

- Shared (preparing experiments, size S)
 - No dedicated resources (users log in nodes and use default settings, etc.)
 - Reserved (similar to Planet lab, size M)
 - Reserved resources but uncoordinated (Users may change node's OS on reserved ones)
 - Batch (automatic, size L or XL)
 - Reserved and coordinated resources experiments run under batch/automatic mode)
 - All these modes with calendar scheduling
- + compliance with local usages (almost every cluster receives funds from different institutions and several projects)



Grid'5000 Communauté d'utilisateurs

- Grid'5000 offrira au début 2005:
 - 8 clusters distribués sur 8 sites en France sites,
 - Environ 2500 CPUs,
 - Environ 2,5 TB de mémoire,
 - Environ 100 TB de Disque,
 - Environ 8 Gigabit/s (bidirectionnel) de bande passante
 - Environ 5 Tera opérations / sec
 - La possibilité de reconfigurer le logiciel
[protocoles/OS/Middleware/Runtime/Application]
 - Outils logiciels pour la Virtualisation et l'émulation de l'Hétérogénéité
- Grid'5000 sera ouvert aux chercheur en Grilles début 2005
- Pourrait être ouvert aux participants des ACI « Data Masse »
- Procédure doit encore être déterminée (certainement fondé sur un appel à projets)



Q&A