An e-Infrastructure for Grid Research and the Industry: the CoreGRID experience

European Research Network on Foundations, Software Infrastructures and Applications for large scale distributed, GRID and Peer-to-Peer Technologies

Thierry PRIOL
Scientific Coordinator
INRIA
http://www.coregrid.eu
Thierry.Priol@inria.fr

European Commission

Information Society
Technologies

Agenda

- **A brief overview of CoreGRID**

- **Requirements for an e-Infrastructure from CoreGRID researchers**

- **Reconfigurable e-Infrastructures for Grid research**
  - **Grid'5000**
  - **DAS-3**

- **CoreGRID and the industry: what we learnt**

- **Some recommendations**

European research network on foundations, software Infrastructures and applications for large-scale, distributed GRID and peer-to-peer technologies

European Commission

- 2 -

Information Society
Technologies

CoreGRID objectives

- **To build a European-wide research laboratory**
  - **To avoid fragmentation of Grid research activities in Europe**
  - **Create the European "Grid Lighthouse" and be seen as such worldwide**
  - **To achieve integration and sustainability**
- **To build solid foundations for GRID and P2P technologies**
  - **Both on a <u>methodological</u> basis and a <u>technological</u> basis**
  - **Support medium and long term research activities**
- **Achieve and promote scientific and technological excellence within & beyond the Grid research community**
- **Collaboration with the industry**
- **Gather and disseminate European research**
- **A think-tank for spin-off projects**
  - **EC funded, bilateral projects, international cooperations, …**

European research network on foundations, software Infrastructures and applications for large-scale, distributed GRID and peer-to-peer technologies
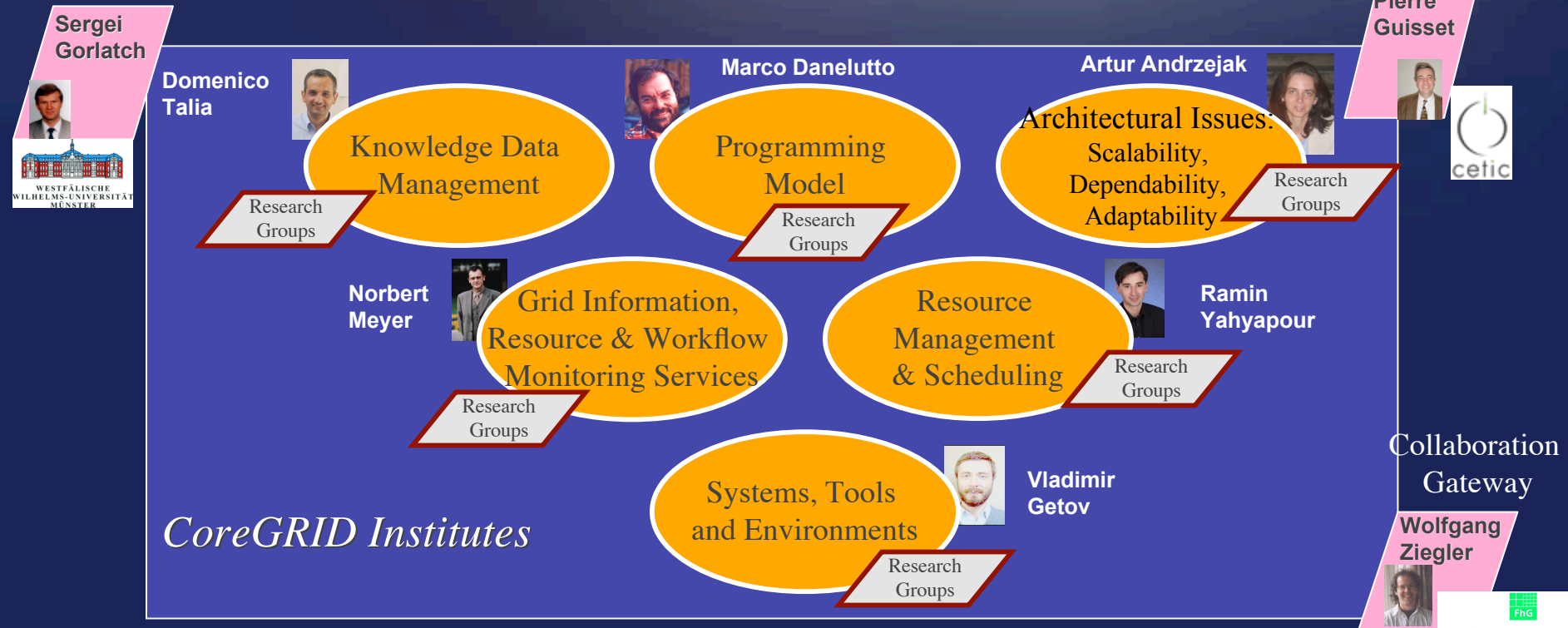
European Commission

- 3 -

Information Society
Technologies

CoreGRID Membership

41 partners from 19 Countries (1 from S. America)

CoreGRID researchers
    Aug 2005: 145 researchers
    Today: 155 researchers

CoreGRID PhD students
    Aug 2005: 169 PhD students
    Today: 168 PhD students

IC
RAL-CCLRC
U. Belfast
U. Manchester
U. Newcastle
U. Wesminster

KTH
SICS

VTT

FhG
FZJ
Delft U.
Vrije U.
USTUTT
U. Muenster
UNI DO
PSC
U. Passau
CYFRONET
CETIC
ZIB
UCL
Masaryk U.
INRIA
U. Innsbruck
EPFL
CNRS
UASF
SZTAKI
ERCIM

UPC
CNR-ISTI
CLPP-BAS
U. Coimbra
INFN
U. Calabria
U. Pisa
ICS-FORTH
U. Chile
U. Cyprus

European Commission

Information Society
Technologies

Dissemination Activities

- **A highly visible initiative within the International Grid research communities**
  - **Sponsorships to GGF and now OGF, CoreGRID label**
  - **CoreGRID Symposium**
- **More than 80 joint Technical Reports, 4 Springer CoreGRID series volumes, a first annual report**
- **Many joint papers published in journals/conferences/workshops**
  - **Publication database (more than 400 references)**

European research network on foundations, software Infrastructures and applications for large-scale, distributed GRID and peer-to-peer technologies

European Commission

- 6 -

Information Society
Technologies

A set of well identified research challenges

Knowledge & Data Management

–   Handling information/data that are required/produced by a wide range of diverse processing power

Programming Model

–   Making the programming of Grid infrastructures as simple and transparent as possible

System Architecture

–   Designing the next generation Grid middleware

Grid Information, Resource and Workflow Monitoring Services

–   Scalable information service to implement a service view of the Grid

Resource Management & Scheduling

–   Scheduling jobs/applications/tasks/computation within a Grid environment

Grid Systems, Tools and Environments

–   Integrating various middleware, tools and applications for problem solving

European research network on foundations, software Infrastructures and applications for large-scale, distributed GRID and peer-to-peer technologies

European Commission

Information Society
Technologies

- 7 -

CoreGRID requirements for a European e-Infrastructures for Grid research

- **Research activities within CoreGRID aims at contributing to the development of Next Generation Grid Middleware**
  - **The Operating System of the Grid Infrastructure !**

- **Experiments have to be carried out to validate the research**
  - **Often at a very large scale (scalability studies)**
  - **In real conditions (resource failures, network congestions, …)**
  - **With different parameters, reproducible, …**

- **Access to e-Infrastructures is one of the integration activities provided by the network (Grid Testbed)**
  - **First plan was to develop our own e-Infrastructure**
  - **Eventually, we rely on existing national or European e-Infrastuctures**
    - **Grid'5000, DAS-2 & 3, CLUSTERIX, VIOLA, EGEE**

European research network on foundations, software Infrastructures and applications for large-scale, distributed GRID and peer-to-peer technologies

European Commission

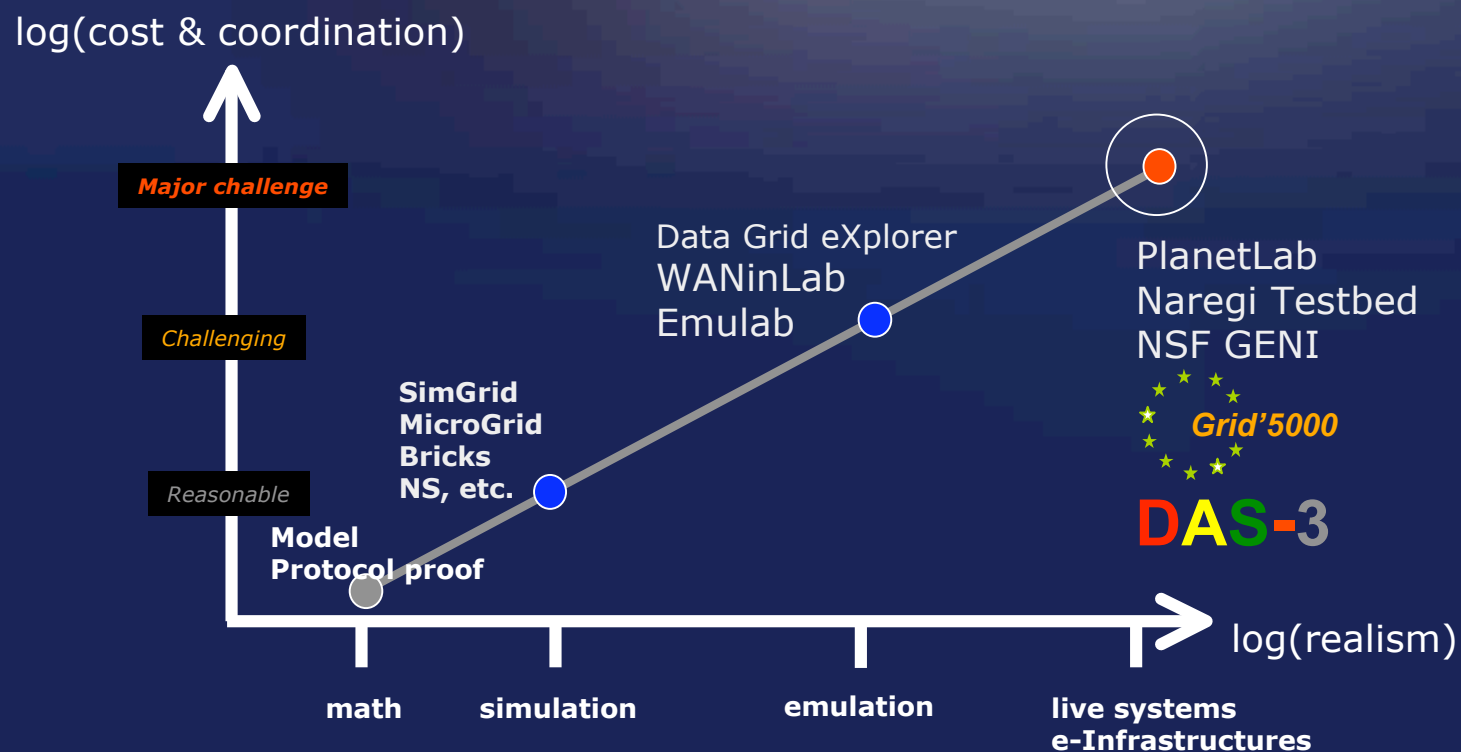- 8 -

Information Society
Technologies

CoreGRID requirements for a European e-Infrastructures for Grid research

- **Various type of Grids**

- **Need to "play" with the network (topology, bandwidth, latency, ...)**
- **Need to be able to experiment various software layers within Grid middleware systems**
  - **We are contributing to the development of Grid operating systems but we are not allowed to install these new OS on existing large-scale e-Infrastructures !**
  - **Incompatible with a production e-Infrastructure**

European research network on foundations, software Infrastructures and applications for large-scale, distributed GRID and peer-to-peer technologies

European Commission

- 9 -

Information Society
Technologies

How we can validate our research ?

Concept of reconfigurable e-Infrastructures for Grid research

- **Grid'5000**
  - **Project from the French ACI GRID**
  - **Reconfiguration of the software stacks**

- **DAS-3**
  - **Project from the Dutch ASCI research school**
  - **Reconfiguration of the optical wide-area network**

European research network on foundations, software Infrastructures and applications for large-scale, distributed GRID and peer-to-peer technologies

European Commission

Information Society Technologies

- 11 -

Grid'5000: an e-Infrastructure for Grid research

Building a nation wide experimental platform for Large scale Grid & P2P experiments
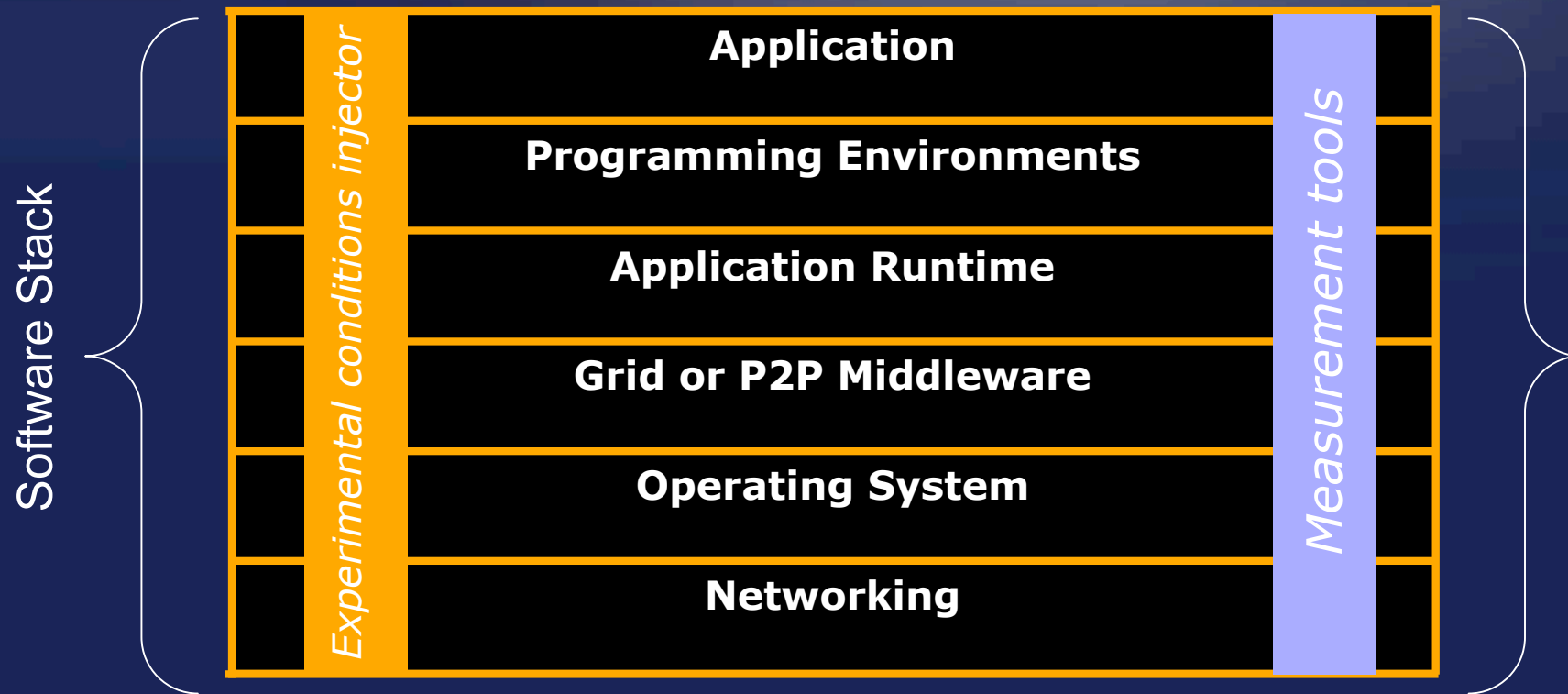
- 9 geographically distributed sites in France
- Every site hosts a cluster (from 256 CPUs to 1K CPUs)
- All sites are connected by RENATER at 10 Gbit/s
  - Grid'5000 is confined: communications between sites are isolated from the Internet and vice versa
- RENATER hosts probes to trace network load conditions
- Design and develop a system/middleware environment for safely test and repeat experiments

Use the platform for Grid experiments in real life conditions

- Port and test applications, develop new algorithms
- Address critical issues of Grid system/middleware:
  - Programming, Scalability, Fault Tolerance, Scheduling
- Address critical issues of Grid Networking
  - High performance transport protocols, Qos
- Investigate original mechanisms
  - P2P resources discovery, Desktop Grids

European research network on foundations, software Infrastructures and applications for large-scale, distributed GRID and peer-to-peer technologies

European Commission

- 12 -

Information Society
Technologies

![CoreGRID]

Grid'5000 principle: A highly reconfigurable experimental platform

Software Stack

Experimental conditions injector

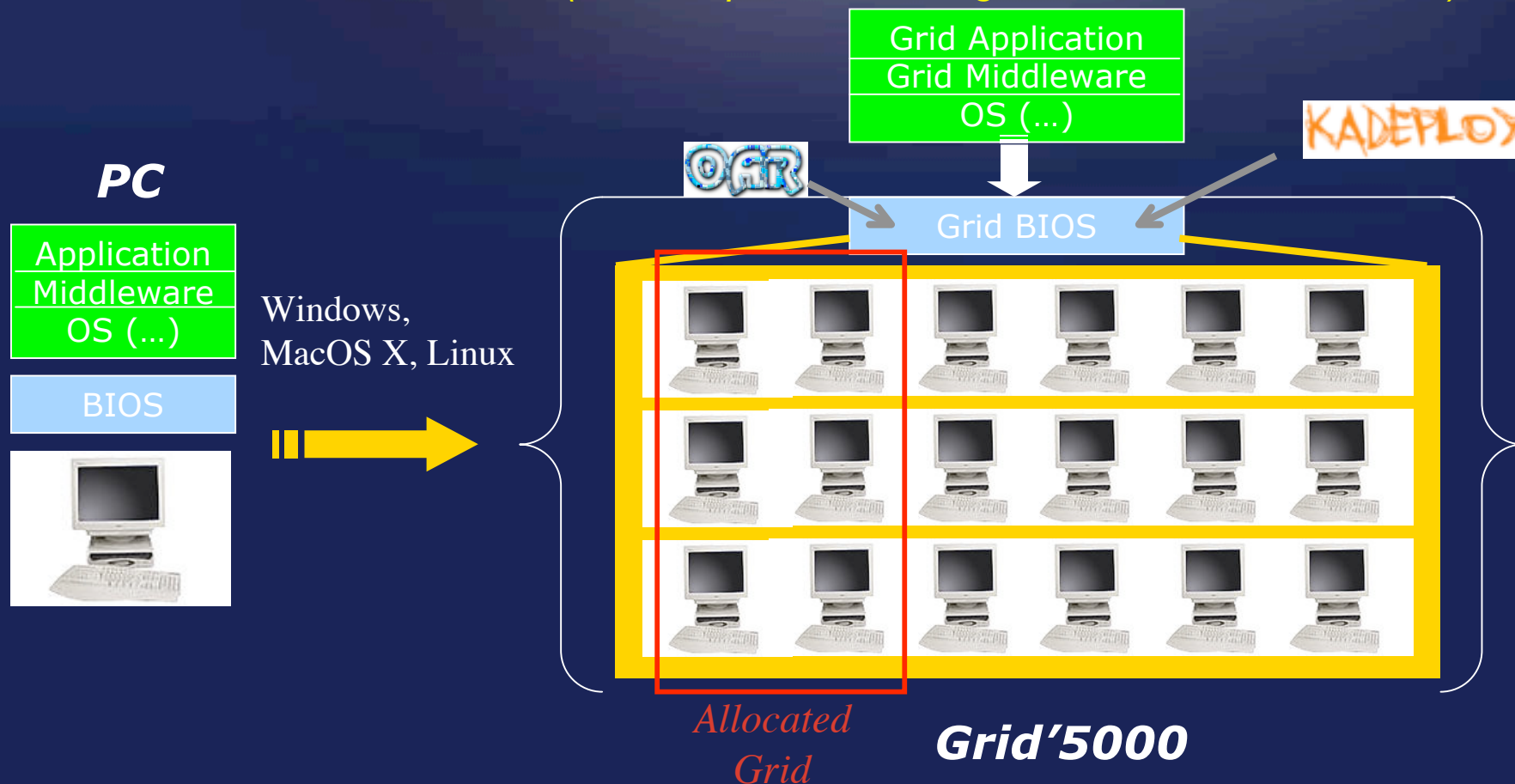| Application |
|---|
| Programming Environments |
| Application Runtime |
| Grid or P2P Middleware |
| Operating System |
| Networking |

Measurement tools

→ Let users create, deploy and run their software stack, including the software to test and their environment + measurement tools + experimental conditions injectors
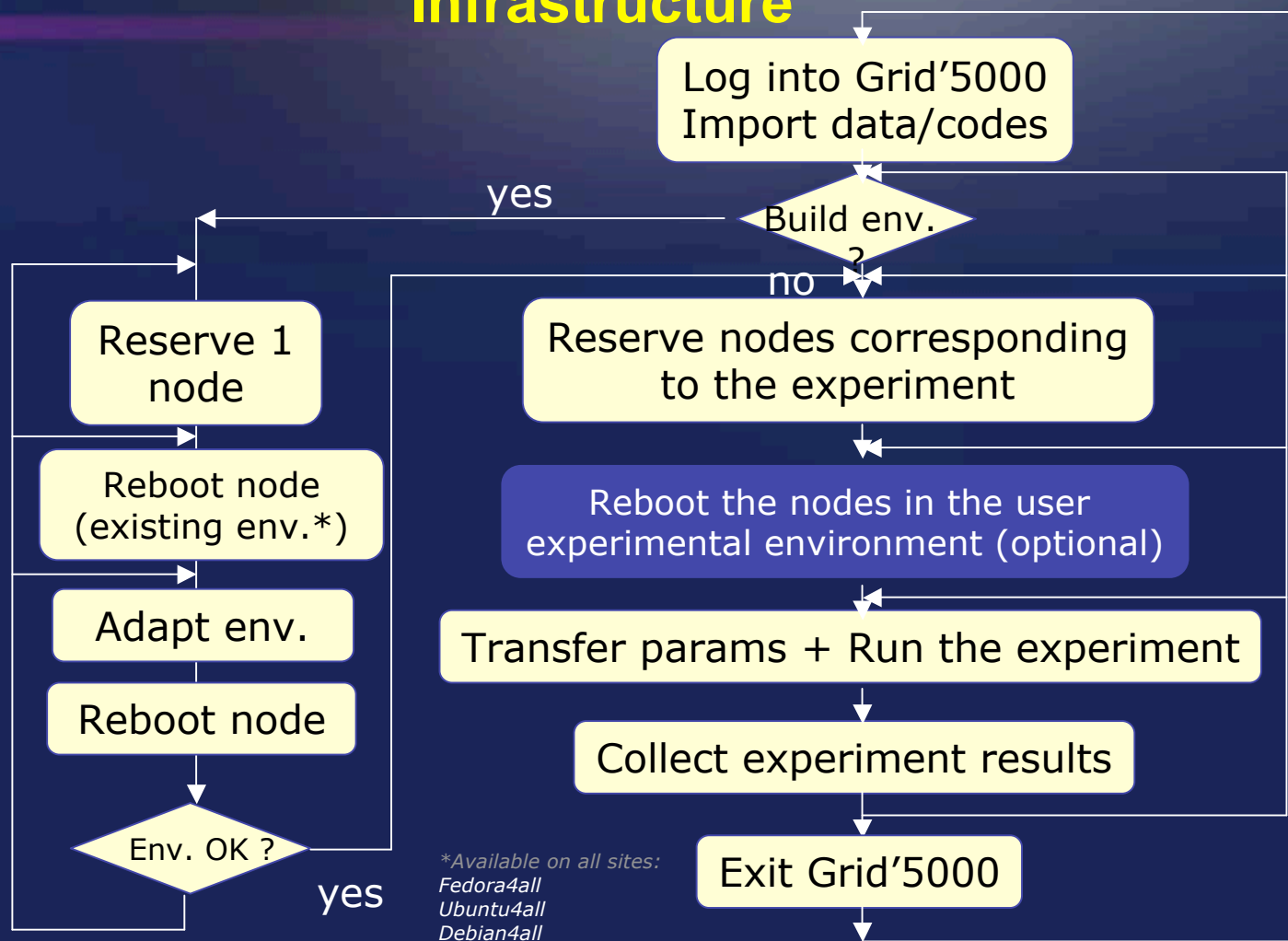
European research network on foundations, software Infrastructures and applications for large-scale, distributed GRID and peer-to-peer technologies

European Commission

- 13 -

Information Society
Technologies

Grid'5000 e-Infrastructure

Software stacks
(User's specific stack, gLite, Unicore, Globus, …)

Grid Application
Grid Middleware
OS (…)

KADEPLOY

OAR

Grid BIOS

PC

Application
Middleware
OS (…)

BIOS

Windows,
MacOS X, Linux

Allocated Grid

Grid'5000

European research network on foundations, software Infrastructures and applications for large-scale, distributed GRID and peer-to-peer technologies

European Commission

- 14 -

Information Society
Technologies

Experiments with the Grid'5000 reconfigurable e-Infrastructure

Log into Grid'5000
Import data/codes

Build env. ?

yes

no

Reserve 1 node

Reserve nodes corresponding to the experiment

Reboot node (existing env.*)

Reboot the nodes in the user experimental environment (optional)

Adapt env.

Transfer params + Run the experiment

Reboot node

Collect experiment results

Env. OK ?

yes

*Available on all sites:
Fedora4all
Ubuntu4all
Debian4all

Exit Grid'5000

European research network on foundations, software Infrastructures and applications for large-scale, distributed GRID and peer-to-peer technologies

European Commission

- 15 -

Information Society
Technologies

DAS-3

VU (85 nodes)

UvA/MultimediaN(46)

UvA/VL-e (40)

SURFnet6

TU Delft (68)

Leiden (32)

10 Gb/s lambdas

272 AMD Opteron nodes
792 cores, 1TB memory
More heterogeneous:
  2.2-2.6 GHz
  Single/dual core nodes
Myrinet-10G (exc. Delft)
Gigabit Ethernet

European research network on foundations, software Infrastructures and applications for large-scale, distributed GRID and peer-to-peer technologies

- 16 -

StarPlane with DAS-3

Key idea:

- Applications can dynamically allocate light paths

- Applications can change the topology of the wide-area network, possibly even at sub-second timescale

VU (Bal, Bos, Maassen)
UvA (de Laat, Grosso, Xu, Velders)

European research network on foundations, software Infrastructures and applications for large-scale, distributed GRID and peer-to-peer technologies

- 17 -

European Commission

Information Society
Technologies
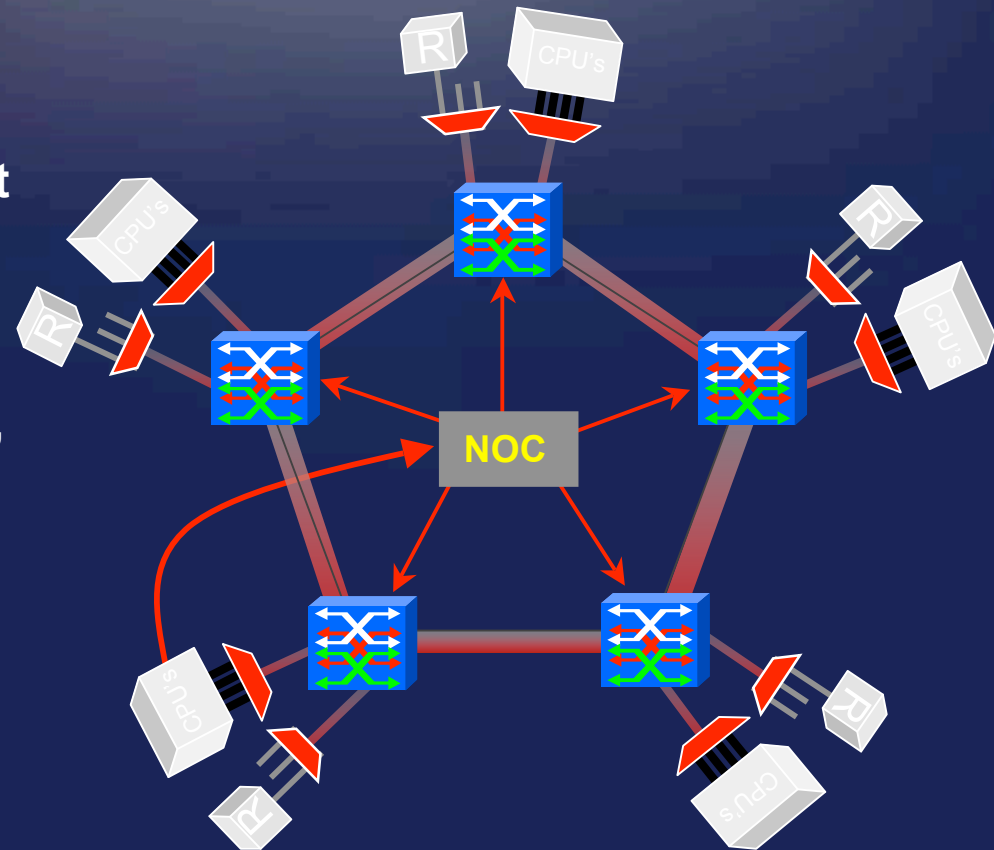
CoreGRID and the Industry

- **Industrial Advisory Board**
  - **To communicate to the Network European Industry's long-term Grid-related technology needs and challenges**
  - **To provide the Network with Grid validation cases**
  - **To help identify the market sector(s) to be addressed within the Grid User Community, in order to ensure a quick adoption of next generation Grid technologies within European industry**
  - **17 IAB Members**
- **Links with NESSI**
  - **NESSI-CoreGRID workshop**
- **Industrial Conference**
  - **Show demo cases to the industry, Grid vision from the industry**
- **Industrial Fellowships**
  - **SLA & Contract Negotiation for the Grid (UPC and ATOS Origin)**
  - **Data Management for High-Performance Distributed Multimedia Content Analysis (VU and Hitachi Europe)**
  - **SLA for supporting use of licensed software in Grid environments (FhG, FzJ and Systems Solutions for Research GmbH)**
  - **Comparative Study over Efficient Data Transport Protocols in Grids (Cetic, Unical, HITACHI sophia antipolis)**

European research network on foundations, software Infrastructures and applications for large-scale, distributed GRID and peer-to-peer technologies

European Commission

Information Society
Technologies

- 18 -

Requirements for a European e-Infrastructure for the Industry based on CoreGRID experience

- **First of all, which industry ?**
  - Telco, software companies, service providers, end-users, …

- **The Grid concept (computing as a utility) seems to be of interest to the industry**
  - One concept but several e-infrastructures to implement the vision
    - Science Grids, Business Grids, …

- **A European initiative to establish a common view for Business Grids: the NESSI-Grid project**
  - To develop a vision of the future for Grids and Service Oriented Infrastructures for business applications
  - Applications involve complex technology stacks <u>often interactive</u> (not batch oriented !)
  - Need to define a <u>general purpose</u> e-Infrastructures for <u>arbitrary</u> business applications

A reconfigurable e-Infrastructure will be flexible enough for this need

European research network on foundations, software Infrastructures and applications for large-scale, distributed GRID and peer-to-peer technologies

- 19 -

Recommendations

#1 Establish a reconfigurable e-Infrastructure at the EU level

– **For Grid research to perform experiments**
– **For the industry to experiment service-oriented utility infrastructures not yet fully defined today**
– **Based on existing national initiatives (G5K, DAS-3, …)**

#2 Encourage strong interactions between reconfigurable and production e-Infrastructures

– **Production e-Infrastructures can be enhanced, refined based on experiments carried out using reconfigurable e-Infrastructures**
– **New identified research challenges from production e-Infrastructures**
– **Complementarity and not competition, we need both to make progress towards next generation Grids**

#3 New requirements will come from FP7 Call1 projects (Software & Service Architectures), need to identify their requirements in term of e-Infrastructures

European research network on foundations, software Infrastructures and applications for large-scale, distributed GRID and peer-to-peer technologies

European Commission

- 20 -

Information Society
Technologies