CENTRO DE SUPERCOMPUTACIÓN DE GALICIA

Providing Grid Services based on Virtualization and Cloud Technologies

A. Simon, C. Fernandez, E. Freire, J. Lopez, S. Diaz
(Grid Systems Technician, Galicia Supercomputing Centre)

VHPC 2009 - Delft, The Netherlands, 25/08/2009

CESGA XUNTA DE GALICIA CONSELLERÍA DE INNOVACIÓN, E INDUSTRIA CSIC CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS MINISTERIO DE CIENCIA E INNOVACIÓN FEDER FONDO EUROPEO DE DESENVOLVEMENTO REXIONAL

ESTABLISHED IN 1993 IN SANTIAGO DE COMPOSTELA (SPAIN)

CESGA

OUTLINE

- INTRODUCTION
- CESGA GRID INFRASTRUCTURE
- BENCHMARKS
- TOWARDS CLOUD COMPUTING
- CONCLUSIONS

INTRODUCTION

CESGA
XUNTA DE GALICIA CONSELLERÍA DE INNOVACIÓN, E INDUSTRIA
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS CSIC
MINISTERIO DE CIENCIA E INNOVACIÓN
FEDER FONDO EUROPEO DE DESENVOLVEMENTO REXIONAL

- **Why virtualizing services?**
  - Grid projects are in continuous development and expanding.
  - Hardware resources are limited.
  - Manpower overhead.

- **How CESGA has solved these issues?**
  - Implementing a totally virtualized grid infrastructure.
    - Middleware is based on gLite.
  - Developing a new accounting procedure.
  - Installing new monitoring tools and resource databases.
    - Cloud computing.

INTRODUCTION

- **Current projects supported at CESGA:**
  - EGEE III
  - int.eu.grid
  - EELA II
  - Spanish NGI (gLite/Globus 4)
  - Ibergrid
  - FORMIGA
  - G-Fluxo
  - e-IMRT
  - …

CENTRO DE SUPERCOMPUTACIÓN DE GALICIA

CESGA GRID INFRASTRUCTURE

CESGA

XUNTA DE GALICIA
CONSELLERÍA DE INNOVACIÓN, E INDUSTRIA

CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS

CSIC

MINISTERIO DE CIENCIA E INNOVACIÓN

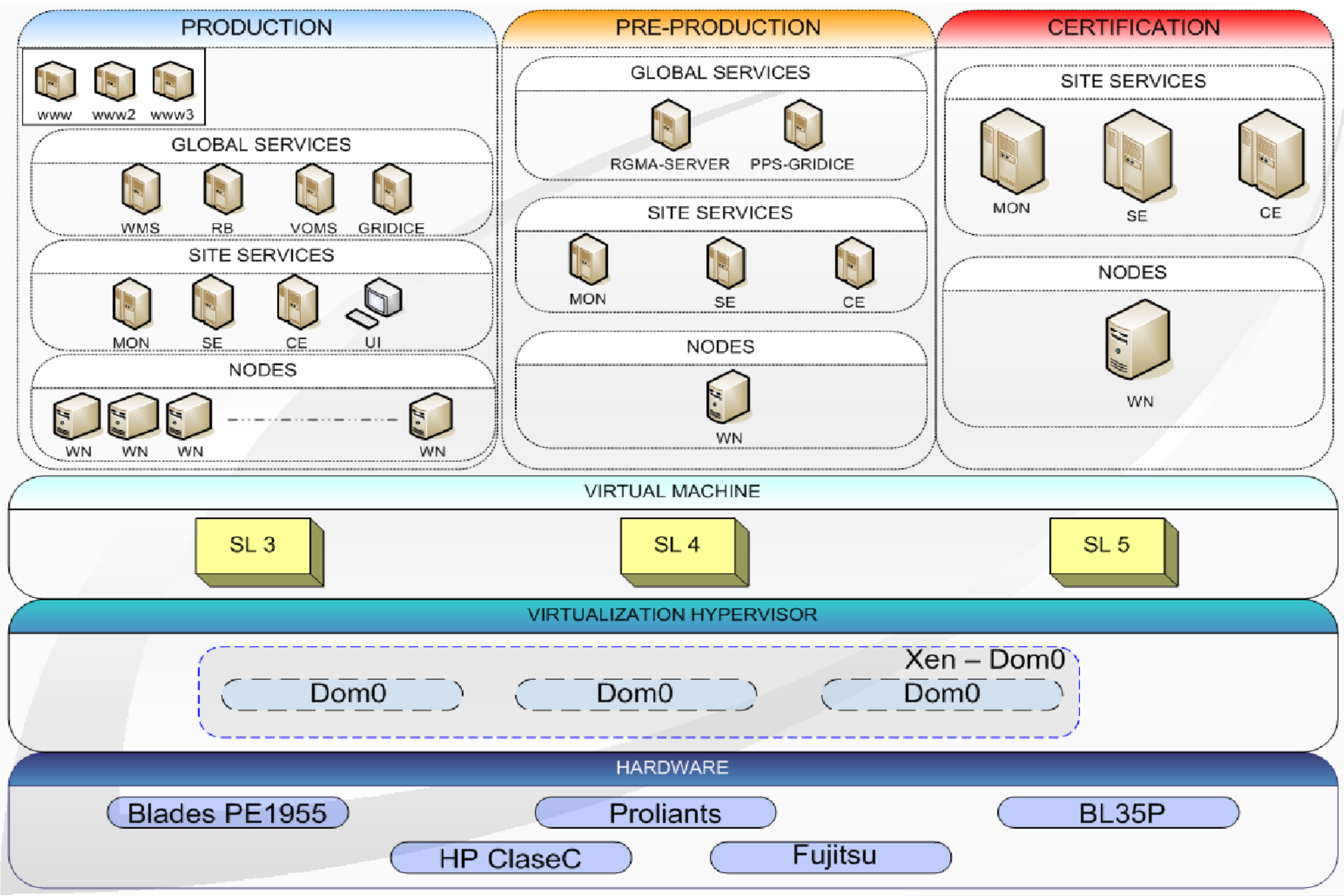
FEDER
FONDO EUROPEO DE DESENVOLVEMENTO REXIONAL

- **Migration was realised two years ago due to limited hardware support in Scientific Linux 3 (SL3)**
- **All grid services are running under a totally virtualized infrastructure**

Servers are automated installed using kickstart and PXE.
All new grid services are configured and installed executing XEN Vms.

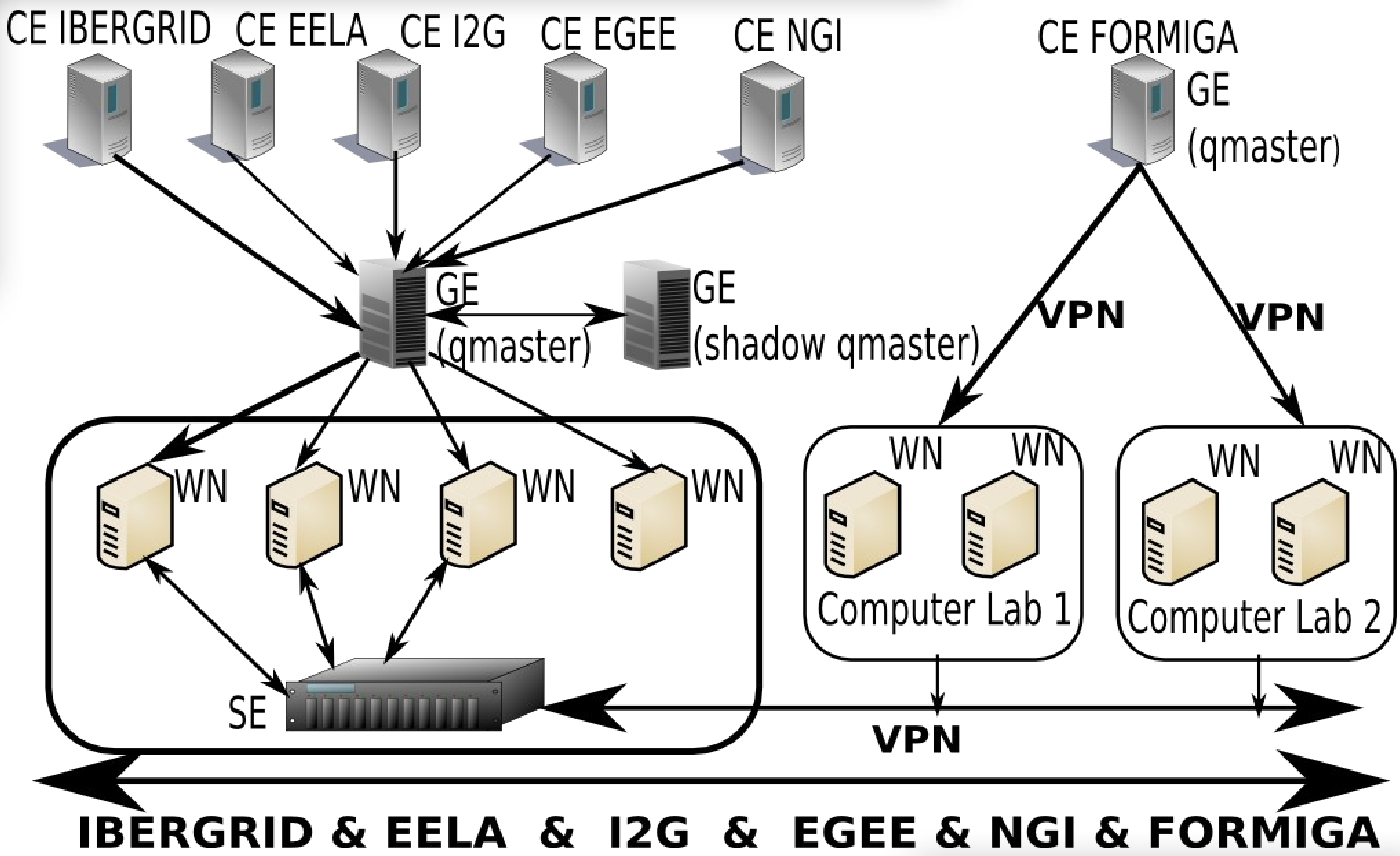
- We can support new projects on demand.
- Better resource utilization.
- Power saving.
- Easy replication.
- Fault-tolerance: Xen combined with Logical Volume Manager (LVM)

CESGA GRID INFRASTRUCTURE



- **Shared batch system**
  - SGE batch system (supported into gLite by LIP, IC and CESGA).
  - Glite middleware requires a Computer Element (CE) for each grid infrastructure.
  - Shared Worker Nodes among the different grid projects supported at CESGA.
  - The batch server is shared using one single SGE qmaster server and a shadow qmaster for fault tolerance purposes.
  - SGE JobManager loads the correct environment for each project (lcgsge.conf)

CESGA GRID INFRASTRUCTURE



- **Integrating computer labs resources**

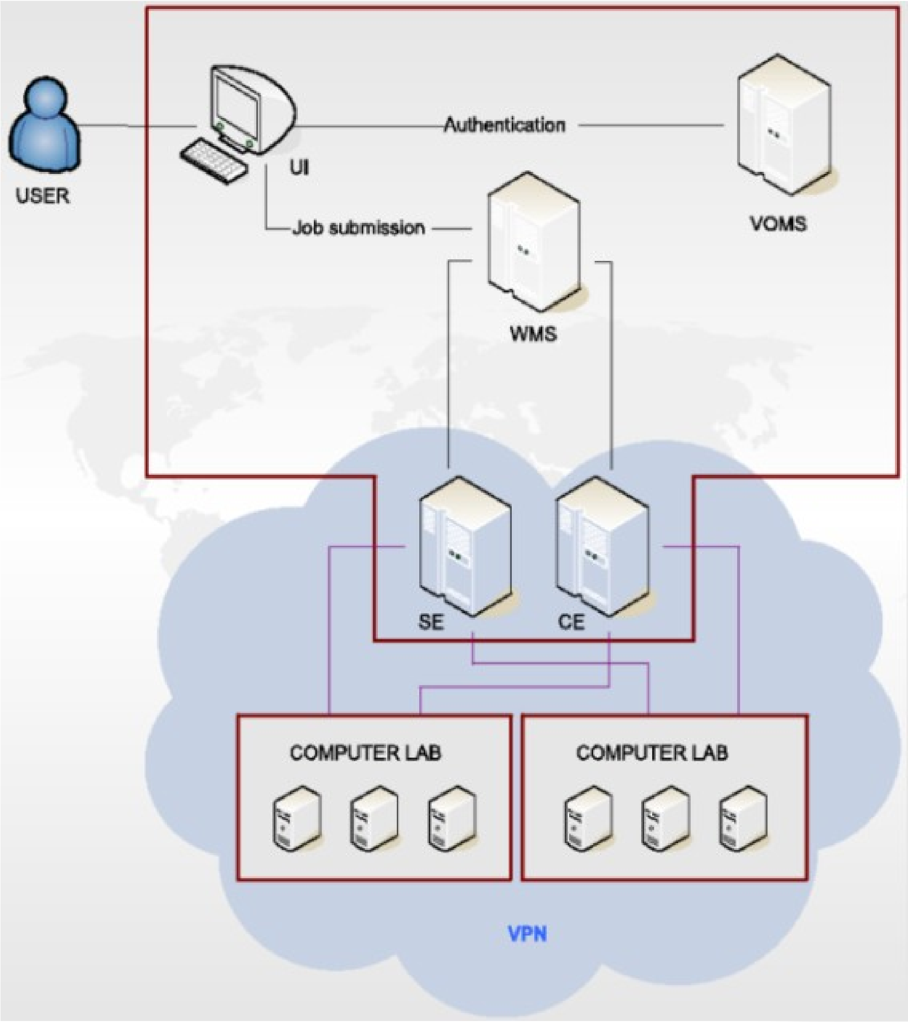
  FORMIGA project has extended the gLite middleware to take advantage of idle resources at computer labs.

  A final prototype is already working in several computer labs of the University of Santiago de Compostela and CESGA.

  Virtual machines are executed under XEN or Vmware depending on the OS installed by the computer lab administrator.

  - These VMs are used as WNs interconnected by a VPN network.
  - Grid users are authenticated using their specific X.509 certificate signed by a CA.

CESGA GRID INFRASTRUCTURE



BENCHMARKS

BENCHMARKS

- **CESGA benchmark configuration:**
  2 x Intel(R) Xeon(R) CPU E5310 @ 1.60GHz QuadCore
  4GB DDR2-667 RAM, 73.4GB HDD
  Gigabit ethernet: Broadcom 5708
  - Dom0: Fedora Core 6 x86_64
  - VM: Scientific Linux 4 i386
  - Xen: Linux Kernel 2.6.18-1.2798.fc6xen

BENCHMARKS

- **Synthetic Benchmarks**

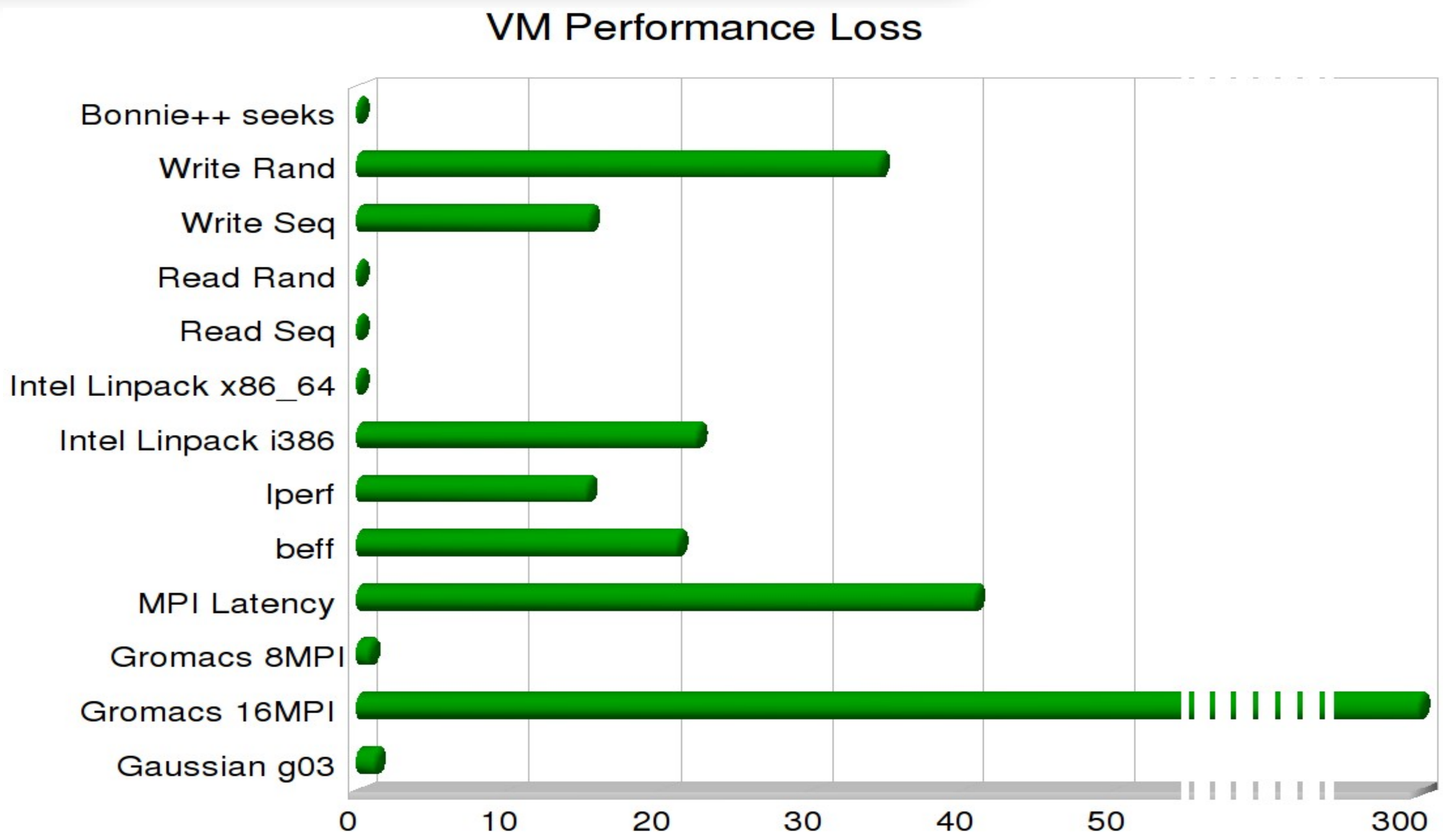
  All the benchmarks were executed using the same parameters both in the virtual and real machines.

  In the sequential write test the VM looses suffers a 15% performance degradation.

  - In the random write test this degradation reaches 30%.
  - CPU performance is quite similar between real and virtual machines.
  - Network bandwidth loss up to 10%.
  - MPI latencies loss up to 40%.

- **Application Benchmarks**
  - Gaussian g03: Computational chemistry simulation package running symmetry example (Test339).
  - Gromacs 3.3.2: Molecular dynamics simulation package running DPPC phospholipid membrane benchmark.
  - Analogous throughput between real and VMs when no intensive I/O operations are needed but...
  - VMs are **294%** slower than real machines running gromacs 16MPI benchmark.

BENCHMARKS


VM Performance Loss
Bonnie++ seeks
Write Rand
Write Seq
Read Rand
Read Seq
Intel Linpack x86_64
Intel Linpack i386
Iperf
beff
MPI Latency
Gromacs 8MPI
Gromacs 16MPI
Gaussian g03
0
10
20
30
40
50
300

CENTRO DE SUPERCOMPUTACIÓN DE GALICIA

TOWARDS CLOUD COMPUTING

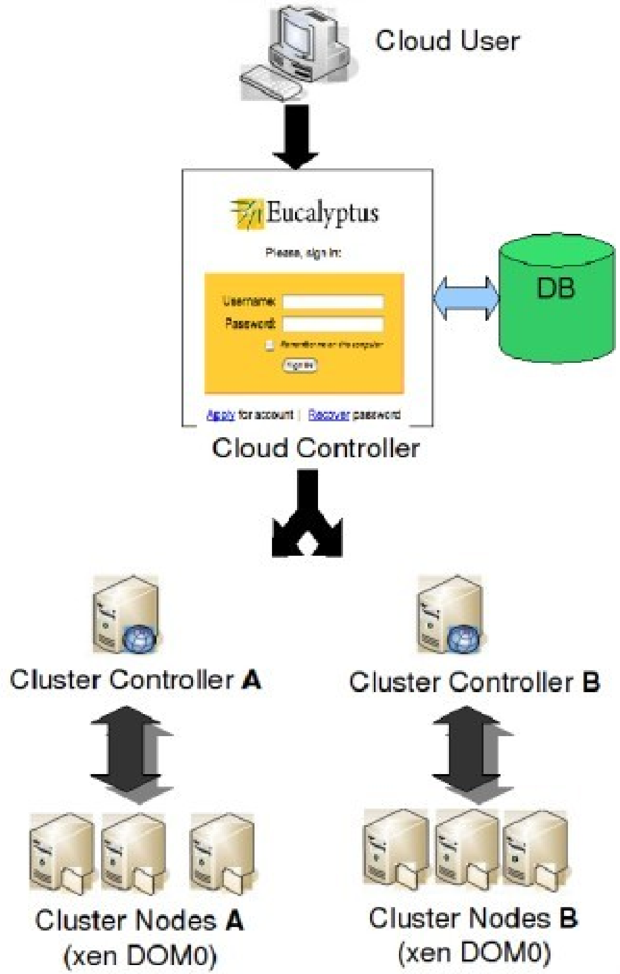
CESGA
XUNTA DE GALICIA
CONSELLERÍA DE INNOVACIÓN, E INDUSTRIA
CSIC CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS
MINISTERIO DE CIENCIA E INNOVACIÓN
FEDER
FONDO EUROPEO DE DESENVOLVEMENTO REXIONAL

- **Why Cloud?**

  Most of the CESGA grid infrastructure is based on VMs.
  VMs "golden copies" are stored in a central repository.
  - New grid services could be started in a short time.
  - XEN resources (Number of CPUs, allocated memory) could be managed from a central cloud controller.

- **Eucalyptus**
  Compatible with Amazon's EC2 interface.
  Only uses open source solutions:
    - Java Developer Kit (SDK) version 1.6 or above.
    - Apache ant 1.6.5 or above.
    - Curl development package.
    - OpenSSL development package.
    - XEN (version >= 3.0.x).
  - Takes care of available resources.
  - Uses different scheduling algorithms:
    - **Greedy** (first node that is found that can run the VM will be chosen).
    - **Roundrobin** (nodes are selected one after another until one is found that can run the VM).

TOWARDS CLOUD COMPUTING



CONCLUSIONS

CONCLUSIONS

- **Totally virtualized infrastructure**

Reduce administration overhead.
- Reduce hardware costs.
- New services are installed in a short time.

CPU throughput is quite similar between VMs and Real machines.
- i386 binaries are executed 22% slower on x86_64 Dom0s.
- Very high network latency decreases MPI throughput, not recommended for intensive MPI jobs.

We can check available resources at a glance.
- Automated start/stop services on demand.
- We can make better use of available Dom0's.

CESGA
XUNTA DE GALICIA CONSELLERÍA DE INNOVACIÓN, E INDUSTRIA
CSIC CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS
MINISTERIO DE CIENCIA E INNOVACIÓN
FEDER FONDO EUROPEO DE DESENVOLVEMENTO REXIONAL

THANK YOU

FOR YOUR ATTENTION

ALVARO SIMON

asimon@cesga.es

www.cesga.es

CESGA | XUNTA DE GALICIA CONSELLERÍA DE INNOVACIÓN, E INDUSTRIA | CSIC CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS | MINISTERIO DE CIENCIA E INNOVACIÓN | FEDER FONDO EUROPEO DE DESENVOLVEMENTO REXIONAL