Management of a Grid Infrastructure in gLite with Virtualization

Miguel Cardenas Montes

Javier Perez-Griffo Callejon

Raul Ramos Pollan

Manuel Rubio del Solar

Ibergrid 2007

Santiago Compostela / 16 of May 2007

MINISTERIO
DE EDUCACIÓN
Y CIENCIA

CENTRO EXTREMEÑO DE
TECNOLOGÍAS AVANZADAS

C E T A Ciemat

Index

- Complexity of a gLite Site

- Classification of gLite Services

- Possibilities for virtualization in gLite

- Virtualization performance in XEN

- XEN with gLite

- CETA-XEN

- Conclusion

MINISTERIO
DE EDUCACIÓN
Y CIENCIA

CENTRO EXTREMEÑO DE
TECNOLOGÍAS AVANZADAS
CETA Ciemat

Complexity of a gLite Site

- Complexity of a site is proportional to the number and nature of the services that compose it.

- Difficult to rapidly deploy new gLite sites.

- Many of the tasks of keeping up-to date a site are repetitive and can be automatized.

  Virtualization can simplify gLite sites management

MINISTERIO DE EDUCACIÓN Y CIENCIA

CENTRO EXTREMEÑO DE TECNOLOGÍAS AVANZADAS **CETA** *Ciemat*

Classification of gLite Services

- Cluster Services **(WN)**
  - ‣ Its complexity relies in keeping up-to date the software installed
  - ‣ Constitute the majority of the site
  - ‣ Configuration is not tied to FQDN

- Basic Services **(CE,SE,Site BDII)**
  - ‣ Strong dependency to core services
  - ‣ Installation can be automated with yaim

- Core Servicies **(RB, LFC, BDII, ...)**

MINISTERIO
DE EDUCACIÓN
Y CIENCIA

CENTRO EXTREMEÑO DE
TECNOLOGÍAS AVANZADAS
CETA Ciemat

Possibilities for virtualization in gLite

- ## Cluster Services **(WN)**
  - ‣ Usage of configure **once and clone many**

- ## Basic Services **(CE,SE,Site BDII)**
  - ‣ Clonage of image with the service pakages
  - ‣ Configuration done automatically when image is **instantiated**

- ## Core Servicies **(RB, LFC, BDII, ...)**
  - ‣ **Snapshots**

MINISTERIO DE EDUCACIÓN Y CIENCIA

CENTRO EXTREMEÑO DE TECNOLOGÍAS AVANZADAS
C E T A Ciemat

XEN Performance

- ## Base System
  - › AMD Opteron Model 270 at 2.0 GHz (Dual-Core) with 4GB Ram **(NO VT)**
  - › Dom 0: Linux Dapper with Xen 3.14 (32 Bits)

- ## Benchmarks Used
  - › FLOPS
  - › IOZONE
  - › BYTEMarks (Numeric Sort, Fourier, IDEA, etc..)

- ## Scenarios
  - › Basic Scenario: Scientific Linux 3 with 2 Virtual CPU's
  - › Lvm Scenario: Scientific Linux 3 with LVM driver
  - › Pin Scenario: Scientific Linux 3 with 2 CPU's

Xen Performance - Basic Scenario

Base System
CPU: 2 Dual Core
Arch: AMD
Memory: 4 G
SO: Scientific Linux 3
Kernel: 2.6

Virtual System
CPU: 4
Memory: 4 G
SO: Scientific Linux 3
Kernel: 2.6
VCPU: virtual
Disk: image

Ubuntu Dapper   **dom0**

hardware

8,2% Cpu Loss
38,6% IO Loss
21,1% BYTEMark Loss

Xen Performance - Basic Scenario

Base System
CPU: 2 Dual Core
Arch: AMD
Memory: 4 G
SO: Scientific Linux 3
Kernel: 2.6

Virtual System
CPU: 2
Memory: 2 G
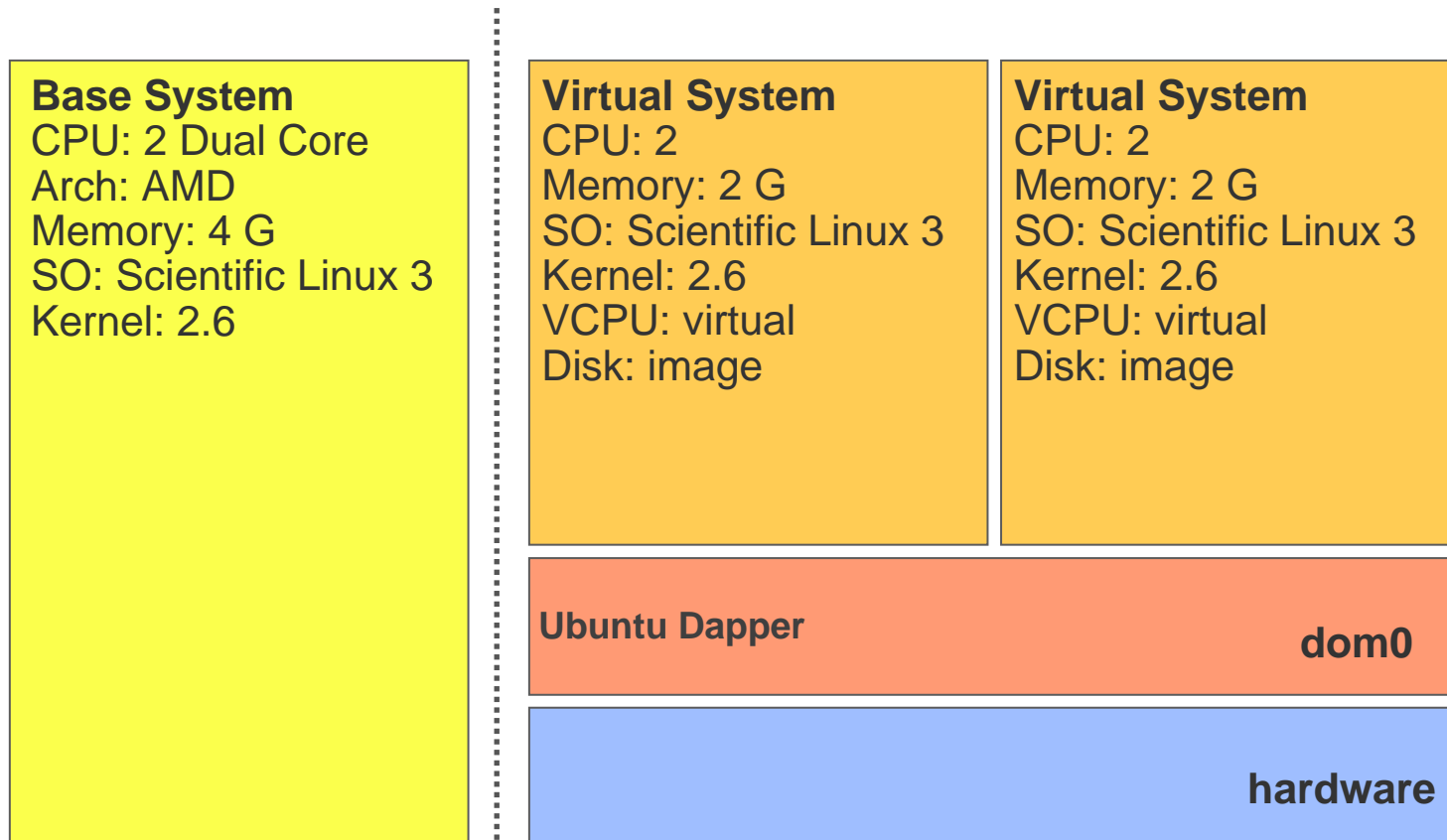SO: Scientific Linux 3
Kernel: 2.6
VCPU: virtual
Disk: image

Virtual System
CPU: 2
Memory: 2 G
SO: Scientific Linux 3
Kernel: 2.6
VCPU: virtual
Disk: image

Ubuntu Dapper   **dom0**

hardware

18,3% Cpu Loss in **each** virtual system
49,3% IO Loss in **each** virtual system
22,5% BYTEMark Loss in **each** virtual system

Xen Performance - LVM Scenario

Base System
CPU: 2 Dual Core
Arch: AMD
Memory: 4 G
SO: Scientific Linux 3
Kernel: 2.6

Virtual System
CPU: 2
Memory: 2 G
SO: Scientific Linux 3
Kernel: 2.6
VCPU: virtual
Disk: **LVM**

Virtual System
CPU: 2
Memory: 2 G
SO: Scientific Linux 3
Kernel: 2.6
VCPU: virtual
Disk: **LVM**

dom0

hardware
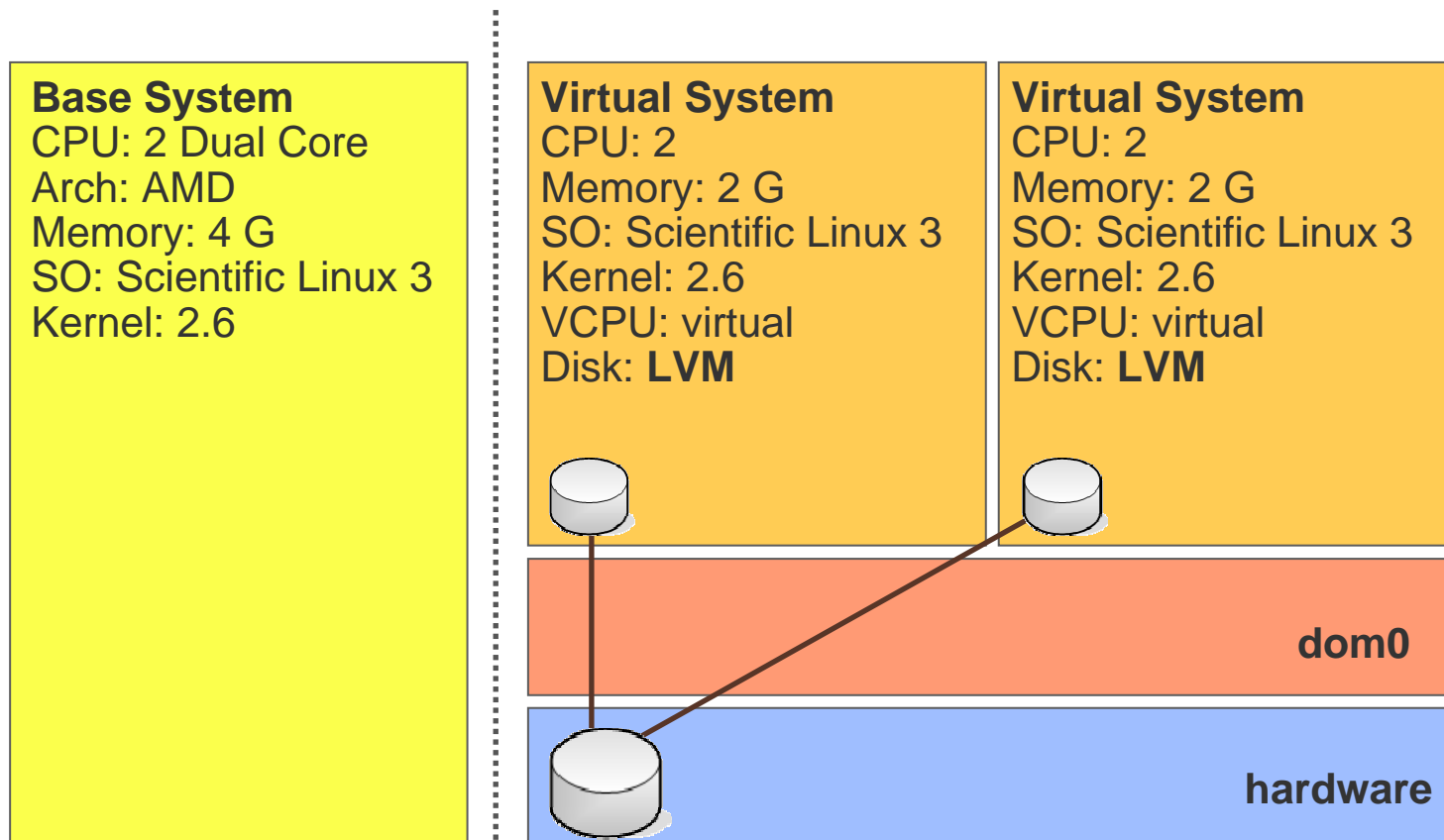
18,6% Cpu Loss in **each** virtual system
46,4% IO Loss in **each** virtual system
22,5% BYTEMark Loss in **each** virtual system

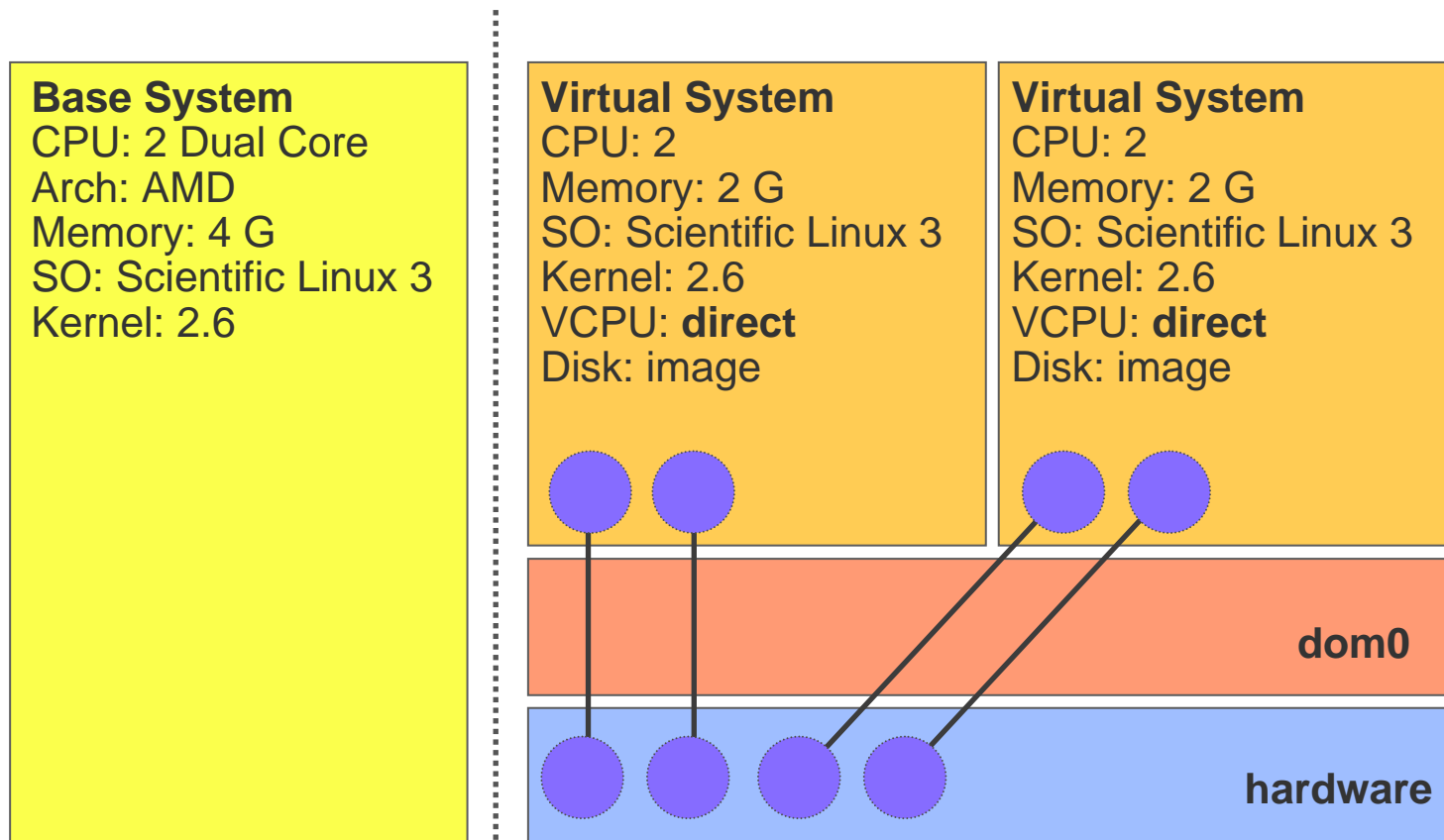# Xen Performance - PIN Scenario

Base System
CPU: 2 Dual Core
Arch: AMD
Memory: 4 G
SO: Scientific Linux 3
Kernel: 2.6

Virtual System
CPU: 2
Memory: 2 G
SO: Scientific Linux 3
Kernel: 2.6
VCPU: **direct**
Disk: image

Virtual System
CPU: 2
Memory: 2 G
SO: Scientific Linux 3
Kernel: 2.6
VCPU: **direct**
Disk: image

dom0

hardware

17,6% Cpu Loss in **each** virtual system
43,8% IO Loss in **each** virtual system
22,1% BYTEMark Loss in **each** virtual system

XEN Performance - Observation

- Strong Penalization for reading

- Direct drivers into file system only give 3% gain

- Direct cpu give a 5% gain in IO operations

- CPU performance can be outweighed against manageability.

- IO performance of LVM is not significant an can be discarded for simplicity.

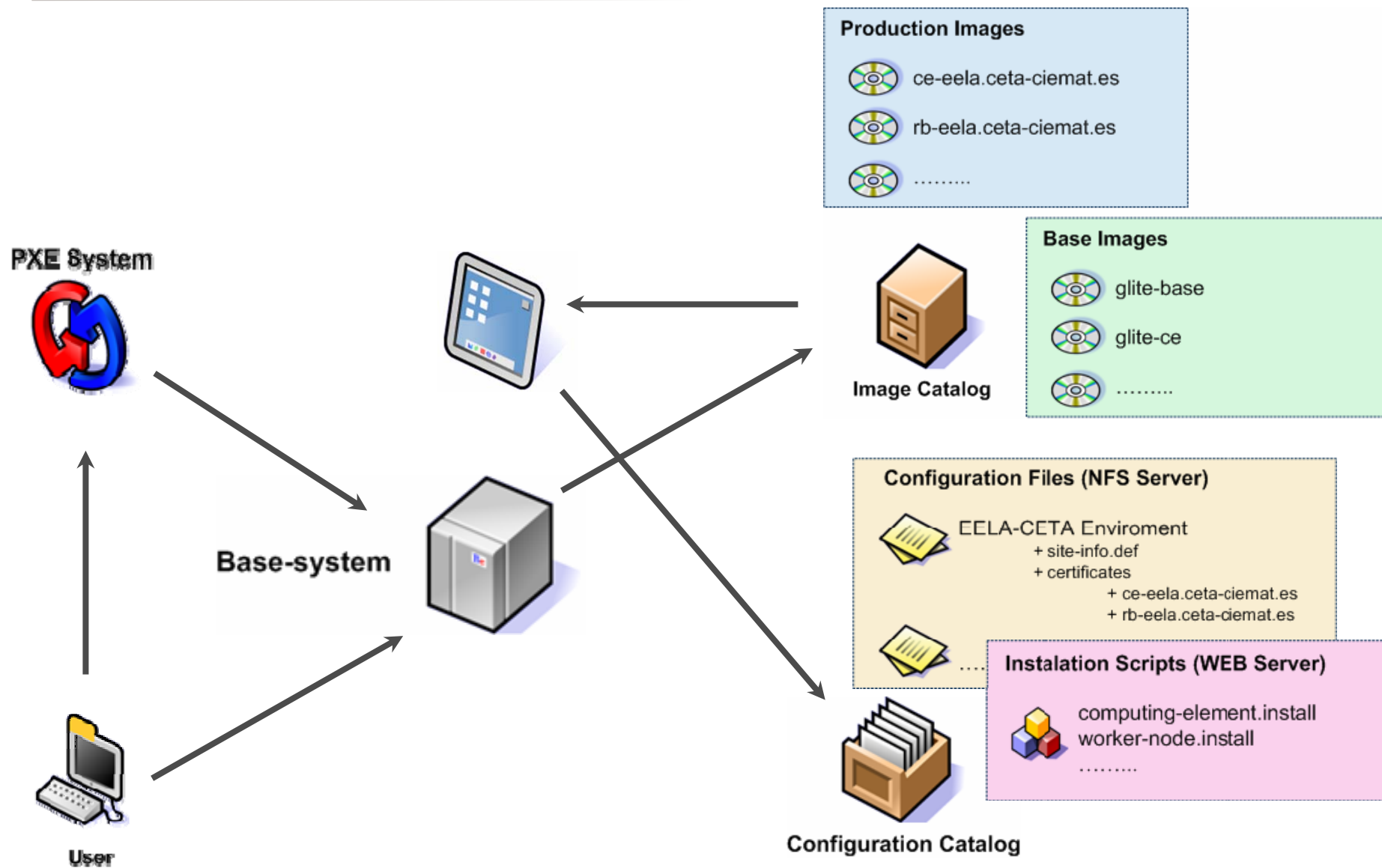- CPU pinning is recommend as improves globally the virtual hosts.

Xen for gLite

- Virtualizable Components:
  - All but....

- We recommend not to virtualize:
  - LFC
  - SE

- Worker Nodes depend on the profile of the application.

CETA-XEN

Enviroment created for the **resource provisioning** of virtual images within a homogenized enviroment. For the creation of **isolated** environments

Features:
- Usage of FQDN as identifier of the virtual host.
    - FQDN to MAC address
- Rapid Cloning deployment
- Easy Snapshoting of virtual images.
- Easy deployment of enviroment
- Easy Architecture

CETA-XEN

Production Images
- ce-eela.ceta-ciemat.es
- rb-eela.ceta-ciemat.es
- ....

Base Images
- glite-base
- glite-ce
- ....

PXE System

Base-system

Image Catalog

Configuration Files (NFS Server)

EELA-CETA Enviroment
+ site-info.def
+ certificates
+ ce-eela.ceta-ciemat.es
+ rb-eela.ceta-ciemat.es

Instalation Scripts (WEB Server)

computing-element.install
worker-node.install
....

User

Configuration Catalog

MINISTERIO
DE EDUCACIÓN
Y CIENCIA

CENTRO EXTREMEÑO DE
TECNOLOGÍAS AVANZADAS
CETA Ciemat

Concluding Remarks

- Virtualization can simplify many of the tasks of maintenance of gLite site.

  ‣ Snapshots of Services

  ‣ Cloning of Environments

  ‣ Rapid Deployment of new Environments

- Virtualization in XEN is not recommend for IO intensive services.

CETA

Thank You for your attention