

# Grids & e-Science 2009

## June 18th, Santander, Spain

### Clouds, Grids & Virtual Machines

Ruben Santiago Montero

[dsa-research.org](http://dsa-research.org)

Distributed Systems Architecture Research Group  
Universidad Complutense de Madrid



- Brief Overview of Clouds
- The IaaS approach for Infrastructure Provisioning
- Management of Virtual Infrastructures
- IaaS for the dynamic provisioning of virtual clusters in Grids
- Grids & Clouds: Trends and Opportunities
- OpenNebula Tutorial

# Cloud Computing in a Nutshell

## What

## Who

### Software as a Service

On-demand access to any application

### End-user

(does not care about hw or sw)



facebook

### Platform as a Service

Platform for building and delivering web applications

### Developer

(no managing of the underlying hw & sw layers)



Windows Azure

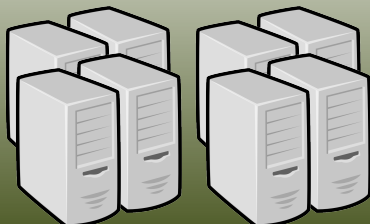
force.com  
platform as a service

### Infrastructure as a Service

Delivery of a *raw* computer infrastructure

### System Administrator

(complete management of the computer infrastructure)



Physical Infrastructure



- **Simple Interface**
- **Raw *Infrastructure* Resources**
  - Total control of the resources
  - Capacity leased in the form of Vms
  - Complete Service-HW decoupling
- **Pay-as-you-go**
  - A single user can not get all the resources
- **Elastic & “*infinite*” Capacity**

**NOTE: This applies to any IaaS Cloud (private, public...)**

# Service Deployment using IaaS



**Service  
End-Users**

## Total control of service layout

- Software Stack
- Type & Number of components
- Service Elasticity
- Placement Constraints

**Web Server  
(Load Balancer)**

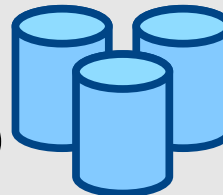
**App  
Server**

**App  
Server**

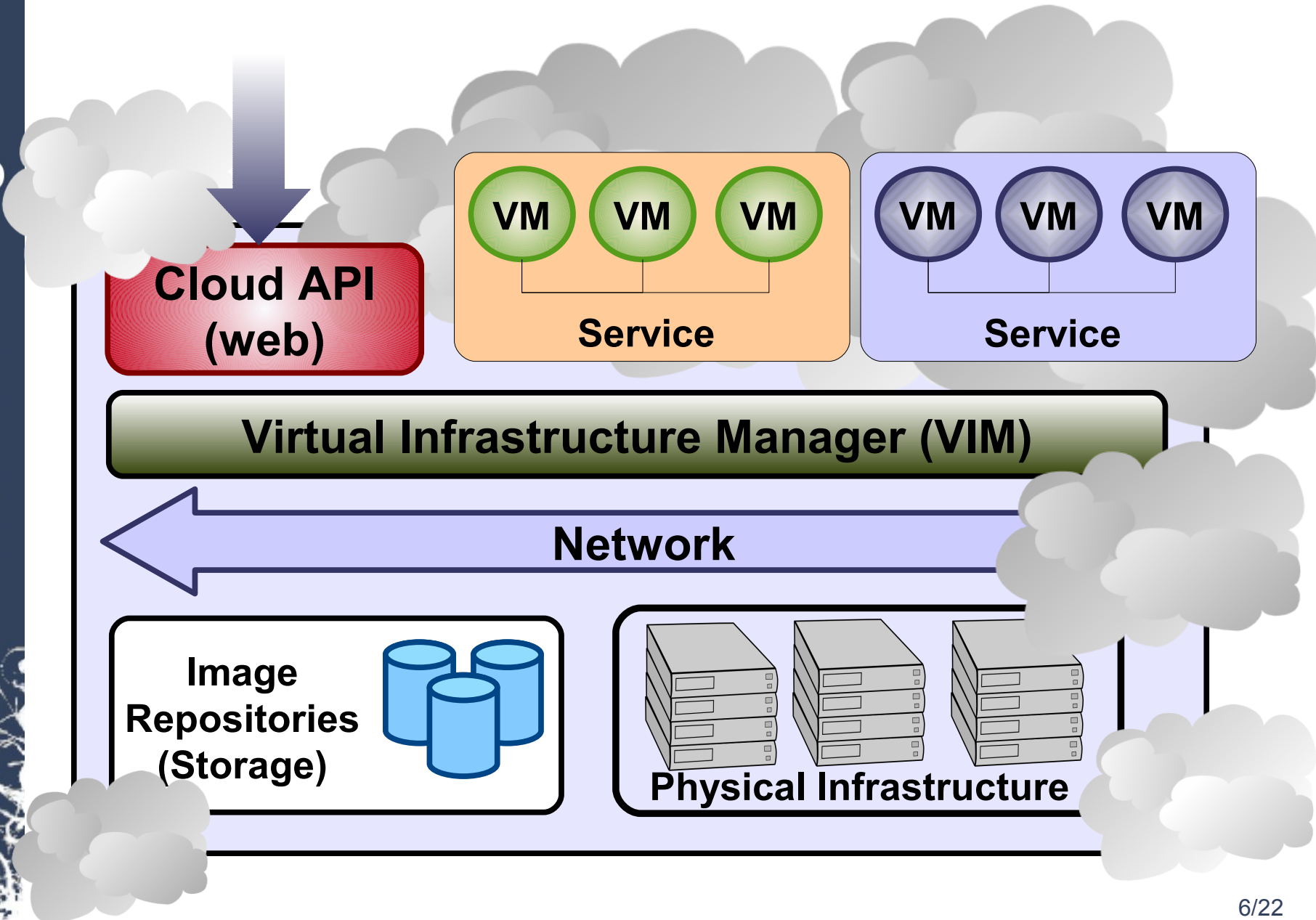
**App  
Server**

**Network**

**DBs  
(storage)**

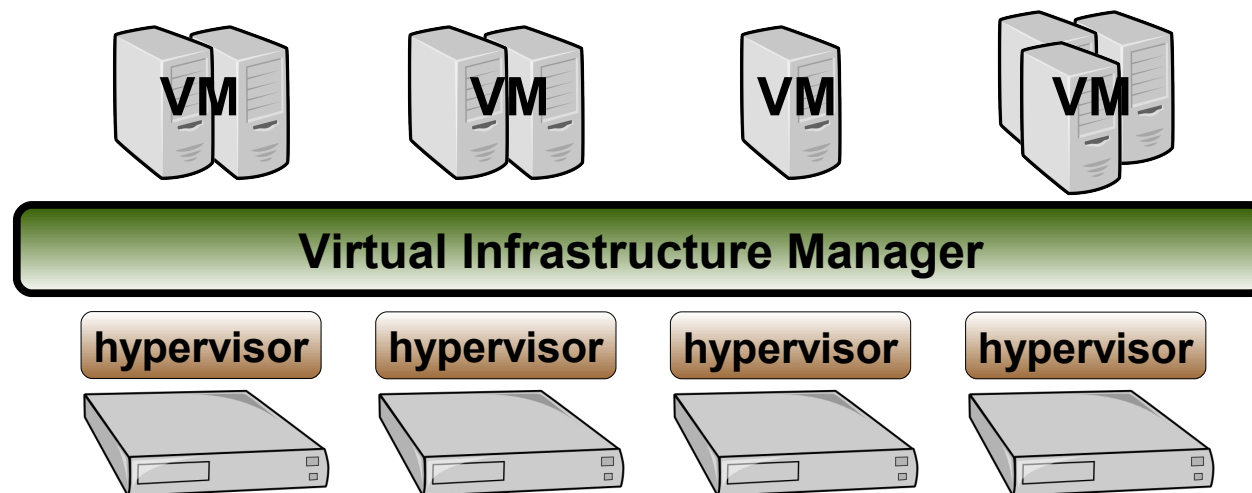


# The Anatomy of an IaaS Cloud



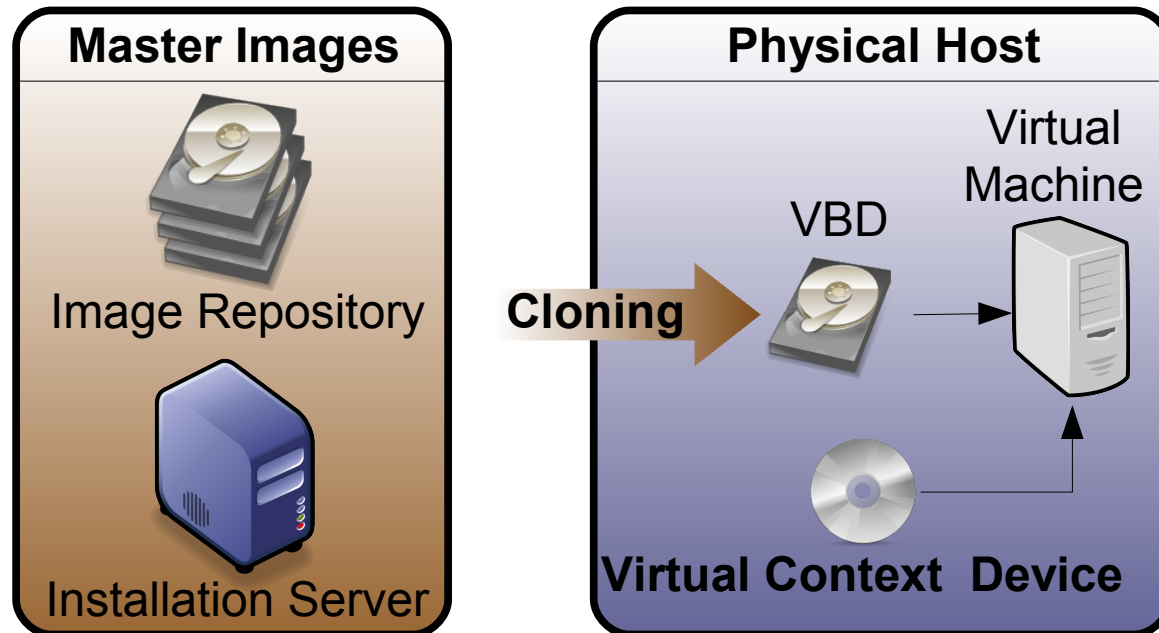
# Virtual Infrastructure Manager

- VMs are great!!...but something more is needed
  - Where did/do I put my VM? (**scheduling & monitoring**)
  - How do I provision a new cluster node? (**clone & context**)
  - What MAC addresses are available? (**networking**)
- Provides a **uniform view** of the resource pool
- **Life-cycle management** and monitoring of VM
- The VIM **integrates** Image, Network and Virtualization



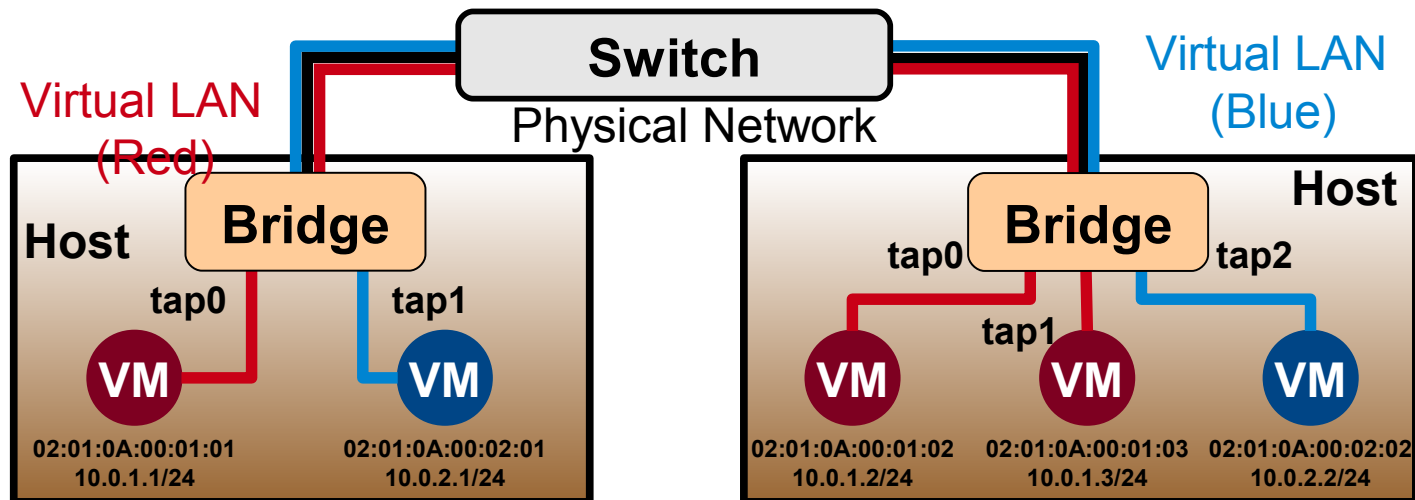


- VM Images Sources:
  - Master images in local repositories
  - Appliance supplier
  - Creation on the fly
- Clones have to be contextualized (Context VBD)





- VMs interconnected through **one or more** networks
  - Isolated, layer 2 LANs
  - Virtual networks are dynamically created
  - Medium size networks (x.x.x.x/20) with limited public IPs
- **TCP/IP services** are not responsibility of the VM Manager



- **Resource Selection:** *Where do I place the VM?*
  - Capacity planning (consolidation)
  - Placement requirements (e.g. affinity)
  - Placement Heuristics (e.g. Green IT, AR...)
- **Resource Preparation:** *What do I need for the VM?*
  - Network preparation
  - Image cloning & contextualization
- **VM Creation:** *How do I start a VM?*
  - Interface with different hypervisors
- **VM Monitoring:** *How is the VM doing?*
- **VM Migration:** *Is there a better resource for the VM?*
  - Adjust placement to better fit to the infrastructure target
- **VM Termination:** *Do I need to save any VM image?*

[www.OpenNebula.org](http://www.OpenNebula.org)

CLI / libvirt

Cloud Interface

OpenNebula Core

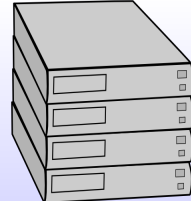
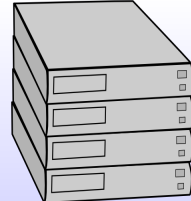
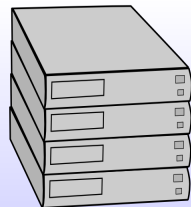
Scheduler

Virt. Plugins

XEN/KVM

CLOUD

OpenNebula



Local Infrastructure

- Flexible & Open Design
  - Third-party components
  - Easily adapted & extended
- Management of *Virtual Services*
  - Image, Network & Context
- Integrated with cloud providers
- Open Source – Apache2
- Included in Ubuntu 9.04 (server)



## Some Limitations of Current Grids

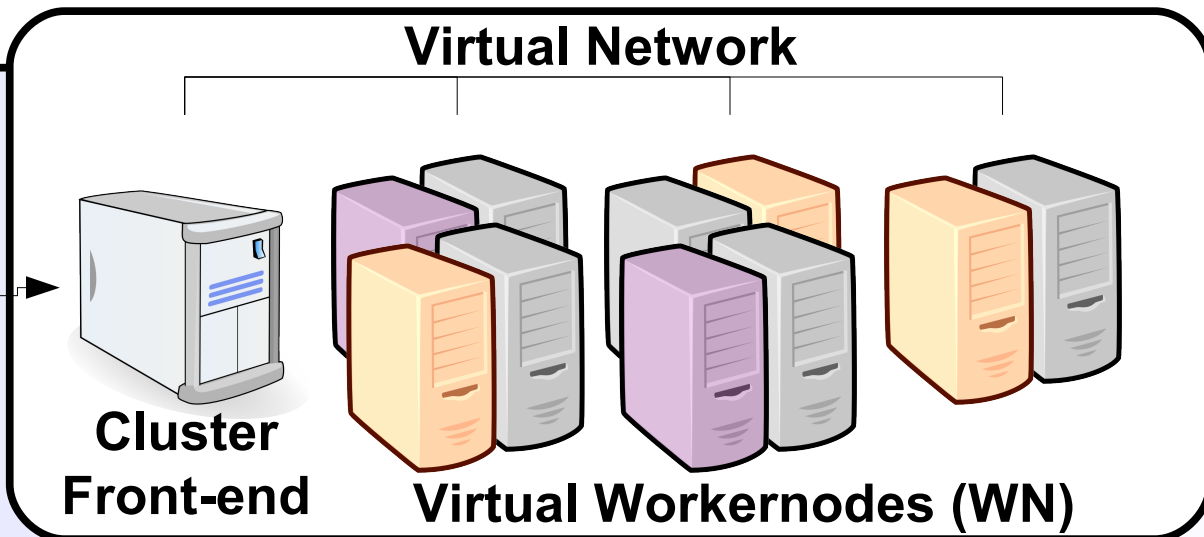
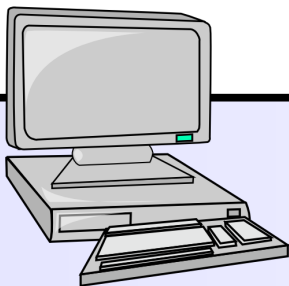
- High degree of heterogeneity (software & hardware)
- High operational costs
- Isolate and partition resources contributed to the Grid
- Specific environment requirements for different Vos
- Users simply do not feel like adopting our execution models (*pilot jobs...*)



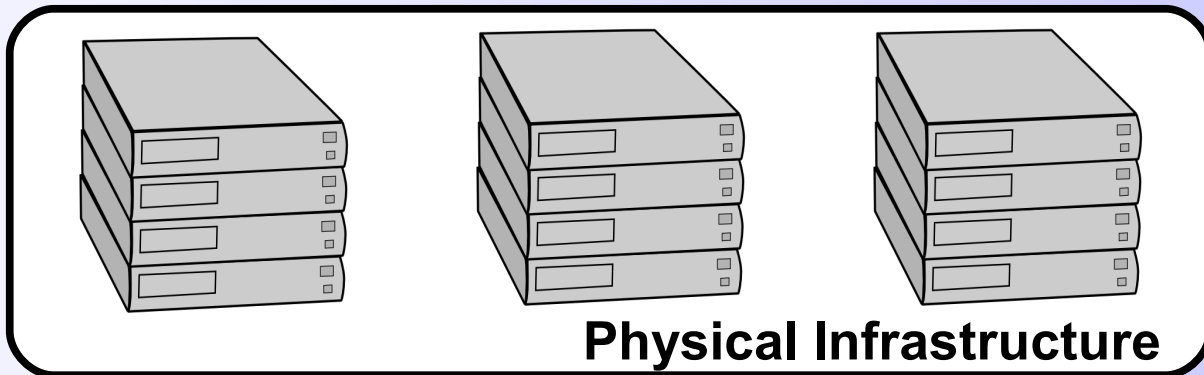
Grids are difficult to maintain, operate and use

# Grids & Virtual Machines

## Cluster users



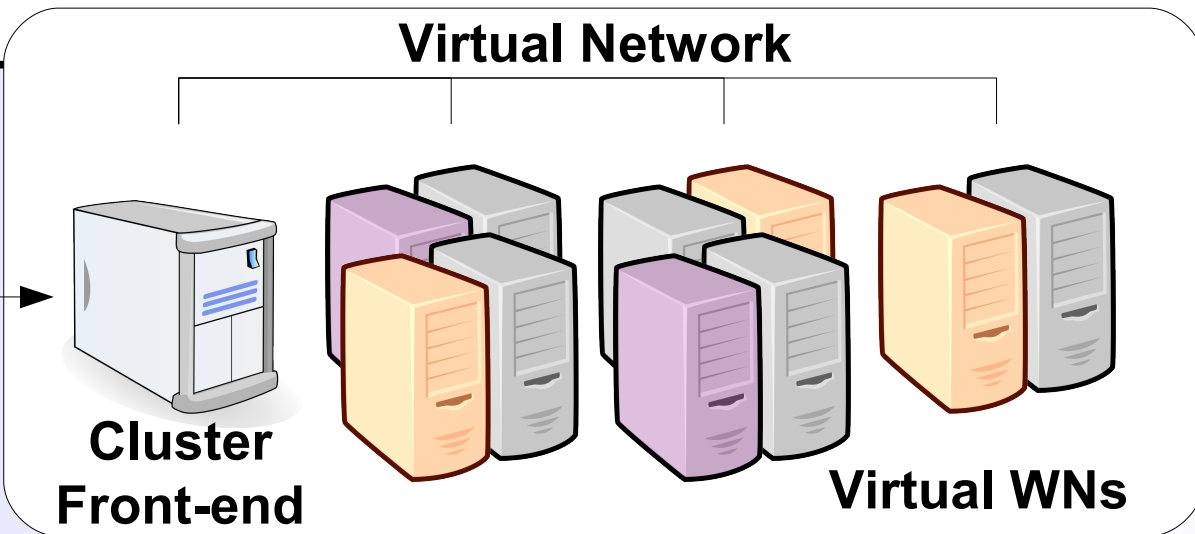
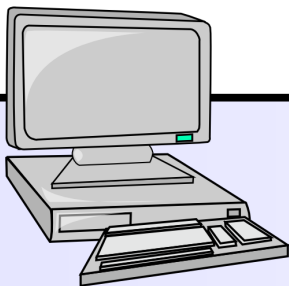
## Service Layer



## Infrastructure Layer

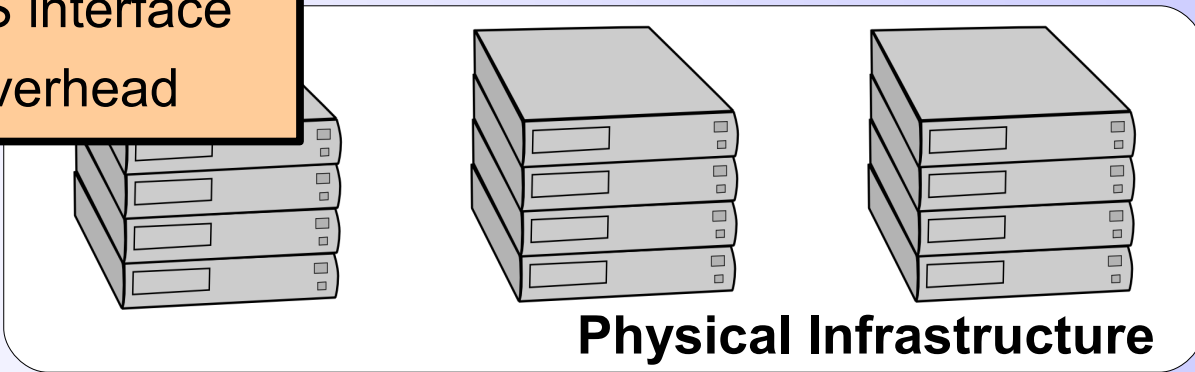
# Grids & Virtual Machines

## Cluster users



## OpenNebula (VIM)

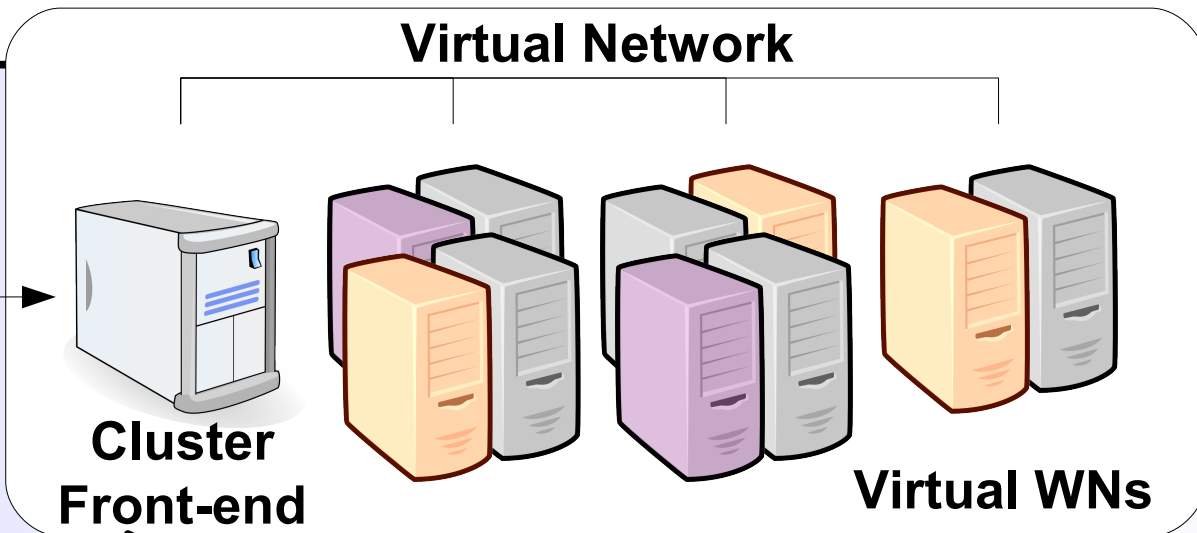
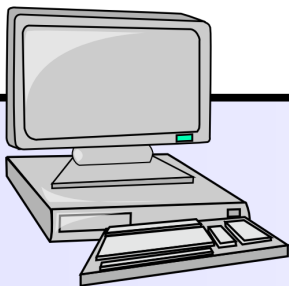
- User Requests**
- “used-to” LRMS interface
- Virtualization overhead



## Infrastructure Layer

# Grids & Virtual Machines

## Cluster users



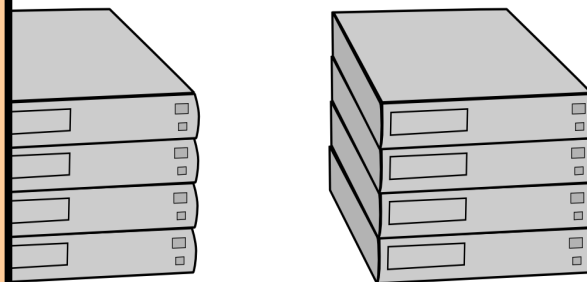
## Service Layer



**OpenNebula (VIM)**

## Cluster Consolidation

- Multiple clusters in a single cluster
- Dynamic provision rules
- Leverage VMM functionality



**Physical Infrastructure**

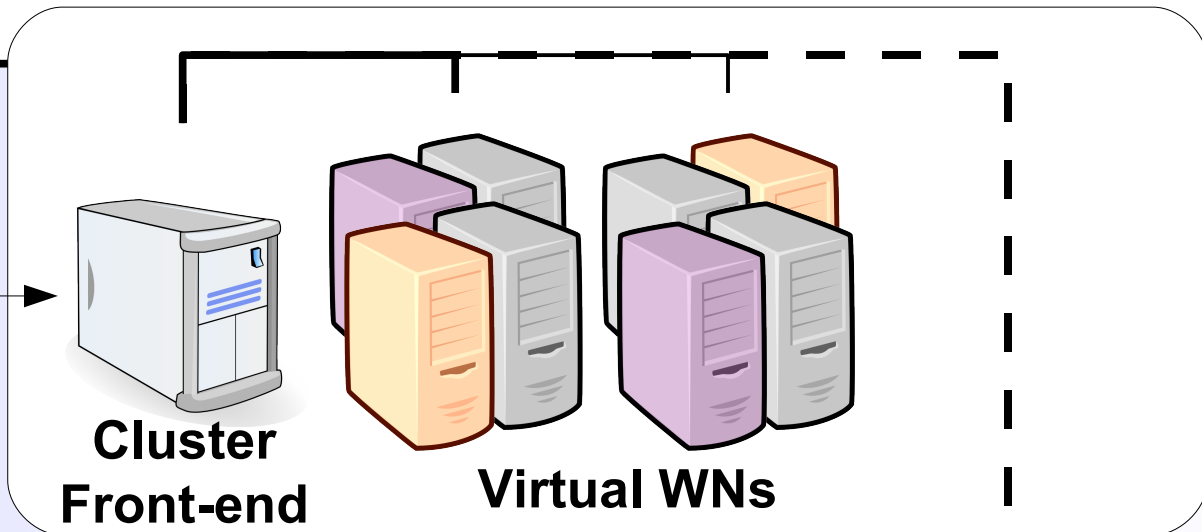
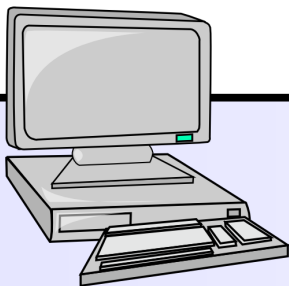
## Infrastructure Layer



# Grids & Virtual Machines



## Cluster users

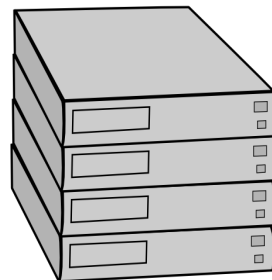
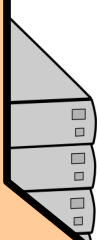


Service Layer

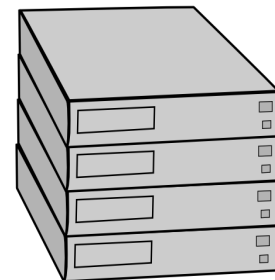
**OpenNebula (VIM)**

## Cluster Partitioning

- Performance partitioning
- Isolate cluster workload
- Dedicated HA partitions



**Physical Infr.**



**Dedicated WN**

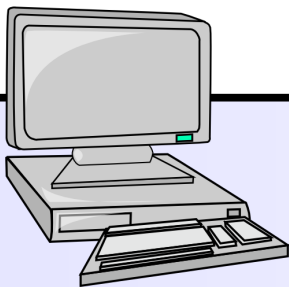
Infrastructure Layer

# Grids & Virtual Machines

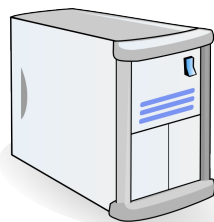


dsa-research.org

Cluster users



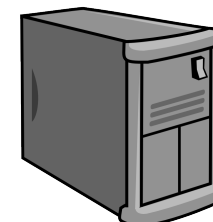
HTTP clients



Cluster Front-end



Virtual WNs



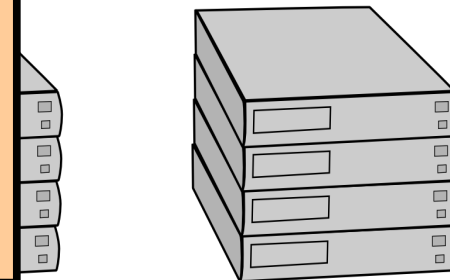
Web Server

Service Layer

Virtual Machine Monitors (VIM)

## Heterogenous Workloads

- Dynamic provision of cluster configurations
- Simultaneous support of different services
- E.g. on-demand VO workernodes in Grids



Physical Infrastructure

Infrastructure Layer

# A Complete Grid Middleware Stack



**Meta-schedulers  
(GridWay, Condor/G...)**

- Unmodified Applications (Grid or local)
- Interfaces preserved (qsub, DRMAA..)

## Applications

**gLite, UNICORE, Globus...**

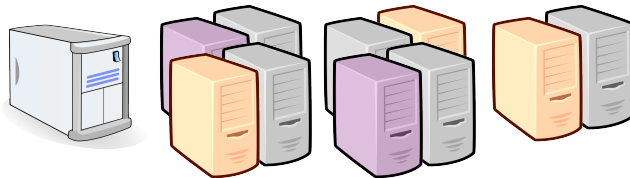
- Virtual resources are exposed by GM
- Dynamic scheduling
- Fault detection & recovery

## Grid Middleware Layer

**Cluster Frontend (SGE...)**

- WNs register to different queues
- Multiple VO-specific clusters

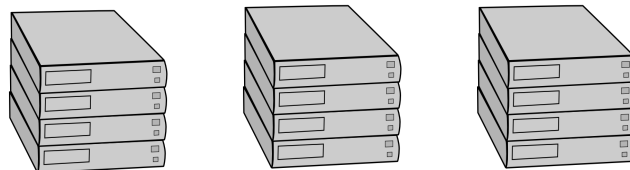
## Computing Service Layer



**OpenNebula (VIM)**

- Infrastructure consolidation
- Infrastructure partitioning
- Infrastructure adaptation

## Infrastructure Layer



# A Complete Grid Middleware Stack



- Unmodified Applications (Grid or local)
- Interfaces preserved (qsub, DRMAA..)

## Applications

**Meta-schedulers  
(GridWay, Condor/G...)**

- Virtual resources are exposed by GM
- Global scheduling
- Fault detection & recovery

**gLite, UNICORE, Globus**

## Grid Middleware Layer

**Cluster Frontends**

- WNs register to different queues
- Multiple VO-specific clusters



## Computing Service Layer

**Virtual Infrastructure (VIM)**

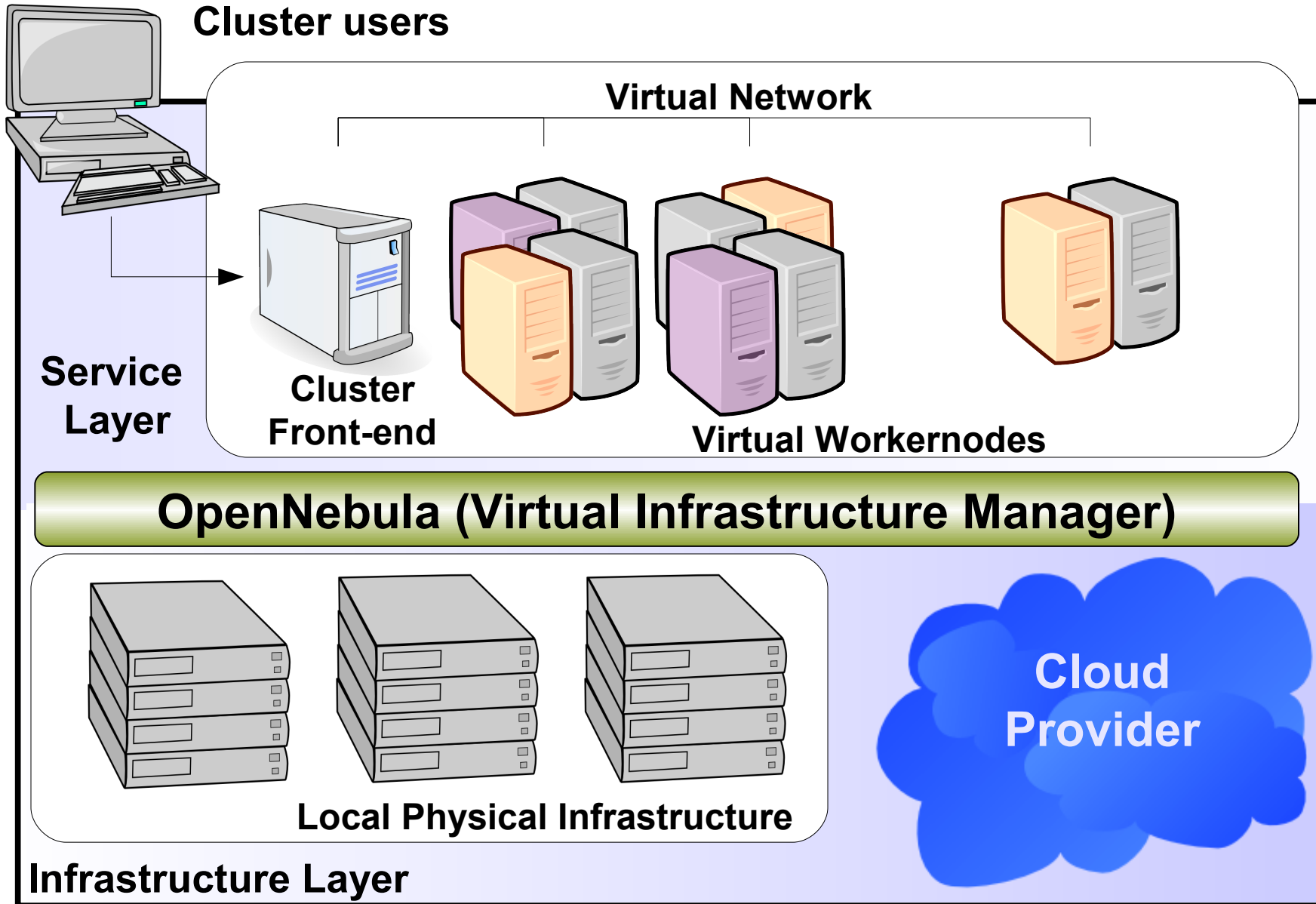
- Infrastructure consolidation
- Infrastructure partitioning
- Infrastructure adaptation



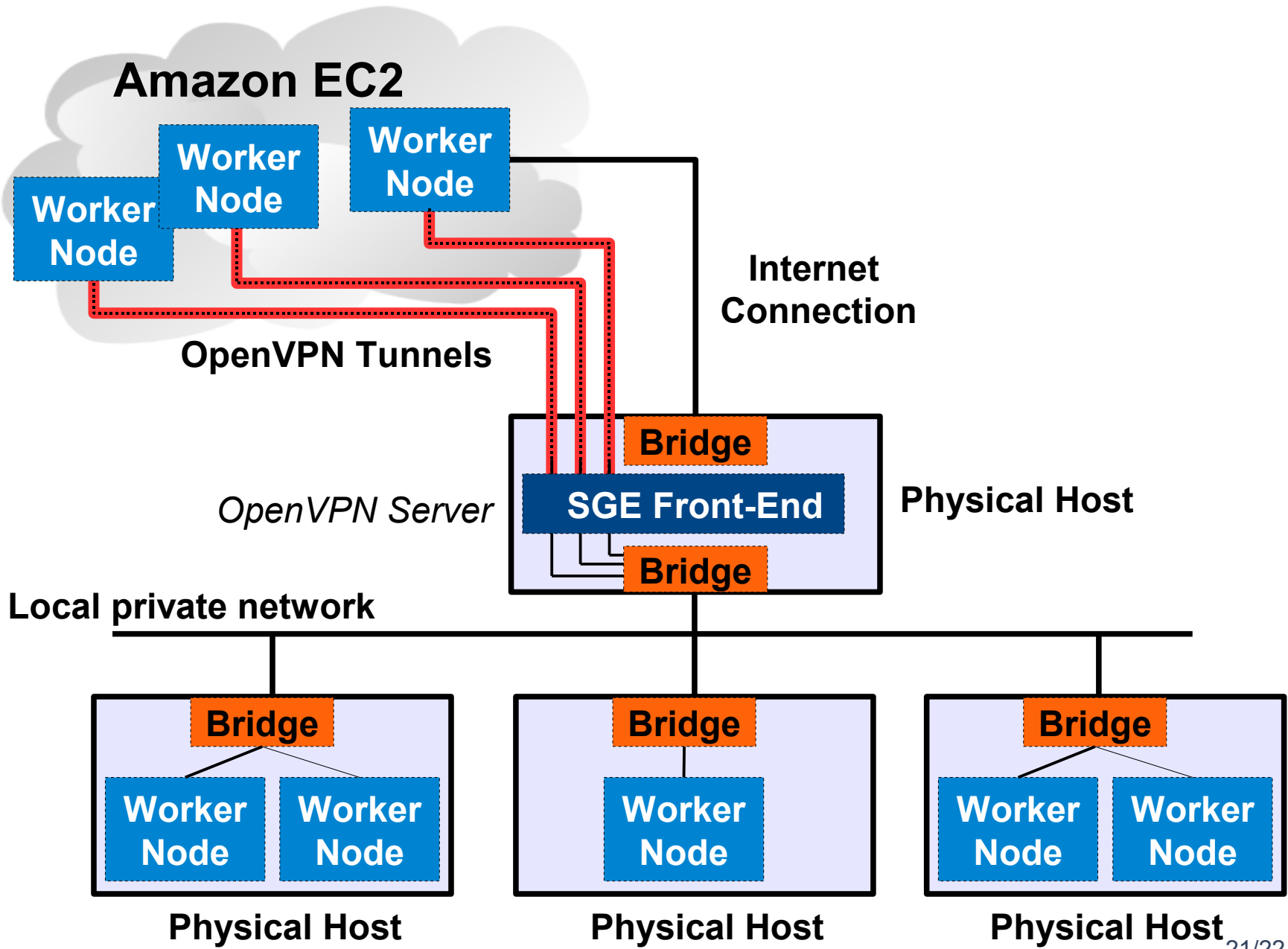
## Infrastructure Layer

**Grid/Cluster as a Service!!!**

# Grids, Clouds and Virtual Machines



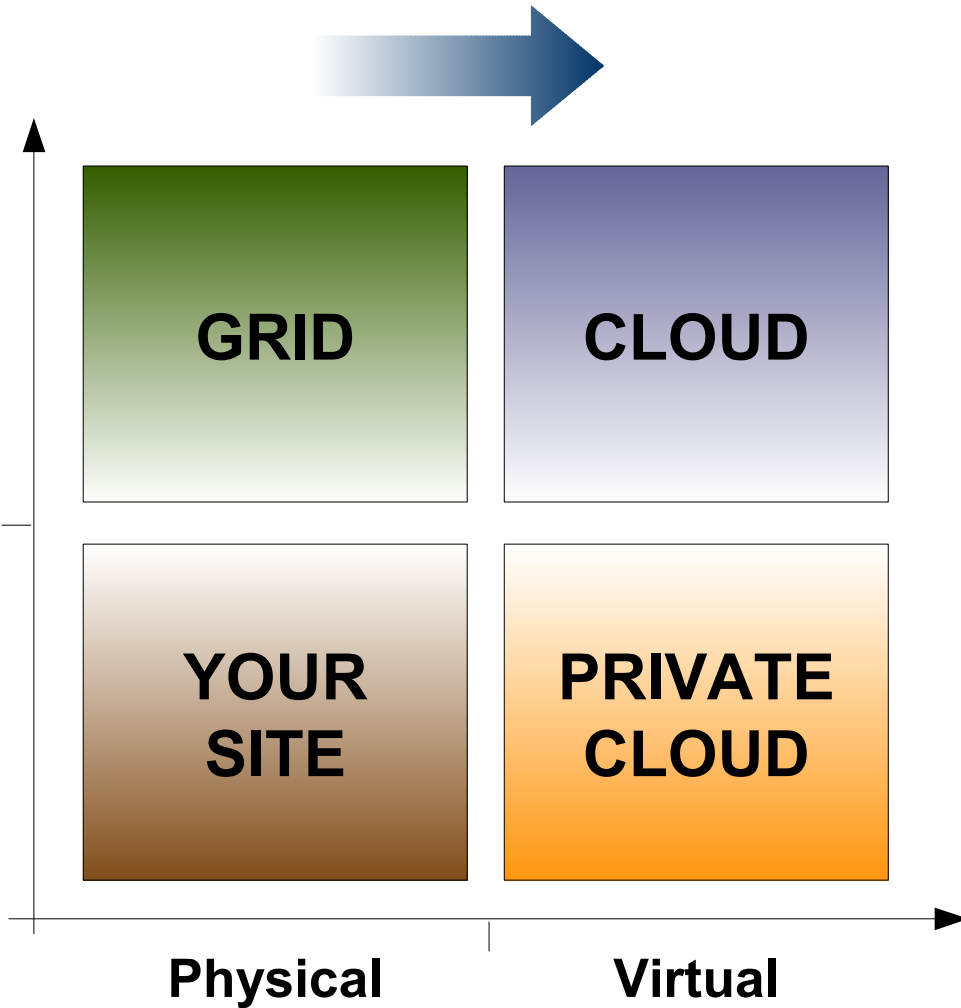
# Tutorial: Hybrid Deployment of a Virtual Cluster



Where are the resources provisioned from?

Remote

Local



How are the resources provisioned?



- Virtualization, cloud, and grid are complementary technologies and will coexist and cooperate at different levels of abstraction
- Virtualization can solve many obstacles for Grid adoption
- Virtualization and cloud do NOT require any modification within service layers (end-user perspective)
- Separation between service and infrastructure layers will allow the application of the utility model to scientific computing *in any form (HPC MPI)*
- Share Hardware not Services (LRMS)!!!

More info, downloads, mailing lists at  
[www.OpenNebula.org](http://www.OpenNebula.org)

OpenNebula is partially funded by the “RESERVOIR– Resources and Services Virtualization without Barriers” project  
EU grant agreement 215605



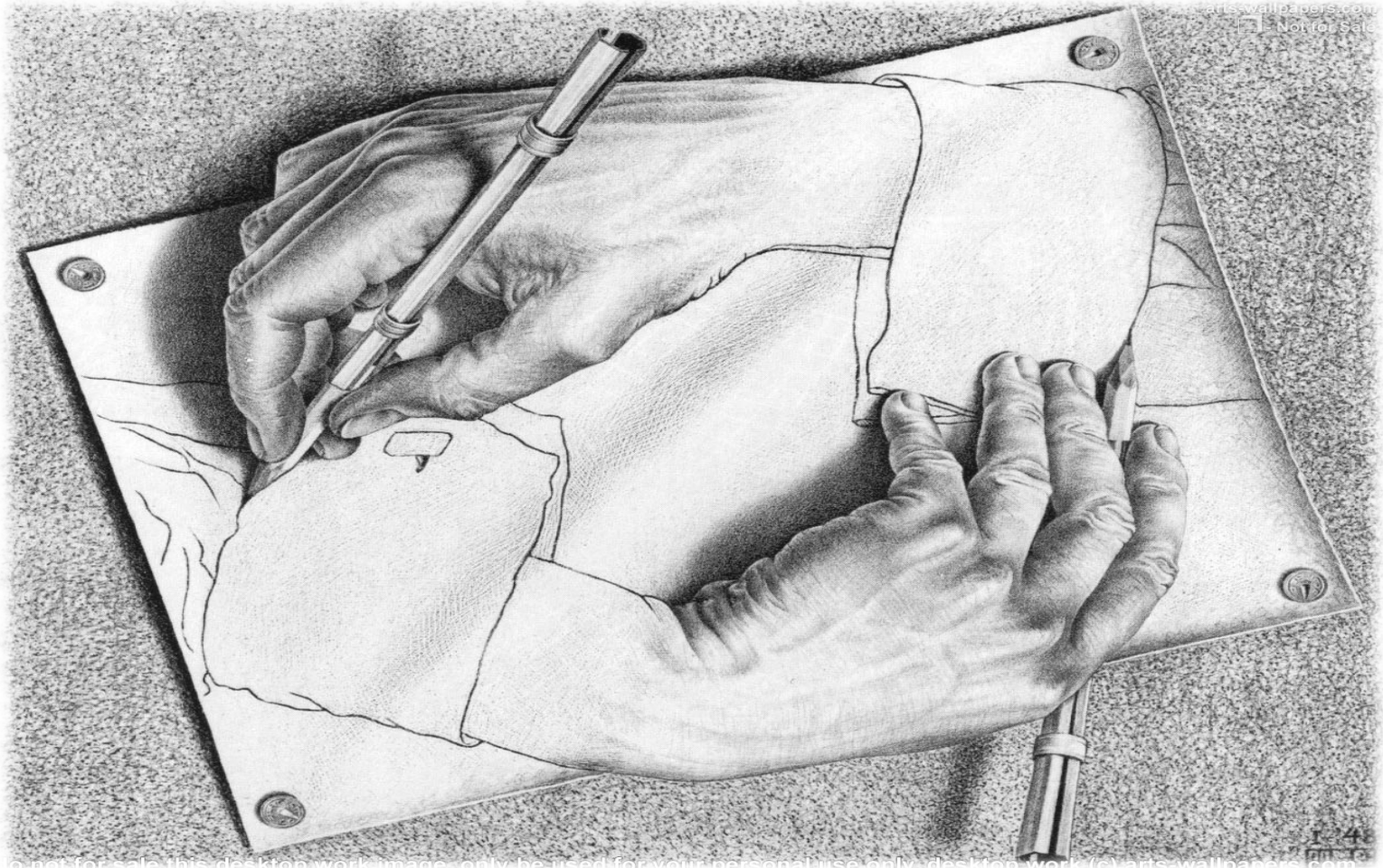
[www.reservoir-fp7.eu/](http://www.reservoir-fp7.eu/)

## The OpenNebula Team

---

- Ignacio M. Llorente ([llorente@dacya.ucm.es](mailto:llorente@dacya.ucm.es))
- Ruben S. Montero ([rubensm@dacya.ucm.es](mailto:rubensm@dacya.ucm.es))
- Rafel Moreno ([rmoreno@dacya.ucm.es](mailto:rmoreno@dacya.ucm.es))
- Tino Vazquez ([tinova@fdi.ucm.es](mailto:tinova@fdi.ucm.es))
- Javier Fontan ([jfontan@fdi.ucm.es](mailto:jfontan@fdi.ucm.es))

# THANK YOU FOR YOUR ATTENTION



## QUESTIONS?