Cloud and Virtualization to Support Grid Infrastructures

Ignacio M. Llorente
dsa-research.org

Distributed Systems Architecture Research Group
Universidad Complutense de Madrid
Objectives

Cloud and Virtualization to Support Grid Infrastructures

• Introduce virtualization and cloud from the perspective of the Grid computing community
• Show the benefits of virtualization and cloud for Grid computing
• Demonstrate how Grid, virtualization and cloud are complementary technologies that will cooperate in future Grid computing infrastructures
• Introduce the RESERVOIR project, European initiative in virtualization and cloud computing
Barriers to Adoption of the Compute Grid Model

Cloud and Virtualization to Support Grid Infrastructures

- High degree of heterogeneity (software & hardware)
- High operational costs
- Difficult isolation and partitioning of resources
- Specific environment requirements for different VOs
- Variability of demand

Grids are difficult to maintain, operate and use
Virtualization Platform

**Cloud and Virtualization to Support Grid Infrastructures**

**Separation of Virtual Machine from Physical Infrastructure**

- A VM is an isolated runtime environment (guest OS and applications)
- Multiple virtual systems (VMs) to run on a single physical system

**Benefits of Virtualization Platforms**

- Natural way to deal with the **heterogeneity** of the infrastructure
- Allow **partitioning and isolating** of physical resources
- Execution of **legacy applications**
Distributed Management of VMs

Cloud and Virtualization to Support Grid Infrastructures

Extending the Benefits of Virtualization to a Physical Cluster

• VM Managers creates a distributed virtualization layer
  • Extend the benefits of VM Monitors from one to multiple resources
  • Decouple the VM (service) from the physical location
• Transform a distributed physical infrastructure into a flexible and elastic virtual infrastructure

Benefits of VM Managers

• Centralized management
• Balance of workload
• Server consolidation
• Dynamic resizing of the infrastructure
• Dynamic cluster partitioning
• Support for heterogeneous workloads
• On-demand provision of VMs
Virtualization of a Computing Cluster

Cloud and Virtualization to Support Grid Infrastructures

Separation of Resource Provisioning from Job Management

- New virtualization layer **between the service and the infrastructure layers**
- **Seamless integration** with the existing middleware stacks.
- **Completely transparent** to the computing service and so end users
Integration of a Virtualized Cluster within a Grid

Cloud and Virtualization to Support Grid Infrastructures

- **Grid Applications**
  - Grid interfaces (DRMAA...)

**Applications**

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

**Grid Middleware**

- Local computing resources

**Local Computing Infrastructure**

- Grid and central services virtualization
- Coexist with other services

**Physical Infrastructure Layer**

- Local computing resources
- Coexist with other services

- Grid and central services virtualization

- Grid and central services virtualization
Integration of a Virtualized Cluster within a Grid

Cloud and Virtualization to Support Grid Infrastructures

Benefits of Virtualization for Existing Grid Infrastructures

The virtualization of the local infrastructure provides:

- Easy support for VO-specific worker nodes
- Reduce gridification cycles
- Dynamic balance of resources between VO’s
- Fault tolerance of key infrastructure components
- Easier deployment and testing of new middleware distributions
- Distribution of pre-configured components
- Cheaper development nodes
- Simplified training machines deployment
- Performance partitioning between local and grid services

Solve many of the obstacles for Grid adoption
Cloud as Provision of Virtualized Resources

- Cloud systems provide **virtualized resources as a service**
- Provide **remote on-demand access to infrastructure** for the execution of virtual machines

**Simple Interfaces for VM Management**
- Submission
- Control
- Monitoring

**Main components of a Cloud architecture:**
- Front-end: Remote interface (Eucalyptus, Globus Nimbus…)
- Back-end: Local VM manager (OpenNebula)

**Infrastructure Cloud Services**
- **Commercial Cloud**: Amazon EC2, GoGrid, Flexiscale…
- **Scientific Cloud**: Nimbus (University of Chicago)
Cloud for Scaling out Local Infrastructures

Cloud and Virtualization to Support Grid Infrastructures

On-demand Access to Cloud Resources

- Supplement local resources with cloud resources to satisfy peak or fluctuating demands
RESERVOIR Project

Cloud and Virtualization to Support Grid Infrastructures

Who?

- IBM (coordinator), Sun, SAP, ED, TID, UCM, UNIME, UMEA, UCL, USI, CETIC, Thales and OGF-Europe
- 17-million and 3-year project partially funded by the European Commission (NESSI Strategic Project)

What?

- The Next Generation Infrastructure for Service Delivery, where resources and services can be transparently and dynamically managed, provisioned and relocated like utilities – virtually “without borders”

How?

- Integration of virtualization technologies with grid computing driven by new techniques for business service management, driven by business use cases

Virtualization-Aware Grid e.g., VM as management unit for metering and billing + Grid-Aware Virtualization e.g., live migration across administrative domains + BSM e.g., policy based manag. of service-level agreement = SOI
RESERVOIR Project

Cloud and Virtualization to Support Grid Infrastructures

The Architecture, main Components and Interfaces

- Service Provider
- Service Manager
- VEE Manager
- VEE Host
- Infrastructure Provider
- VHI
- VMI
- SMI

Create, request, monitor and control services and enforce SLA compliance by managing number and capacity of service components (VEEs).

Organize the placement of groups of VEEs on a distributed pool of resources to meet optimization policies and constraints, supporting the migration of VEEs to partner infrastructure providers.

Extend existing virtualization platforms to support advance functionality for performance and reallocation optimization.
Conclusions

Virtualization, cloud, grid and cluster are complementary technologies and will coexist and cooperate at different levels of abstraction.

Virtualization and cloud do NOT require any modification within service layers from both the administrator and the end-user perspectives.

Separation between service and infrastructure layers will allow the application of the utility model to Grid/cluster/HPC computing.
THANK YOU FOR YOUR ATTENTION!!!
More info, downloads, mailing lists at 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/

Acknowledgements

• Javier Fontan    • Tino Vazquez
• Rubén S. Montero • Rafael Moreno
Computing Cluster Virtualization

Cloud and Virtualization to Support Grid Infrastructures

Cluster users

Service Layer

Cluster Front-end

Virtual workernodes

Distributed Virtualizer

Infrastructure Layer

Physical Infrastructure
Computing Cluster Virtualization

Cloud and Virtualization to Support Grid Infrastructures

Cluster users

User Requests
- Typical LRMS interface
- Virtualization overhead

Infrastructure Layer

Distributed Virtualizer

Physical Infrastructure

Cluster Front-end
Virtual workernodes
Computing Cluster Virtualization

Cloud and Virtualization to Support Grid Infrastructures

Cluster users

Service Layer

Cluster Consolidation
- Multiple worker nodes in a single resource
- Dynamic provision rules (inf. adaptation)
- VMM functionality (e.g. live migration)

Infrastructure Layer

Physical Infrastructure

Cluster Front-end

Distributed Virtualizer

Virtual worker nodes
Computing Cluster Virtualization

Cloud and Virtualization to Support Grid Infrastructures

Cluster users

Cluster Partitioning
- Performance partitioning (dedicated nodes)
- Isolate cluster workload
- Dedicated HA partitions

Service Layer

Distributed Virtualizer

Infrastructure Layer

Physical Infrastructure
Computing Cluster Virtualization

Cloud and Virtualization to Support Grid Infrastructures

Cluster users

HTTP clients

Service Layer

Infrastructure Layer

Distributed Virtualizer

Cluster Front-end

Virtual rnodes

Web Server

Heterogenous Workloads

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

Cloud and Virtualization to Support Grid Infrastructures

Cluster users

Service Layer

Cluster Front-end

Virtual Network

Virtual Workernodes

OpenNebula

Local Physical Infrastructure

Infrastructure Layer

Cloud Provider