GridARS: Resource Management Framework for Multi-domain Cloud

Hidemoto Nakada, Atsuko Takefusa
Tomohiro Kudoh, Ryousei Takano

National Institute of Advanced Industrial Science and Technology (AIST)

This work is partly supported by NICT.
## Cloud computing and IaaS

- IaaS provides IT infrastructure (hardware) on demand.
- IaaS provider owns a large amount of physical resources (computers, storages), and provide a fragment of them to each user through virtualization.

<table>
<thead>
<tr>
<th>Application software</th>
<th>SaaS (Software as a Service)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operating system</td>
<td>PaaS (Platform as a Service)</td>
</tr>
<tr>
<td>System software</td>
<td>IaaS (Infrastructure as a Service)</td>
</tr>
<tr>
<td>Hardware such as</td>
<td></td>
</tr>
<tr>
<td>computers, storages</td>
<td></td>
</tr>
</tbody>
</table>

### Conventional Computing system

### Cloud
IaaS

IaaS provides “Virtual Infrastructures (VI)” to users on demand.

What is VI?

- A VI is an isolated infrastructure
  - Isolated from other VIs or physical infrastructure
- A user can use a VI as if it is a dedicated infrastructure of the user.
- Providers construct VIs using their physical resources (computers, storages), and provides them to users

Shift from “owning” to “using”

- Users do not have to own their own resources
- Infrastructure is provided when needed
Well-organized infrastructure
• Multiple isolated VIs (per request)
• Easy to use, high utilization

(1) Request

(2) Provide

(3) Use

Virtual Infrastructure

Resource Manager

IaaS provider

Data Center
Data centers are the key components

- Data center: a large number of computers and storages located at a place
  - A large number of homogeneous resources reduces the operation cost
- A data center is dynamically shared by a large number of users through virtualization
  - Utilization ratio of physical resources will be high

Role of network

- Network has been considered as an “as is” resource
  - It is not easy to provide data intensive services
  - User experience may be worse than dedicated HW
- Network should be a manageable resource
For data intensive services, chunks of data are moved (copied) between IaaS and user, and inside IaaS.
Multi-domain cloud

(1) Request

(2) Provide

(3) Use
Required technologies for Multi-domain Cloud with manageable network

- A unified interface and provisioning system to request network and other resources (computers storages)
- A mechanism to set up network related application execution environment
  - Dynamic assignment of IP addresses, VLAN ids etc.
- A mechanism to provide users with monitoring capability of virtual infrastructures while keeping isolation
  - Monitoring of not physical but virtual infrastructure
Unified interface and provisioning system

- G-lambda project has been defining an interface
  - Joint project of KDDI R&D labs., NTT, NICT and AIST, started in 2005. http://www.g-lambda.net/
  - The goal of this project is to define a web services interface (GNS-WSI) to request heterogeneous resources (network, computers, storages etc.)

- AIST has been developing a reference implementation of GNS-WSI called GridARS
GridARS provisioning system

- **GRC**: Global Resource Coordinator
- **CRM**: Compute Resource Manager
- **NRM**: Network Resource Manager
- **SRM**: Storage Resource Manager

Scheduling and Allocation

Note: In G-lambda architecture, NRM corresponds to provider NSA of NS-framework of OGF NSI-WG

Allocated Virtual Infrastructure
Application execution management

- A user job is executed on a VI
  - Environment management
    - Determine IP addresses of the hosts
    - Share the IP address list among the hosts
    - Configure network interface of the hosts
      - VLAN, IP address, routing
    - Set up ssh keys, authorized_hosts, known_hosts
    - Set up file system of each hosts
  - Execution management
    - Launch jobs, terminate jobs etc.
    - Confirm network connectivity before launching a job
AEM overview (1)

Scheduling and Allocation

Application Execution Manager

Provisioning

Allocated Virtual Infrastructure
AEM behavior

Request from user/requester

(2) Schedule resources, Including VLAN id, IP addresses

AEM root

GRC

(1) Advertise available IP range, etc.

AEM leaf

CRM

(1) Advertise availability including VLAN

AEM daemon

LS

NRM

NRM

NRM

NW

LS: Local Scheduler (cf. Sun Grid Engine)
AEM behavior

GRC

(3) Reserve Resources, notify IP range, VLAN.

AEM root

AEM leaf

CRM

NRM

NRM

NRM

LS: Local Scheduler (cf. Sun Grid Engine)

NW
AEM behavior

(4) Notify reservation parameters
(5) Register the reservation

LS: Local Scheduler (cf. Sun Grid Engine)

NW: Network
AEM behavior

Start time of provisioning

AEM root
GRC

AEM leaf
CRM

AEM daemon

(6) Launch Node Managers

LS: Local Scheduler (cf. Sun Grid Engine), NM: Node Manager

(6) Network provisioning

LS

NRM

NW

NM

NM

NM

NM

NM

NM

NM

NM

NM

NM
AEM behavior

- NIC config:
  - IP addresses
  - VLAN routing
  - ssh set up
  - File system set up

(7) Gather and distribute compute node configuration information, and configure the nodes

LS: Local Scheduler (cf. Sun Grid Engine), NM: Node Manager

LS: Local Scheduler (cf. Sun Grid Engine), NM: Node Manager

LS: Local Scheduler (cf. Sun Grid Engine), NM: Node Manager
AEM behavior

(8) NM exchange ping to confirm connectivity

LS: Local Scheduler (cf. Sun Grid Engine), NM: Node Manager
AEM behavior

LS: Local Scheduler (cf. Sun Grid Engine), NM: Node Manager

(9) NM Launches user's job

LS: Local Scheduler (cf. Sun Grid Engine), NM: Node Manager
Monitoring

Monitoring information of not physical infrastructure
But provisioned VI should be provided to user

(1) Request

(2) Provide

(3) Use
Monitoring overview

(1) User makes a reservation with a credential

(2) Reservation ID is returned at each level of tree

(3) User requests monitoring with the Reservation ID and the credential. The request is hierarchically propagated using reservation information from GRC

(4) Mapping between rsv. ID and physical infrastructure is provided by GRC and RM

(5) DMS filter monitoring information based on policy, and return to the user

During provisioning, Leaf DMS periodically gather monitoring info.

Provisioning

Allocated Virtual Infrastructure

User

GRC

DMS

NRM

CRM

SRM

GRS

DMS

NRM

CRM

SRM
Demonstration at GLIF2010

Geneva

Tokyo, Japan

Network Domain D1

Network Domain D2

Emulated WAN environment using hardware network emulators

 delay (10msec)

 delay (10msec)

 delay (100msec)
Summary: GridARS key features

- Unified provisioning system of network and compute resources
- Dynamic scheduling and automatic set up of IP addresses, VLAN (host and network), ssh and file system
- Monitoring of virtual infrastructure
  - Provide requester with monitoring information of the provisioned infrastructure (network and CPU) only
  - The monitoring system interoperates with the provisioning system
- Policy based filtering
  - Provider can define information to be provided to a particular requester
Come and see our demonstration this evening!