Virtual Resource Management (VRM) in Cloud Environment

draft-junsheng-cloud-vrm-00

Chu JunSheng, B. Khasnabish, and Meng Yu

IETF-79 Clouds Mtg.
Sunday 7 Nov 2010
Pearl Rm., Shangri-La hotel, Beijing
Introduction

- Applications/Services and Cloud provided basic functions are based on the Virtual Resources which are abstracted from physical resources.
- Virtual Resources include V-CPU, V-Storage, V-Network etc, and Virtual Resource can be further abstracted to Virtual Network Resources, such as V-Router, V-Switch, V-Firewall, VPN, V-Interface, V-Link.
- Virtual Resources are integrated as a whole (Resource Pool) and supplied to upper layers. Virtual Resources can be allocated and released on demand.
- In Cloud, Physical Resources can be divided into many Server Clusters, and Virtual Resources can be further divided into many Virtual Resource Clusters in Resource Pool.
- VRM can allocate resources for applications/services on multiple Virtual Machines which may belong on several Physical Servers in the same Cluster or different Clusters of the same Cloud, or even different Clouds.
Configuration of VRM

Our Scope

VRC: Virtual Resource Client
VRM: Virtual Resource Management
VRP: Virtual Resource Provider
Virtual Resource Description

- Virtual Resource Info
  - Type: V-CPU, V-Memory, V-Storage, V-Firewall, V-Router, V-Switch, V-Interface, V-Link, etc.
  - Subtype: V-Storage(V-Database, V-Filesystem, V-Object, V-Block, etc).
  - Status: Idle, Running
  - Capacity
  - Usage Ratio
  - Virtual Machine
  - Virtual Cluster
  - Physical Cluster
  - Physical Server
  - etc

- Two kinds of resource info for virtual resource pools need be maintained
  - Proprietary resource pool for private network
  - Open resource pool for public network
Hierarchical Resource Organization

Application Service Layer
- Customer
- Customer
- Customer

Virtual Resource Client
- Service Portal
- Application/Service Control

Cloud Management
- Configuration Mgmt.

Resource Control Layer
- Virtual Resource Manager

Virtualization Layer
- Virtual Resource Cluster 1
- Virtual Resource Cluster 2
- Virtual Resource Cluster 3
- Virtual Resource Cluster 4

Virtual Resource Provider

Physical Layer
- Cluster 1
- Cluster 2
- Cluster 3
- Other Cloud
- Server
Virtual Resource Management

- VRM provides Virtual Resource Register & Deregister, Virtual Resource Discovery function to VRC and VRP.
  - When powered on, VRP registers to VRM immediately to deliver resources.
  - When powered off, VRP deregisters to VRM to cancel delivered resources.
  - When VRC needs resources to perform services, it queries VRM to discover suitable resources.
  - VRM checks the required resources from delivered resource pool and responses eligible virtual resources to VRC.
  - VRC interacts with VRP to bind and utilize the allocated resources for allocated time.
Virtual Network Resource (VNR)

- Virtual Machines further abstracts virtual resources in the shared resource pools into Virtual Network Resources, such as
  - V-Switch
  - V-Router
  - V-Firewall,
  - VPN
  - V-Network Interface/Link

- Based on business demands, these Virtual Network Resource can be constructed to provide specific service-levels or meet particular business needs.

- By defining a set of standardized/generalized Virtual Network Resources, Applications/Services can be easily realized, deployed and extended by VRM.
Resource Admission Control (RAC)

- Aggregate computing/storage/network capacity across a collection of servers into logical resource pools.
- The cloud can share resources of another cloud, RAC also has the responsibility to aggregate them into logical resource pools.
- Initiate one or a group of Virtual Machine(s), when Application/Service is deployed on the Cloud.
  - RAC will check if there are enough unreserved resources in the shared resource pools can be provided to the Virtual Machine(s).
  - If enough unreserved resources are available, the virtual machine is powered on. Otherwise, an Insufficient Resources warning appears.
Resource Authentication and Authorization (RAA)

- When Application/Services request RAM to allocate (extra) resources, RAA will check whether they are authorized the capabilities.
- When the Cloud provides resources to another Cloud, it will authenticate the inter-cloud resource request. If passed, the required amount of resources will be provided.
- Also, in case of another Cloud providing resources to the Cloud, it will check the security of the resources. If reliable, it will authorize the resources and put them into the virtual resource pools.
Resource Scheduler (RS)

- RS defines rules and policies to decide how resources should be prioritized among virtual machines and intelligently allocates available resources from virtual resource pools among the virtual machines based on the pre-defined rules that reflect Applications/Services requirements and changing priorities.
- Consolidate workloads during off-peak hours and power off hosts to make datacenters more energy efficient.
- Bring powered off hosts online to meet virtual machine requirements either at a pre-defined time or when the utilization of virtual machine increases.
Inter-Cloud Resource Control (ICRC)

- Resources in one cloud can be shared with the resources in another cloud in certain circumstances.
- ICRC may provide the following functions:
  - Signaling with other cloud
  - Control and monitoring the resources in other cloud
  - Transport and security with other cloud
  - Naming, Addressing and Translation if they have different format.
Conclusion

- Classifying virtual resources into clusters is the efficient way to manage and schedule virtual resources.
- Virtual Network Resources can be further abstracted from Virtual Resources, which can be flexibly used to realize virtual network, telecom network virtualization, and VLAN related services.
- Delivery of Virtual Resources to VRM needs to be considered when construct Virtual Cluster Resource Pool.
- VRM schedules resources in an efficient way among virtual machines based on pre-defined policy and rules to satisfy business requirements.
- VRM mainly focuses on allocation of Virtual Resources efficiently, securely, flexibly and reliably.
Next Step

- Further develop the requirements
- Define the interfaces and protocols
- Develop the profile for the protocols
- Work on the draft
  - draft-junsheng-cloud-vrm-00
Comments/Suggestions
Thanks!