



Deploying Business Virtual Appliances on Open Source Cloud Computing

Tran Van Lang¹ and Nguyen Trong Duc²

¹Institute of Applied Mechanics and Information Technology (IAMI/VAST)

²Post and Telecommunications Institute of Technology, HCMCity

¹tvlang@vast-hcm.ac.vn, ²nguyentrongduc@gmail.com

Abstract– The Cloud computing emerges as a new computing paradigm which aims to provide reliable, customized and QoS guaranteed dynamic computing environments for end-users. Virtual machines (VMs) provide a virtualization solution to the security and resource management issues that arise in isolated environment. We encounter the need to deploy an open source cloud computing and integrate VM (a virtual business server) with business applications into cloud models and the existing resource management infrastructure. To address such requirements, the paper describes an approach to deploying business virtual appliances on Open-source Scientific Cloud computing based on the Globus Virtual Workspace Service (Nimbus project) as an Infrastructure as a Service.

Index Terms– Cloud Computing, Grid Computing, Virtualization, IaaS and Virtual Appliance

I. INTRODUCTION

Recently, cloud computing [1] has been under a growing spotlight as a possible solution for providing a flexible, on demand computing infrastructure for a number of applications. Cloud computing is recently offered by companies like Google, IBM [2], Amazon [3], and Sun, etc. With increasing demand from business for higher efficiency, productivity, business agility, and lower cost, for several years, information communication technologies (ICT) have been shifting dramatically from static local environment with manually managed resources and applications towards dynamic virtual environments with automated and shared services with service-oriented architectures. In order to find other computing infrastructure solution for the flexible and effective computing resource for enterprises, one of solution that has come recent years is cloud computing.

The research try to deploy an internal or private cloud similar to Amazon EC2/S3 services that provide Infrastructure as a Service (I-a-a-S), it means a set of virtual servers that work together through the Internet and can be dynamically managed, monitored, and maintained. Users are expected to develop their own virtual images or use existing

ones as an executable environment on the cloud. Using virtual machines (VMs) that can be configured before deployment has the potential to reduce inefficient resource allocation and excess overhead. A VM can create an environment on a resource that is configured independently from that resource, allowing multiple such environments to be deployed on the same resource at the same time. In this manner of separation, each environment is kept secure from any others. Because sharing can be much more flexible, this also can also increase resource utilization [4].

The approaching of this research was focused on open source grid computing, cloud computing and virtualization technologies. That approach based on the grid computing middleware Globus Toolkit-4 [15], Globus Virtual Workspace service [5], Virtual Machine Monitor (VMM) Xen Hypervisor [6], and some networking technologies on Linux OS (Ubuntu Server). In addition, I also tried to deploy some configured virtual machine images as virtual appliances with some popular business applications that run on open source environments with Linux OS, Apache Web Server, MySQL, PostgreSQL RDBMS, PHP or JAVA.

The main goal of this work was to test the applicability of science cloud computing to business application environments that are usually run on the local resource computing or private IT infrastructures.

The remainder of the paper was organized as follows. Section 2 presents the approach methods to solve problem. In section 3, the experimental environment was presented. Then, some conclusions and evaluations were given in final section.

II. APPROACH

This approach relied on implementing the Globus Workspace Services for IaaS cloud that enable resource leasing with VMs and deploying business applications which based on the configured VMs or virtual appliances.

A. The Use of Virtual Machine (VM) Technology to Implement Virtual Appliance

VMs present the image of a dedicated raw machine to the business application; Virtual appliance is a configured VM from the operating system to the applications and the use of VMs for business application becomes more common. A

configured VM can be deployed on many different sites or hosting machine without requiring the resource providers to understand the application and its dependencies. An application running on a VM is decoupled from the system software of the underlying host machine.

VMs also enable resource allocation for specific business jobs or applications. It is hence feasible to restrict the memory, network, disk size, and even the CPU cycles allocated to a given VM. Furthermore, the use of VMs allows the target execution environment for a business application to be completely customized, thereby enabling support for jobs with special requirements like root access or legacy dependencies. VMs also enable process migration without requiring any modification or re-linking of the business application.

B. Globus Virtual Workspace Services

Virtual Workspaces (VW) are described by workspace metadata (XML Schemas), which contains all the information needed to deploy a workspace in various contexts. An atomic workspace, representing a single execution environment, specifies the data (e.g., VM images) that must be obtained and deployment information (such as networking setup) that need to be configured on deployment. At deployment time, a workspace is associated with a resource allocation, which describes how much resource (CPU, memory, etc.) is assigned to the workspace.

Globus Virtual Workspace Services (VWS) [5] is a Globus Toolkit 4 (GT4) [15] component based service used to deploy workspaces. The Workspace Service implements a WSRF-based protocol allowing remote clients to start, stop, manage, and query the workspaces. All client actions are authorized using the Grid Security Infrastructure (GSI). VWS functionality is exposed via web services implemented in the GT4 container. Providing essential functionality are the Workspace Factory Service, for generating new workspaces; and the Workspace Service, for managing existing workspaces. The Workspace Client communicates with these services, to authenticate users and control virtual workspaces. VWS maintains an internal resource list of worker node hosts (each available node is identified as a resource slot), stored in the workspace persistence database.

Currently, Xen Hypervisor [6] is the main VMM available to use with VWS. Xen is a software process that manages the hardware resources of the real machine among instances of VMs, thus allowing multiple instances of VMs to run simultaneously on the same hardware. Xen requires source modifications to guest operating systems to run them in a hypervisor.

C. Virtual Workspace Service Interactions

Assume that the Workspace Service executes on a service node of a physical cluster and provides a secure gateway to a set of resources that can support the deployment of virtual machines. All VM images necessary for deployment has already been staged to a node in the trusted computing base (TCB). The service node of the physical cluster runs a GT4 container and the Workspace Service. The hosts are configured with the Xen hypervisor, Workspace Service back-

end scripts and some means of transferring image files and other data relevant to the workspace from within the TCB.

To deploy a VW on a specific resource, a client contacts a VW Manager Grid service on that resource and presents it the workspace's end-point reference (EPR). In order to create a workspace instance, a Grid client contacts the VW Factory with a workspace description. A negotiation process may take place to ensure that the workspace is created in a policy controlled way. The newly created workspace is registered with a VW Repository, which provides a Grid service interface allowing for inspection and management of workspaces and keeps track of resources implementing workspaces such as virtual machine images. As a result of creation the client is returned a WSRF EPR to the workspace.

The Workspace Service maintains a database of information about physical hosts available for workspace deployment. For each physical host it records availability, CPU type, total/available memory size, total/available disk size, and system information. When the Workspace Service receives the workspace creation request, it searches the database for a set of resources matching the resource allocation request, defines a matching set, marks it as reserved, and maps the resource allocation onto it. When the workspace is terminated, the resources are reclaimed and the database is modified accordingly.

In addition to allocating resources, the Workspace Service also handles local IP address allocation. The first step of workspace deployment involves propagating the images to the target nodes: workspace scripts executing on each node download the images from a specified location for pre-staging of the images without actually starting the VMs. The back-end scripts work with the Xen hypervisor and complete the configuration of the workspace. Configuration information that cannot be processed by Xen (such as networking) is set up by calling an OS boot script preinstalled in the VM images.

After a workspace is deployed, a program can be started by using Grid infrastructure mechanisms (e.g., Globus Resource Allocation Manager, or GRAM) or by using other methods such as preconfigured program startup or a continuation of a previous execution. The VW Manager can also stop, pause, or undeploy a stopped or paused workspace by invoking start and stop operations with different parameters [7].

III. EXPERIMENTAL SETUP

First we need to build a science cloud using the Globus Virtual Workspace (or Nimbus Toolkit project) in our testing environment and then try to deploy some virtual appliances that will run business servers have configured applications. Some installation manuals reference to Globus Toolkit 4 Installation Guide, Nimbus Administration Guide [8], and Ubuntu Server Guide [9].

A. Testing Environment

Experimental environment is a simple private cloud model with some networked physical servers which were setup one head node, one cloud client node and four worker nodes. Most of nodes must be installed a stable and long term support

version of Linux Ubuntu Server 8.04.3 LTS (Hardy Heron), On cloud client node can be used linux desktop with graphic user interface. All worker nodes must be installed VMM – Xen Hypervisor 3.2 with Xen OS kernel and DHCPd service. Password-less SSH and NFS were configured within these

servers to allow automated file transfers and remote command executions. The nodes were time-synchronized using NTP as required by Globus Virtual Workspace. Fig. 1 shows the network diagram of testing environment and Table 1 shows the characteristics of machines I used in my experiments.

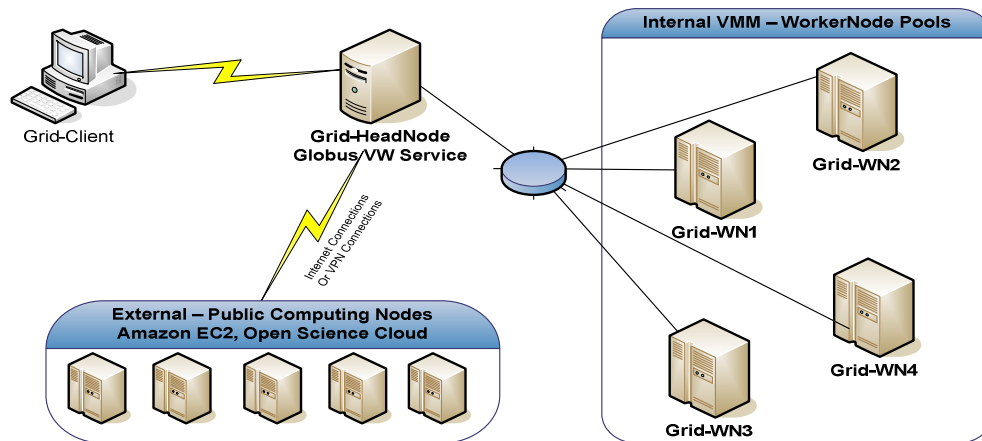


Fig. 1: The network diagram for testing private cloud computing

Table 1: The characteristics of physical servers in testing environment

Node Name	Hardware Configurations	Networks	Role
Grid-HN	CPU Core 2 dual 2.2 GHz, RAM 512MB, HDD 04GB, 02 NIC	192.168.0.130 (private NIC) 192.168.100.130 (public NIC)	Head Node or VW Service Node.
Grid-WN1	CPU Core 2 dual 2.4 GHz, RAM 02GB, HDD 250GB, 01 NIC	192.168.0.131	Worker or Resource Node.
Grid-WN2	CPU Core 2 dual 2.4 GHz, RAM 02GB, HDD 250GB, 01 NIC	192.168.0.132	Worker Node
Grid-WN3	CPU Core 2 dual 2.4 GHz, RAM 02GB, HDD 250GB, 01 NIC	192.168.0.133	Worker Node
Grid-WN4	CPU Core 2 dual 2.2 GHz, RAM 512MB, HDD 04GB, 01 NIC	192.168.0.134	Worker Node
Grid-Client	CPU Core 2 dual 2.2 GHz, RAM 512MB, HDD 04GB, 01 NIC	192.168.0.254	Cloud Client Node
Network Configurations:	Gateway: 192.168.0.1 Subnet Mask: 255.255.255.0 Primary DNS: 192.168.0.1 Secondary DNS: 192.168.0.1 Domain: grid-hn.mycloud.com		
IP arranges for VMs	192.168.0.151 ~ 192.168.0.160 = 20 IPs.		

B. Science Cloud Computing Installation with IaaS Model

Several tools were used to set up this experiment, the first of which are the Globus Toolkit 4.0.8 (stable version that is compatible with current using linux version Ubuntu Server 8.043 LTS), Virtual Workspace Service TP2.2 and Workspace cloud client 1.4 [7]. The Virtual Workspace Service is installed on a Grid architecture which was aim to create the science cloud, and allowing users to create virtual environments using the workspace cloud client that allows a user to upload virtual machine images, download, modify, delete, or save copies of preexisting images, and deploy

images as virtual machines. In addition, there were many linux networking tools and system services need to be installed and configured in order to match all system requirements such as NFS, DHCP, RFT (Reliable File Transfer), GridFTP (a high speed transfer service that expands the File Transfer Protocol to include features needed for Grid computing applications), GSI-X509, Password Less-SSH, Xen VMM as well as Java & Python runtime. Before any workflow could be run on the science cloud, the appropriate above softwares and toolshad to be installed and configured on the experimental environment [8].

In the following layout, the paper describes the layout of science cloud computing model with Infrastructure as a Service. In this architecture, there was one node for virtual machine disk image repository which will store all cloud user's VM images. The specific requested VM image will be transferred to resource worker node and then will be taken by Xen Hypervisor in order to boot the VM's guest operating

system. The physical cluster head node, was installed with Globus virtual workspace service, was as a GT4 Web service interface which receive all virtual workspace deploying request from virtual workspace clients node(s). The steps were taken by the VM-based GVW science cloud model to establish a virtual machine session for a user as follows (refer to Fig. 2):

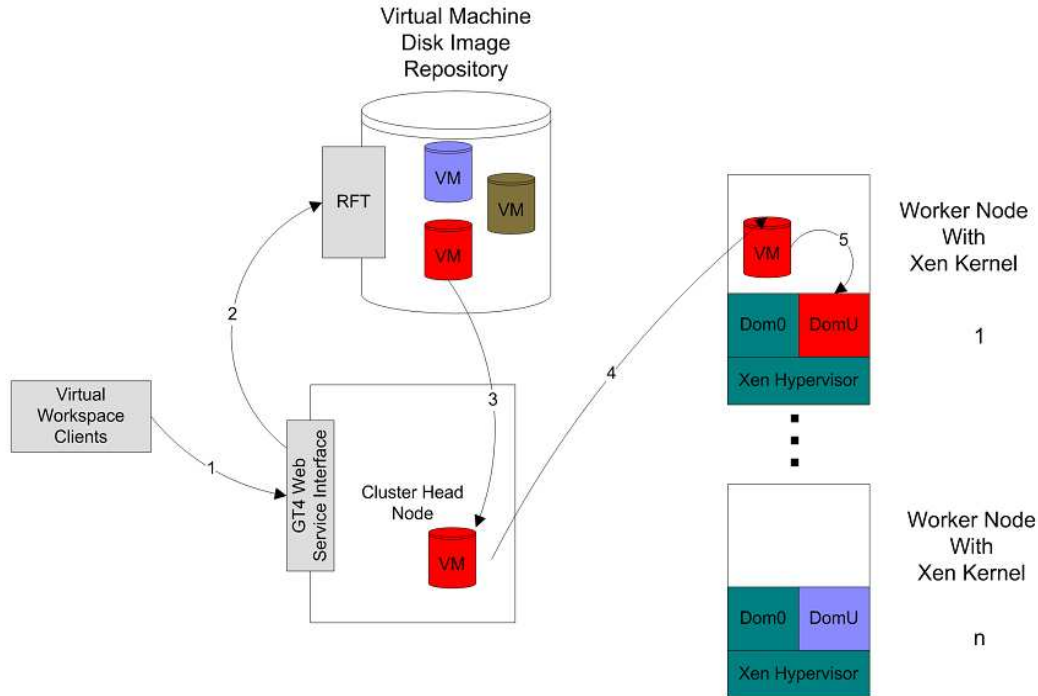


Fig. 2: Science cloud in experimental environment with IaaS Model

The diagram above depicts the basic steps: (1) A special workspace client called the "cloud-client" invokes workspace deployment operations on the cluster head node service; (2) Files are transferred from the cloud-client to a client-specific directory on the repository node (RFT or other types of GridFTP based transfers); (3,4) The service invokes commands on the VMs to trigger file transfers from the repository node; (5) The VM instance is running in Xen DomU and starts its lifecycle events, if the workspace state changes, the cloud-client will reflect this to the screen (and log files) and depending on the change might also take action in response. User can access and control the VM via SSH Telnet login session, or a virtual display session such as VNC.

C. Virtual Business Servers Deployment

Our built cloud computing in the above test environment are resources for the IT computing infrastructures, we need to build some VM images which are fully configured images (virtual appliance) with installed business applications, and coordinate the mapping of virtual appliances onto those available resource platforms. We can create the virtual appliance by ourselves or buy from appliance providers.

Some popular appliance providers such as VMware [10], JumpBox [11], TurnKey [12] or BitrixSoft, provided fully customized virtual appliances (disk images) that will launch them in the cloud or computing environment that supports virtualization (all major forms of virtualization like VMware, Parallels, Microsoft, Xen Open Source, VirtualBox and Amazon EC2).

According to experiments environment with Globus Virtual Workspace Services and Xen Hypervisor, so we could only deploy open source applications for business solutions with more than fifty-five different applications spanning all major product categories. The deploying virtual server were installed Open Source server software (such as Ubuntu 8.0.4 LTS operating system, Apache2 HTTP Server, PHP 5.2.13, MySQL5 with InnoDB support, IP address acquired via DHCP) and some popular applications are JoomlaCMS 1.5.12, vTigerCRM 5.1.0, and ProjectPier PM 0.8.3. These virtual servers need resources with 03GB disk, 512-1024KB RAM and 1GHz CPU. As experiment results, these business virtual servers ran smoothly on the GVW Open-source Scientific Cloud computing.



Fig. 3: Running business virtual servers on Open Source Cloud Computing

IV. CONCLUSIONS

Cloud computing will become important dynamic components of enterprise Grid computing, that provide Infrastructure as a Service (IaaS) and add a new external dimension of business flexibility by enhancing their capacity whenever needed, on demand. Providers of compute cycles in the cloud, such as Amazon EC2 [3] or the Science Clouds [5], enable users to acquire on-demand compute resources, usually in the form of virtual machines (VMs). Pre-configured VM images are called virtual appliance which can be deployed on the remote resources like cloud computing and contain stacks of open source softwares for business requirements.

This paper has described the approaching of using Globus Virtual Workspace [5] for building open source science cloud computing. Then it had showed an abstraction of hardware architecture and required softwares installation in order to be built the private cloud in an experimental environment. This cloud environment will provide resource service as an 'Infrastructure as a Service' model, and provide an isolated virtualized execution environment for deploying workspaces, then launching VMs. To evaluate the feasibility of open source cloud implementations for business applications, we have created and deployed some business virtual appliances and they have run smoothly on experimental cloud environment.

Future works will be needed in order to fully assess the usefulness of these ideas: (1) Because of the limitation of business application on open source environments, we will focus on research the virtual appliance with Windows Desktop Server (as guest OS). (2) There are several open source cloud solution such as OpenNebula [13], Eucalyptus [14] which need to be researched and estimated for business applications.

REFERENCES

- [1] Ricadela A., Computing Heads for the Clouds, in Business Week, November 16, 2007.
- [2] Lohr S., Google and IBM Join in Cloud Computing Research, New York Times, 2007.
- [3] Amazon Elastic Compute Cloud, <http://aws.amazon.com/ec2/>.
- [4] Foster I., Freeman T., et al., Virtual Clusters for Grid Communities, CCGRID, 2006.
- [5] Globus Virtual Workspace Services, <http://workspace.globus.org>.
- [6] Barham P., Dragovic B., et al., Xen and the art of virtualization, ACM SOSP, pp. 164-177, 2003.
- [7] Kate Keahey, Tim Freeman, Science Clouds: Early Experiences in Cloud Computing for Scientific Applications, 2008.
- [8] Nimbus Administration Guide, <http://workspace.globus.org/vm/TP2.0/admin/index.html>
- [9] Ubuntu Server Guide, <https://help.ubuntu.com/8.04/serverguide/C/index.html>
- [10] VMware Virtual Appliance Marketplace, <http://www.vmware.com/appliances/>
- [11] JumpBox Virtual Machine & Virtual Appliance Library, <http://www.jumpbox.com/library>
- [12] TurnKey Linux Virtual Appliance Library, <http://www.turnkeylinux.org/>
- [13] OpenNebula: The Open Source Toolkit for Cloud Computing, <http://www.opennebula.org/>
- [14] Eucalyptus Systems: Cloud computing software, <http://www.eucalyptus.com/>
- [15] Tran Van Lang, Grid computing: Building computing system deploying applications, Vietnam Education Publishing House, 2008, 196p.