Middle-East Journal of Scientific Research 19 (12): 1664-1668, 2014
ISSN 1990-9233

DOI: 10.5829/idosi.mejsr.2014.19.12.11285

Virtual Machines for Cloud Computing in Copy on Write Method

S. Archana and B. Nagalakshmi

Department of CSE,
Bharath University, India

Abstract: The commercial success of Cloud computing and recent developments in Grid computing have brought platform virtualization technology into the field of high performance computing. Virtualization offers both more flexibility and security through custom user images and user isolation. In this paper, we deal with the problem of distributing virtual machine (VM) images to a set of distributed compute nodes in a Cross-Cloud computing environment, i.e., the connection of two or more Cloud computing sites. Ambrust et al. identified data transfer bottlenecks as one of the obstacles Cloud computing has to solve to be a commercial success. Several methods for distributing VM images are presented, and optimizations based on copy on write layers are discussed. The performance of the presented solutions and the security overhead is evaluated.

Key words: Cross-Cloud computing · VM images are presented · Optimizations based on copy

INTRODUCTION

The Cloud computing paradigm is sometimes viewed as the commercial successor of the academic Grid computing paradigm. Although this view is debatable, there is a strong trend towards this new paradigm. Cloud computing as offered by companies such as Amazon (also called Infrastructure as a Service (IaaS)) is easier to use than the typical Grid and offers more user rights (i.e. super-user rights) and tighter security. This has led to much greater commercial adoption of Cloud computing than of Grid computing. However, unlike a Grid, a Cloud does not assist its users with the selection and use of different remote Cloud sites. Due to commercial interests, a Cloud provider usually has a proprietary system where a user creates an image that only runs on the provider's site. This represents a significant step back compared to Grid computing. While many cross-site technologies applied in Grid computing such as virtual organization management and single sign-on can quite easily be adopted by Cloud computing, the actual configuration and management of user software in the Cloud is significantly different compared to the traditional Grid. In the Cloud, users expect to be able to install complex software with super-user privileges on-demand. First, we briefly present a Cross-Cloud virtual machine creation solution that is based on a layered approach to allow a single user image to be deployed to multiple Cloud sites, including desktop Cloud sites. Next, we analyze several methods for the bulk transfer of these layered images to the Cloud sites. Unlike the transfer of virtual machines in Grid computing, which can be planned and executed in advance by the metascheduler, this transfer must be performed on-demand without pre-knowledge of the transfer or the recipients of the transfer. This creates a much tighter time constraint.

Literature Survey

Secure On-Demand Cloud Computing: M. Smith, M. Schmidt, N. Fallenbek, T. Doernemann, C. Schridde and B. Freisleben [1]: Virtual machines provide a promising vehicle for controlled sharing of physical resources, allowing us to instantiate a precisely defined virtual resource, configured with desired software configuration and hardware properties, on a set of physical resources. We describe a model of virtual machine provisioning in a Grid environment that allows us to define such virtual resources and efficiently instantiate them on a physical Grid infrastructure. We argue that to properly account for and manage, the overhead resulting from instantiating and managing virtual resources, overhead must be scheduled at the same level as virtual resources, instead of being deducted from a user's resource allocation. We present preliminary results that demonstrate the benefits of such an approach.
2.2 Secure Service-Oriented Cloud Computing with Public Virtual Worker Nodes: M. Schmidt, N. Fallenbeck, M. Smith and B. Freisleben [2].

Corresponding Author: S. Archana, Department of CSE, Bharath University, India.

Cloud computing uses virtualization technologies to offer a non-shared use rental of computer resources with publicly accessible worker nodes. However, unlike grid computing, cloud computing as implemented by Amazon, IBM, Google and Microsoft only offers compute and storage resources from a single organization. Many of the cross-site and cross-organizational advantages offered by grid computing are lost. In this paper, we present a novel infrastructure that combines the benefits of grid and cloud computing: Cheap multi-organizational resources and private compute nodes with root access reachable from the Internet. Our previously introduced virtualization of grid resources is extended by an approach to offer the same freedom of network access cloud computing offers, but in a multi-organizational and shared use environment without endangering existing users or resources. An approach is presented for the dynamic network isolation of grid users from each other as well as a mechanism for shielding the grid infrastructure from malicious users and attacks from the Internet. This solution overcomes the traditional limitation that grid worker nodes are kept in private networks and enables new multi-site service-oriented applications to be deployed securely.

Above the Clouds:a Berkeley View of Cloud Computing: M. Armbrust, A. Fox, R. Griffith and A. Joseph [3]: Cloud Computing, the long-held dream of computing as a utility, has the potential to transform a large part of the IT industry, making software even more attractive as a service and shaping the way IT hardware is designed and purchased. Developers with innovative ideas for new Internet services no longer require the large capital outlays in hardware to deploy their service or the human expense to operate it. They need not be concerned about overprovisioning for a service whose popularity does not meet their predictions, thus wasting costly resources, or underprovisioning for one that becomes wildly popular, thus missing potential customers and revenue. Moreover, companies with large batch-oriented tasks can get results as quickly as their programs can scale, since using 1000 servers for one hour costs no more than using one server for 1000 hours. This elasticity of resources, without paying a premium for large scale, is unprecedented in the history of IT. Cloud Computing refers to both the applications delivered as services over the Internet and the hardware and systems software in the datacenters that provide those services. The services themselves have long been referred to as Software as a Service (SaaS). The datacenter hardware and software is what we will call a Cloud. When a Cloud is made available in a pay-as-you-go manner to the general public, we call it a Public Cloud; the service being sold is Utility Computing. We use the term Private Cloud to refer to internal datacenters of a business or other organization, not made available to the general public. Thus, Cloud Computing is the sum of SaaS and Utility Computing, but does not include Private Clouds. People can be users or providers of SaaS, or users or providers of Utility Computing [4].

We focus on SaaS Providers (Cloud Users) and Cloud Providers, which have received less attention than SaaS Users.

Providing and Managing Virtual Machine Execution Environments for Cloud Computing: I. Krsul, A. Ganguly, J. Zhang, J. A. B. Fortes and R. J. Figueiredo. Vmplants: As virtual appliances become more prevalent, we encounter the need to stop manually adapting them to their deployment context each time they are deployed. We examine appliance contextualization needs and present architecture for secure, consistent and dynamic contextualization, in particular for groups of appliances that must work together in a shared security context. This architecture allows for programmatic cluster creation and use, as well as mitigating potential errors and unnecessary charges during setup time. For portability across many deployment mechanisms, we introduce the concept of a implementation of the entire architecture using the Virtual Workspaces toolkit, showing real-life examples of dynamically contextualized Grid clusters.

CONCLUSION

The long dreamed vision of computing as a utility is finally emerging. The elasticity of a utility matches the need of businesses providing services directly to customers over the Internet, as workloads can grow (and shrink) far faster than 20 years ago. It used to take years to grow a business to several million customers-now it can happen in months. From the cloud provider's view [5], the construction of very large datacenters at low cost sites using commodity computing, storage and networking uncovered the possibility of selling those resources on a pay-as-you-go model below the costs of many medium-sized datacenters, while making a profit by statistically multiplexing among a large group of customers. In this paper we described a new technique, called contextualization, enabling the dynamic creation of functioning virtual constructs aware of their context. We discussed two existing implementations providing generic contextualization information, their respective assumptions and capabilities and gave examples of they can be used in conjunction with a context broker to deploy virtual clusters. Our purpose in this paper was to describe a general solution and a process that can be used with any deployer and any appliance provider that fulfill the specified conditions of secure transfer of information. Based on this process, we highlighted the need for standards on the deployers and appliance provider's side. Making contextualization an accepted technology will require the collaboration of many branches of technology. Besides the obvious ones of appliance configuration and deployment, better and more flexible methods of context information delivery to appliances [6] will need to be developed to allow for recontextualization. Further, applications will also need to develop the awareness of the potential of contextualization in order to leverage it.

In this paper, we have presented several methods for distributing virtual machine images to physical machines in a Cloud computing environment no matter if it is a private Cloud connected via a fast local area network or several remote Cloud computing sites connected via the Internet. We have studied their performance in case of encrypted and unencrypted data transfer: unicast distribution, binary tree distribution, peer-to peer distribution based on Bit Torrent and finally multicast. Our evaluation showed that multicast offers the best performance. Nevertheless, when transferring VM images between remote sites, Bit Torrent is the method to choose.

It is even possible to combine two methods to overcome limitations caused by the transfer over network boundaries. To avoid the retransmission of a VM image that is already present at a destination node and to make Cross-Cloud computing work, an approach based on copy-on-write layers has been presented. The evaluation of the layered file system showed that it saves a considerable amount of traffic, up to 90%.There are several areas for future work. For example, integration of strong encryption and decryption directly into the Bit Torrent client would significantly increase the performance and security of encrypted Bit Torrent image deployment. Furthermore, the scalability of the presented solutions should be investigated in a test environment with machines in a wide area network environment. Finally, studying efficient change detection methods for storage blocks to avoid retransmissions is an interesting area for further research.

System Analysis

Existing System: In the existing system, Unicast is the term used to describe communication where a piece of information is sent from one point to another point. In this case there is just one sender and one receiver.

Unicast transmission, in which a packet is sent from a single source to a specified destination, is still the predominant form of transmission on LANs and within the Internet. All LANs (e.g. Ethernet) and IP networks support the unicast transfer mode and most users are familiar with the standard unicast applications (e.g. http, SMTP, ftp and telnet) which employ the TCP transport protocol.

Proposed System: In the Proposed system, Multicast is the term used to describe communication where a piece of information is sent from one or more points to a set of other points. In this case there is may be one or more senders and the information is distributed to a set of receivers [7].

One example of an application which may use multicast is a video server sending out networked TV channels. Simultaneous delivery of high quality video to each of a large number of delivery platforms will exhaust the capability of even a high bandwidth network with a powerful video clip server. This poses a major salability issue for applications which required sustained high bandwidth. One way to significantly ease scaling to larger groups of clients is to employ multicast networking.

Multicasting is the networking technique of delivering the same packet simultaneously to a group of clients. IP multicast provides dynamic many-to-many connectivity between a set of senders (at least 1) and a group of receivers. The format of IP a multicast packet are identical to that of unicast packets and is distinguished only by the use of a special class of destination address (class D Ipv4 address) which denotes a specific multicast group. Since TCP supports only the unicast mode, multicast applications must use the UDP transport protocol.

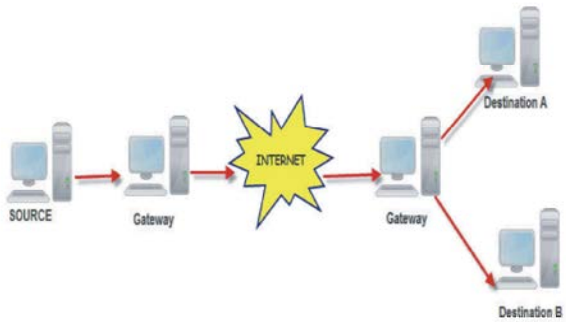
System Specification

Software Requirements:

- Operating system : - Windows XP Professional.
- Coding Language: - JAVA (JSP)
- Tool used: -Eclipse

Hardware Requirements:

• System:	Pentium IV 2.4 GHz.
• Hard Disk:	40 GB.
• Floppy Drive:	1.44 Mb.
• Monitor:	15 VGA Colour.
• Mouse:	Logitech.
• Ram:	256 Mb.



Architecture:
Module Description Modules:

- MODULE 1:
- Multicast Module
- MODULE 2:
- VM image
- MODULE 3:
- Virtual Private Network
- MODULE 4:
- Tree-based distribution

Multicast Module: The distribution of VM images via multicast is the most efficient method in local area network environment. The design of the multicast module can be kept simple if no ordering guarantees are given by the master node. Ensuring reliability in this case is delegated to both the sender and the receivers. It is possible to distribute the multicast packets only over selected links according to their multicast group membership. The network policy forbids that data is transferred via multicast over the network backbone connecting the two desktop Clouds. Choosing Bit- Torrent or the binary tree based distribution mechanisms to overcome this limitation is an interesting option.

VM image: The problem of distributing VM images to physical compute nodes may be negligible, but in a common Cloud computing setup with hundreds or thousands of nodes, the time needed to distribute virtual machines images to their destinations is crucial. The VM image distribution time, the network load caused by the distribution process itself is a critical factor. A simple and inconsiderate distribution method could lead to significant traffic on the core network components. The redistribution of already transferred VM images. If a user's task is started for the first time on a new infrastructure, his/her data VM image is distributed to all destination nodes. After the computation, the image remains on the local disk of the compute node, only the parts that change are stored on shared storage. If the user starts another task that is scheduled on the same compute nodes, it would be a waste of time and bandwidth to retransfer the VM image [8].

Virtual Private Network: The time needed to transfer a VM image to all compute nodes as compared to the unicast distribution method. The method can be used for Cross-Cloud computing, if either all compute nodes are located inside the same subnet (e.g., if two remote sites are connected with a Virtual Private Network (VPN) or if the compute nodes have public IP addresses. to distribute VM images between two dedicated VM image pool nodes on remote sites if the networks where the actual compute nodes reside are private or not connected in a VPN. The distribution from the pool node to the compute nodes.

Tree-based Distribution: The tree-based distribution method only scales to a certain degree, but it are ready to go and needs no preliminary work like generating and distributing torrent files. Thus, choosing the right method depends on the amount of data to distribute and the number of receiving nodes. Network congestion and to allow parallel transfers, a binary-tree based distribution method can used. In this method, all compute nodes are arranged in a balanced binary tree with the source node as its root. the tree needs to be rebalanced to guarantee seamless transfer. Besides rebalancing, there is some further action involved to resume the transfer to all nodes that are now rebalanced.

CONCLUSION

we have presented several methods for distributing virtual machine images to physical machines in a Cloud computing environment no matter if it is a private Cloud connected via a fast local area network or several remote Cloud computing sites connected via the Internet. We have studied their performance in case of encrypted and unencrypted data transfer: unicast distribution, binary tree distribution, peer-to-peer distribution based on Bit Torrent and finally multicast. Our evaluation showed that multicast offers the best performance. Nevertheless, when transferring VM images between remote sites, Bit Torrent is the method to choose. It is even possible to combine two methods to overcome limitations caused by the transfer over network boundaries. To avoid the retransmission of a VM image that is already present at a destination node and to make Cross-Cloud computing work, an approach based on copy-on-write layers has been presented [9].

The evaluation of the layered file system showed that it saves a considerable amount of traffic, up to 90%. There are several areas for future work. For example, integration of strong encryption and decryption directly into the Bit Torrent client would significantly increase the performance and security of encrypted Bit Torrent image deployment. Furthermore, the scalability of the presented solutions should be investigated in a test environment with machines in a wide area network environment. Finally, studying efficient change detection methods for storage blocks to avoid retransmissions is an interesting area for further research [10-14].

REFERENCES

1. Amazon Web Services LLC. Feature Guide: Amazon EC2 Elastic IP Addresses, 2008. http://developer.amazonwebservices.com/connect/entry.jspa?externalID=1346,
2. Andrews, T., F. Curbera, H. Dholakia, Y. Goland, J. Klein, F. Leymann, K. Liu, D. Roller, D. Smith, S. Thatte, I. Trickovic and S. Weerawarana, 2003. Business Process Execution Language for Web Services, Version 1.1. Microsoft, IBM, Siebel, BEA und SAP, 1.1 edition.
3. Armbrust, M., A. Fox, R. Griffith and A. Joseph, 2009. Above the Clouds: A Berkeley View of Cloud Computing. University of California.
4. Schmidt, M., N. Fallenbeck, M. Smith and B. Freisleben, 2009. Secure Service-Oriented Grid Computing with Public Virtual Worker Nodes. In Proceedings of 35th EUROMICRO CONFERENCE on Internet technologies, quality of service and applications (ITQSA), page (to appear). IEEE press.
5. Kerana Hanirex, D. and K.P. Kaliyamurthie, 2013. Multi-classification approach for detecting thyroid attacks, International Journal of Pharma and Bio Sciences, 4(3): B1246-B1251.
6. Khanaa, V., K. Mohanta and T. Saravanan, 2013. Comparative study of uwb communications over fiber using direct and external modulations, Indian Journal of Science and Technology, 6(6): 4845- 4847.
7. Kumar Giri, R. and M. Saikia, 2013. Multipath routing for admission control and load balancing in wireless mesh networks, International Review on Computers and Software, 8(3): 779-785.
8. Kumarave, A. and K. Rangarajan, 2013. Routing alogrithm over semi-regular tessellations, 2013 IEEE Conference on Information and Communication Technologies, ICT.
9. Kumarave, A. and K. Rangarajan, 2013. Algorithm for automaton specification for exploring dynamic labyrinths, Indian Journal of Science and Technology, 6(6).
10. Shafaq Sherazi and Habib Ahmad, 2014. Volatility of Stock Market and Capital Flow Middle-East Journal of Scientific Research, 19(5): 688-692.
11. Kishwar Sultana, Najm ul Hassan Khan and Khadija Shahid, 2013. Efficient Solvent Free Synthesis and X Ray Crystal Structure of Some Cyclic Moieties Containing N-Aryl Imide and Amide,Middle-East Journal of Scientific Research, 18(4): 438-443.
12. Pattanayak, Monalisa and P.L. Nayak, 2013. Green Synthesis of Gold Nanoparticles Using Elettaria cardamomum (ELAICHI) Aqueous Extract World Journal of Nano Science & Technology, 2(1): 01-05.
13. Chahataray, Rajashree. and P.L. Nayak, 2013. Synthesis and Characterization of Conducting Polymers Multi Walled Carbon Nanotube-Chitosan Composites Coupled with Poly (P-Aminophenol) World Journal of Nano Science & Technology, 2(1): 18-25.
14. Parida, Umesh Kumar, S.K. Biswal, P.L. Nayak and B.K. Bindhani, 2013. Gold Nano Particles for Biomedical Applications World Journal of Nano Science & Technology, 2(1): 47-57.