International Journal of Information Technology and Knowledge Management
December 2012, Volume 6, No. 1, pp. 33-38

COMPARATIVE STUDY AND PERFORMANCE ENHANCEMENT OF SCALED OUT CLOUD ARCHITECTURE OF AMAZON WEB SERVICES

Rupali Dalpatia[1] and Neeraj Mangla[2]

Abstract: The introduction of affordable infrastructure on demand, specifically Amazon's Elastic Compute Cloud (EC2), has had a significant impact in the business IT community and provides reasonable and attractive alternatives to locally-owned infrastructure. For scientific computation however, the viability of EC2 has come into question due to its use of virtualization and network shaping and the performance impacts of both. Several works have shown that EC2 cannot compete with a dedicated HPC cluster utilizing high-performance interconnects, but how does EC2 compare with smaller departmental and lab-sized commodity clusters that are often the primary computational resource for scientists? Because of the increasing volume of information distribution and data accumulation by means of broadband networks, the amount of available information is increasing in recent years. In this study, we aim to acquire insufficient resources dynamically from cloud computing systems while basic computation is performed on its own local clusters. Scalable resource management is achieved by monitoring resource usage of own local cluster and insufficient resources are acquired dynamically from cloud computing. We have confirmed that the execution time becomes shorter with the migration of virtual machines even the cost of migration is taken into account, compared with the case of accessing data over a network when an I/O-intensive application is executed. Moreover, we have measured basic performance on Amazon Elastic Compute Cloud (EC2) as basic data of a real cloud computing system.

Keywords: Elastic Compute Cloud, High Performance, Clusters.

1. INTRODUCTION

In this study, we aim to acquire insufficient resources dynamically from cloud computing systems while basic computation[3] is performed on its own local clusters. Because of the increasing volume of information distribution and data accumulation by means of broadband networks[2], the amount of available information is increasing in recent years.

Scalable resource management is achieved by monitoring resource usage of own local cluster and insufficient resources are acquired dynamically from cloud computing[1]. In terms of cluster systems, we have proposed a virtual machine PC cluster in which virtualization is applied to a PC cluster that uses a general-purpose personal computer for each node[4].

In addition, we have used IP-SAN for storage access which can realize long-distance connection at low cost over a high latency network[3].

In this paper, we have confirmed that the execution time becomes shorter with the migration of virtual machines even the cost of migration is taken into account, compared with the case of accessing data over a network when an I/O-intensive application is executed[5]. Moreover, we have measured basic performance on Amazon Elastic Compute Cloud (EC2) [1], throughput of network and iSCSI disk access, as basic data of a real cloud computing system. VMWare [2] and Virtual PC [3] are known to realize a virtual computing environment. They enable us to install guest OS on top of host OS, and therefore, they have a problem of processing performance degradation compared with the case of real machine because guest OS works as an application on host OS. On the other hand, Xen [4] provides a basic platform of virtualization called Virtual Machine Monitor, on which multiple OSes operate as shown in Figure 1. Since the overhead of virtual machine is reduced, Xen can achieve higher performance which is close to the case of real machine. Xen is used even in a business field recently, because Xen achieves remarkably high performance as opensource software[5].

In the architecture of Xen, Virtual Machine Monitor is foundation for virtualization[2] and virtual machines called domain are allocated on top of it. Domain0 behaves as a host OS and DomainU behaves as guest OS. Domain0 has a privilege to access physical hardware resources and to manage other domains[1].

[1] CSE, MM University, Mullana, Ambala (Haryana) *E-mail: sundeep.galaxy@gmail.com*
[2] CSE, MMEC, MM University Mullana, Ambala (Haryana), *E-mail: erneerajynr@gmail.com*

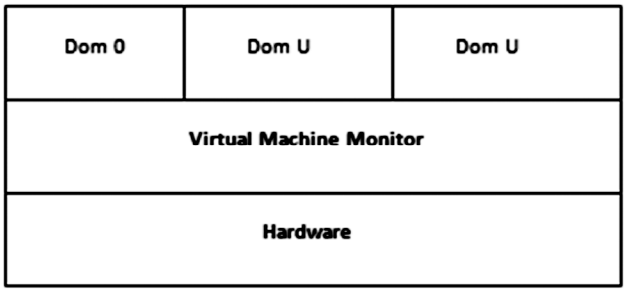

Figure 1: Architecture of XEN

Using virtualization software, a system can be constructed as multiple computers that operate virtually on a single computer [7]. In this study, we have used virtual machine in worker nodes of virtual machine PC cluster [8]. As we use virtual machine as worker nodes, flexible management of infrastructure can be introduced depending on system load and service demand [9]. In addition, the migration mechanism can be introduced that migrate virtual machine to another node while maintaining the state of running applications.

2. BACKGROUND TECHNOLOGIES

2.1 Cloud Computing

Cloud computing is an idea that users are allowed to use computing resources including server machines and storage across the Internet without [6] being conscious of its existence nor inner structure. The usage of remote data center and cloud computing should increase in the near future, because it is expected to reduce the introduction and management cost of information systems, which is better than the case that all computing resources must be purchased at the user site, like Hardware as a service (Haas)[8].

Amazon EC2 is one of typical cloud computing systems, which is a rental service of virtual machines on Internet. In the back-end of Amazon EC2 [2][4][5], server virtualization technology is used and users can back up whole image of own OS. Flexible system operation is possible because a new machine can start up in minutes when a system load increases suddenly.

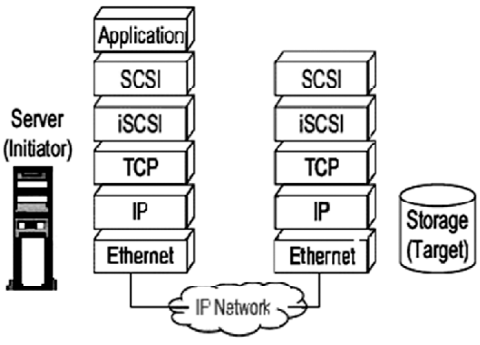

Figure 2: Configuration of iSCSI

Recently, since the amount of data processed in an information system has become huge, SAN is introduced to connect storage, and it has come to be widely used. SAN unifies the distributed storage held by each node, and realizes efficient practical use with central control for disk resources [6]. IP-SAN is expected as SAN of the next generation built with IP network. Internet Small Computer System Interface (iSCSI) [5] is the most popular protocol of IPSAN and we can build SAN with inexpensive Ethernet and TCP/IP using iSCSI [3]. In addition, iSCSI is expected for the realization of the long-distance remote access because IP network infrastructure is widely deployed and maintained in wide area networks. IP-SAN is expected to be used not only remote storage like data center but also outsourcing computer resources in a cloud computing framework. Figure 2 shows [5] the layered structure of iSCSI. iSCSI encapsulates a SCSI command within a TCP/IP packet and transmits the volume of data between server (Initiator) and storage (Target). In the future, since the gigabit/ 10 gigabit class line is expected to be more popular by development of the Internet, iSCSI will be effective furthermore. In this paper, we have used iSCSI protocol and constructed a virtual machine PC cluster assuming wide area network environment communication.

3. EXECUTION OF DATA-INTENSIVE APPLICATIONS

3.1 Evaluation of PC Cluster Consolidated with IP-SAN

PC cluster consolidated with IP-SAN is introduced in [6] which consolidate back-end SAN that connects between a node (server) and storage and front-end LAN that connects among nodes by using iSCSI. In the case of PC cluster consolidated with IP-SAN, both the back-end SAN and front-end LAN can be unified into a single commoditized network built with TCP/IP and Ethernet by using iSCSI, as shown in Figure 3. Therefore, the reduction [6] of network construction cost and the increase in efficiency of operational management can be expected. However, since both back-end and front-end networks use the same network resources, it is concerned the communication packets transmitted between nodes collide with packets of storage access on the same network, and performance deteriorates as a result. On PC cluster consolidated with IP-SAN, we have executed Hash Partitioned Apriori (HPA) [7] and Parallelized FP-growth (PFP) which are parallel association rule mining, and mpiBLAST [8] which is a sort of parallelized large-scale scientific computation in bioinformatics.

According to the result of above experiments, the feature of total performance of the system becomes CPU-bound or I/O-bound, not network-bound in these cases even though iSCSI network is consolidated.

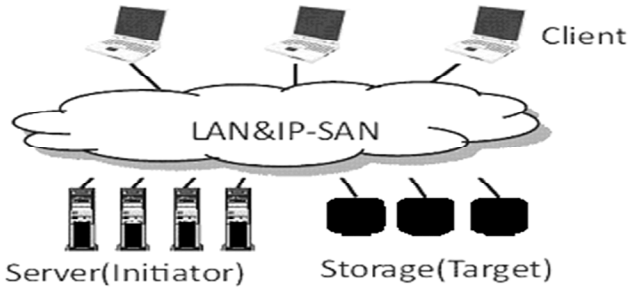

Figure 3: PC Cluster Consolidated with IP-SAN

3.2 Benchmark Execution with Remote Storage Access

In [9], we have analyzed the behavior of a virtual machine PC cluster executing an I/O-bound application over high latency networks.

Specification of each node of the cluster is shown in Table 1. A virtual machine (DomainU) has been created for each worker node. We have monitored a virtual machine PC cluster using a monitoring tool Ganglia [10] including iSCSI communication to access remote storage. We have built a virtual machine PC cluster like Figure 4 that contains four servers (Initiator) and two iSCSI storage nodes (Target). We have executed a database benchmark Open Source Development Labs Database Test3 (OSDLDBT3) [12] and compared the performance of two cases in which job is given to Domain0 and DomainU, respectively.

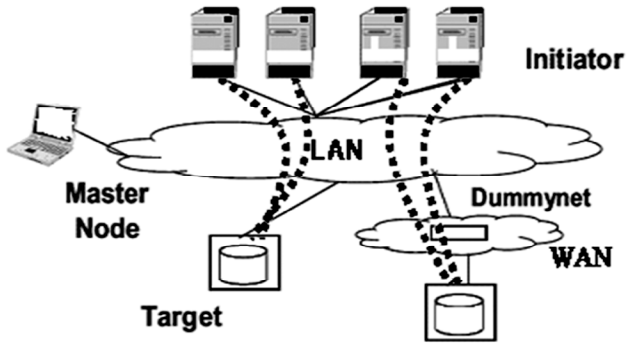

Figure 4: Experimental Environment

A virtual machine (DomainU) has been created for each worker nodes. For storage access we have used iSCSI. Two server access to storage in a local area and other two servers access to remote storage. To construct a remote access environment, we have inserted Dummynet [11] that simulates delay artificially between a local site and a remote site, and we have inserted delay as RTT of remote iSCSI access is 1msec, 2msec, 4msec, 8msec, 16msec.

The execution time of OSDL-DBT3 is shown in Figure 5. According to the graph, execution time is increasing as delay becomes longer from 1msec to 16msec in iSCSI access[1][5]. Especially, severe increase in execution time is observed when delay is longer than 4msec. On the contrary, Figure 6 shows execution time of HPA with the data size of 20 megabytes. HPA is a parallel application that parallelizes association rule mining which is based on Apriori algorithm using a hash function, and generates frequent itemsets from candidate itemsets repeatedly. Although HPA is data mining that processes huge amount of transaction data, since it includes heavy CPU processing,[11] it is not I/O-bound. Therefore the difference of execution time is small even in the case of longer delay. On the other hand, OSDL-DBT3 that is simplified from Transaction Processing Council Benchmark-H (TPC-H) [13] benchmark simulates a decision support system, and this is an I/O-bound application that inserts and deletes data to databases and queries are executed repeatedly. Remote storage access has caused I/O-bound in the application execution, and therefore significant execution time difference was observed as the delay becomes longer.

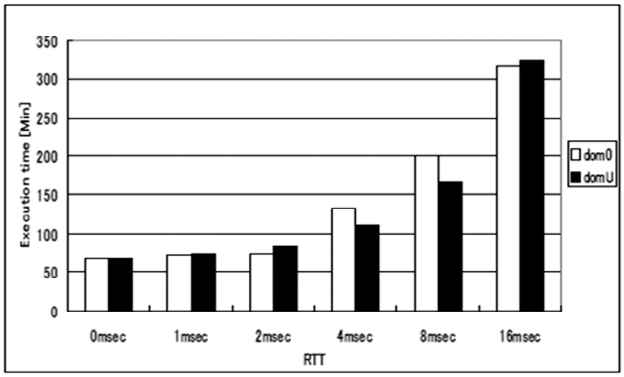

Figure 5: Execution Time of OSDL-DBT3 with Remote Storage Access

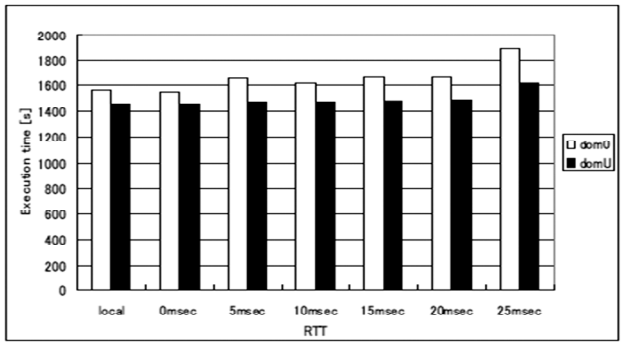

Figure 6: Execution Time of HPA with Remote Storage Access

4. EVALUATION OF CLUSTER WITH SERVER MIGRATION

4.1 Basic Experiment

First, we have compared performance of a real machine PC cluster and a virtual machine PC cluster with a basic

experiment. We have built the cluster connecting four servers (Initiator) and a storage node (Target) with iSCSI, and executed OSDL-DBT3 on each environment (Figure 7) [5]. Worker nodes on the real machine cluster have 4 gigabytes memory, and we have devided it to two 2 gigabytes and assigned them for Domain0 and DomainU respectively in the case of the virtual machine cluster. The execution time is shown in Figure 8. From this result, performance of DomainU (virtual machine) degrades a little compared with that of the real machine cluster. However, taking flexible management of infrastructure into account, using a virtual machine is considered to have greater benefits. In this graph, performance of DomainU is better than that of Domain0. This is considered to be the result of optimization of DomainU in the mechanism of Xen [14] [15].

Table 1
Experimental Setup: PCS

OS	initiator : Linux 2.6.18-53.1.14.el5(CentOS5.3)
CPU	initiator : Intel (R) Xeon(TM) 3.6GHz target : Intel (R) Xeon(TM) 3.6GHz
Main Memory	initiator(Domain0) : 2GB
iSCSI	initiator : iscsi-initiator-utils
Monitoring Tool	Ganglia

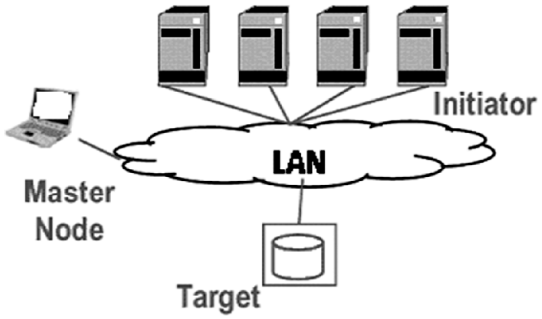

Figure 7: Basic Experiment: Experimental Environment

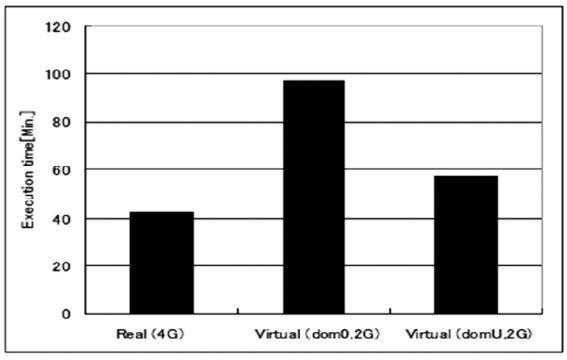

Figure 8: Comparison of Real Machine Cluster and Virtual Machine Cluster

4.2 Experiment Including Virtual Machine Migration

The results of previous research have shown that parallel data processing applications that is not I/O-bound like HPA[11] can be executed with sufficiently practical performance even when the data is stored at a remote site. In contrast, in the case of I/O intensive applications like OSDL-DBT3, we have confirmed remarkable performance decline when data exists on the site through high-latency network. Thus in this study, we propose a technique to migrate virtual machine to a cloud environment over high latency networks site that stores data, in order to achieve load balancing and optimization of storage access instead of iSCSI remote access[13]. We have built a virtual machine PC cluster as shown in Figure 9 and 10 that contains six servers (Initiator) and two iSCSI storage nodes (Target). Four servers and a storage node are located in the local site, two servers and a storage node are located in the remote site. Assuming remote iSCSI access, we have inserted 1msec, 2msec, 4msec, 8msec, 16msec RTT by Dummynet between the local site and the remote site as the same with existing research. First, Figure 11 shows migration time of a virtual machine from a local site to a remote site on each RTT. The result is 21 seconds when RTT is from 0msec to 4msec, and 52 seconds when RTT is 16msec that is the longest RTT measured in this study. Figure 12 shows total amount of time that sums up migration time to migrate a virtual machine to a remote site and execution time of OSDL-DBT3 at the remote site. Figure 12 includes

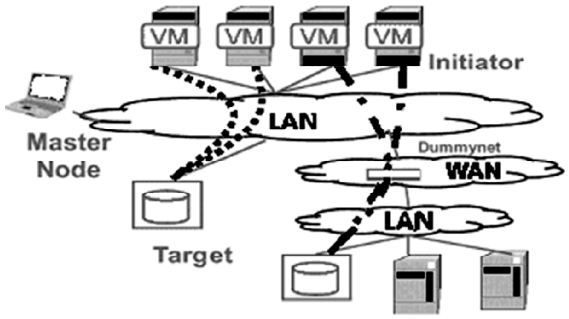

Figure 9: Experimental Environment of Before Migration

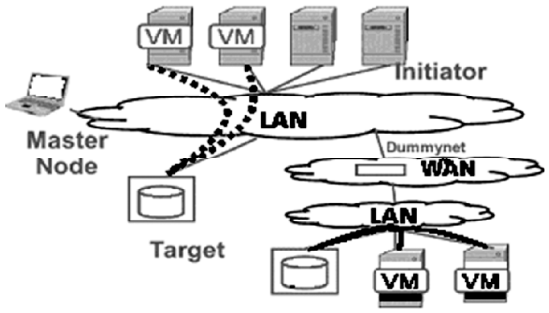

Figure 10: Experimental Environment of After Migration

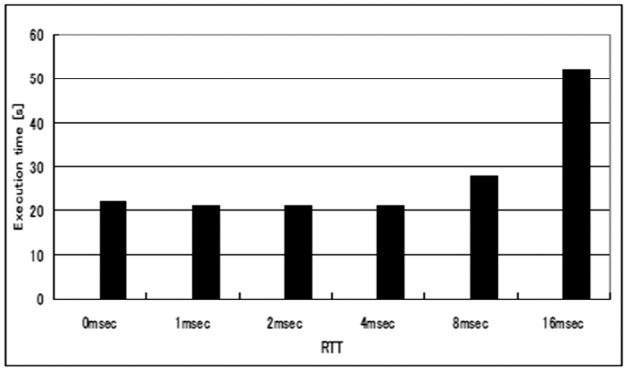

Figure 11: Migration Time

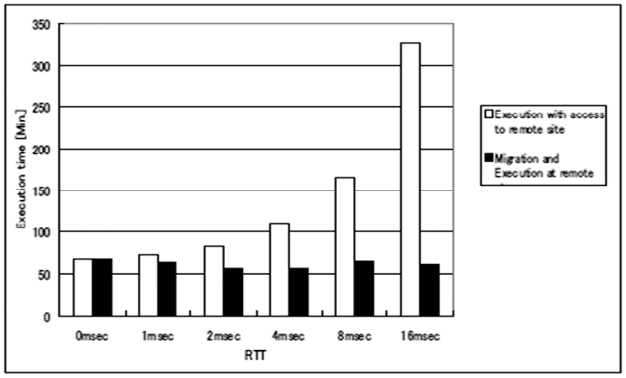

Figure 12: Execution Time

DomainU execution time of existing research, in which data at the remote site is accessed directory through WAN, for comparison. From this figure, the execution of this experiment is faster than that of existing research when RTT is long. We can execute an application without delay on iSCSI by means of migration of virtual machines to a remote site. Therefore, as RTT becomes longer, it is effective to migrate a server to a remote site where storage is located.

5. BASIC PERFORMANCE OF AMAZON EC2

5.1 An Overview of Experiments

We aim to realize the load balancing between local own cluster and remote real cloud dynamically, estimating application execution performance by means of network throughput and performance of disk access in data intensive application. Performance evaluation of each element should be essential for performance prediction of data intensive application.

First we have measured throughput of network and iSCSI disk access of Amazon EC2 because we are using iSCSI for storage access. We have used Open-iSCSI [16] for iSCSI Initiator and scsi-target-utils for iSCSI Target.

For the installation of Open-iSCSI, it is common to compile a kernel of OS at first and then compile OpeniSCSI in order to make a kernel module of Open-iSCSI. However, in Amazon EC2, users can not boot up a kernel that users have compiled by themselves. Thus we have prepared Amazon Machine Image (AMI) which includes gcc version 4.0, which is suitable for compiling both the kernel and Open-iSCSI, and created a kernel image in which kernel modules can be installed. Using the kernel image, we have deployed AMI and installed iSCSI.

5.2 Network Bandwidth Inside Cloud

Figure 13 shows network throughput between virtual machines on Amazon EC2. It includes throughput between virtual machines on the local site cluster for comparison.

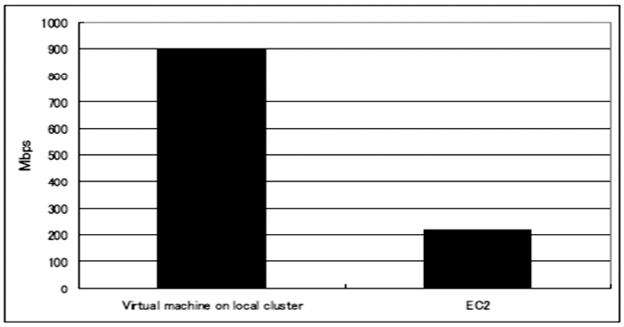

Figure 13: EC2 - Network Bandwidth

In network throughput, there is a difference in performance about 4.5 times in local cluster and Amazon EC2. We have examined network throughput between numbers of virtual machines on Amazon EC2, and observed that the performance of all cases is almost the same. This indicates that all the virtual machine should be deployed on the same real (physical) machine in this case.

5.3 Performance of Disk Access

Next, we have measured performance of disk access on Amazon EC2. Figure 14 shows performance of Write and Figure 15 shows Read.

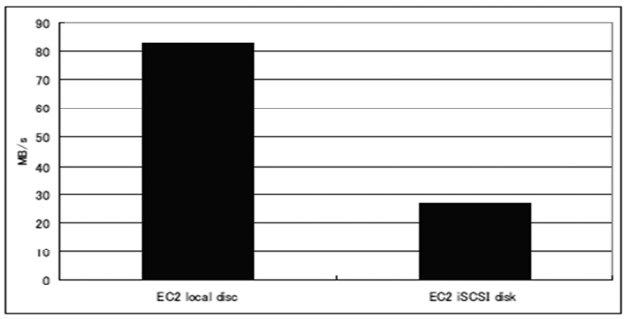

Figure 14: EC2 - Performance of Disk Access (Write)

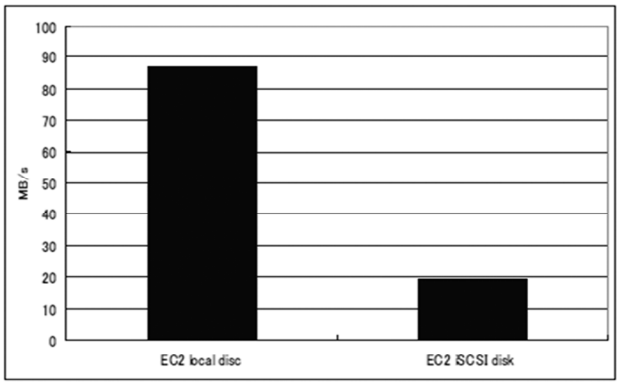

Figure 15: EC2 - Performance of Disk Access (Read)

These graphs include performance of local disk access and iSCSI disk access on EC2. Performance of iSCSI disk access is significantly lower than that of local disk access. This must come from the low throughput of network on EC2, which is shown in Figure 13. According to this result, we should take network performance inside cloud into account, including iSCSI disk access, which we outsource a job to a cloud computing system.

6. CONCLUSION

In existing research, we have experimented server access to remote storage directly during runtime. We have confirmed as RTT is longer the execution time becomes longer in iSCSI remote access. Since remote access is bottleneck of execution of the application, we have migrated a virtual machine on a local site to a remote site, and OSDL-DBT3 is executed with storage access in the remote site. As a result, total amount of migration and execution time is shorter than that obtained in existing research as RTT is longer. We have confirmed that our method is effective to relocate a local server to a remote site when it takes longer to access to a remote storage. Moreover, we have introduced iSCSI on Amazon EC2, measured throughput of network and iSCSI disk access for estimating execution performance of applications on real cloud.

As a future work, we will build a system of load balancing dynamically if the load is heavy which migrates virtual machine to a remote site automatically, and analyze its behavior. In addition, we will perform experiments that make cooperation between local cluster and real cloud.

REFERENCES

[1] O'Riordan, Dan. Cloud Computing Breaking Through the Haze. s.l.", *IBM Corporation 2009.*

[2] Lijun Mei,W.K. Chan,T.H. Tse, "A Tale of Clouds: Paradigm Comparisons and Some Thoughts on Research Issues", *2008 IEEE Asia-Pacific Services Computing Conference.*

[3] Hong Cai, Ning Wang, Ming Jun Zhou, "A Transparent Approach of Enabling SaaS Multi-tenancy in the Cloud", *2010 IEEE 6th World Congress on Services.*

[4] Chang Jie Guo, Wei Sun, Ying Huang, Zhi Hu Wang, Bo Gao", "A Frameworkfor Native Multi-Tenancy ApplicationDevelopment and Management", *The 9th IEEE International Conference on E-Commerce Technology and The 4th IEEE International Conference on Enterprise Computing, E-Commerce and E-Services* (CEC-EEE 2007) 0-7695-2913-5/07 $25.00 © 2007.

[5] Zeeshan Pervez, Sungyoung Lee, Young-Koo Lee, "Multi-Tenant, Secure, Load Disseminated SaaS Architecture", *Department of Computer Engineering, Kyung Hee University, South Korea.*

[6] Jack Brass, "Physical Layer Network Isolation in Multi-tenant Clouds", *International Conference on Distributed Computing Systems Workshops 2010.*

[7] Pankaj Goyal, "Computing Policy-based Event-driven Services-oriented Architecture for Cloud Services Operation & Management", *Presented at 2009 IEEE International Conference on Cloud.*

[8] Guoling Liu. "Research on Independent SaaS Platform", *School of Information Science and Technology Shandong Institute of Light Industry Jinan*, China.

[9] Shiori Toyoshima, Saneyasu Yamaguchi, Masato Oguchi: "Analyzing Performance of Storage Access Optimization with Virtual Machine Migration", *CPSY*, **109(296)**, CPSY2009-37,pp. 13-18, Kyoto, Nobember 2009.

[10] Ganglia Monitoring System:http://www.ganglia.info/

[11] Luigi Rizzo:"Dummynt"http://info.iet.unipi.it/ luigi/ dummynet/

[12] OSDL-DBT3:http://ldn.linuxfoundation.org/

[13] TPC-H:http://www.tpc.org/tpch/

[14] Aravind Menon, Alan L. Cox, Willy Zwaenepoel: "Optimizing Network Virtualization in Xen", *2006 USENIX Annual Technical Conference*, pp. 15-28, May 2006.

[15] Jose Renato Santos, Yoshio Turner, G. (John) Janakiraman, Ian Pratt: "Bridging the Gap between Software and Hardware Techniques. for I/O Virtualization", *2008 USENIX Annual Technical Conference*, pp. 29-42, June 2008.

[16] Open-iSCSI:http://www.open-iscsi.org/

[17] Tanimura Yusuke, Ogawa Hirotaka, Nakada Hidemoto, Tanaka Yoshio, Sekiguchi Satoshi: "Comparison of Methods for Providing An IP Storage to A Virtual Cluster System", *IPSJ SIG Notes 2007*, pp. 109-114, March 2007.