Volume 6, Issue 3, July 2019 ISSN: 2348-8565 (Online)

International Journal of Modern Engineering and Research Technology

Website: http://www.ijmert.org Email: editor.ijmert@gmail.com

VM Load Balancing Technique for Cloud

Hrithik Sanyal
Research Scholar
Department of Electronics & Telecommunication,
Bharati Vidyapeeth College of Engineering,
Pune (M.S.) India

Rajneesh Agrawal
\Mentor,
Comp-Tel Consultancy,
Jabalpur (M.P.) India
Email: rajneeshag@gmail.com

ABSTRACT

Cloud computing has spread over the Internet very deep and overriding technique among all parallel technologies in the current era. Major reasons for the same are available and flexible services being received from it. It is all due to the use of virtual machines for managing the user-specific services for cloud operations. Although SaaS, PaaS, IaaS have provided a great service model for the cloud, there are many challenges too. The most important challenge occurs due to availability of services using virtual machines. For virtual machines to work smoothly a lot of migration processes occurs due to variable loads which may increase or decrease over the cloud. When load decreases it all goes smoother but when load increases on a particular virtual machine or more than one virtual machines then a load balancer should be applied which will carry the responsibility of choosing a specific virtual machine and migrate it on a free server. It means there are three major challenges i.e. proactively identifying the load enhancement, selecting a specific virtual machine and then migration of the same to another server. In this paper, these problems are being discussed thoroughly and a better solution is being provided for the same. The work shall provide the high performance and accuracy of the load balancing of the virtual machines over the cloud.

Keywords:— *Virtual Machines, Cloud Computing, Fault Tolerant, Load Balancing, CloudSim.*

I. INTRODUCTION

Cloud computing is turning into a popular expression in the PC industry and everybody is hoping to relate in one manner or other with this shiny new idea. Cloud Computing is a current point and the term has increased a ton of footing being worn on notices everywhere on the Internet from web space facilitating suppliers, through server farms to virtualization programming suppliers. Slicing through the publicity of Cloud Computing is certifiably not a simple assignment as a basic web search gets the job done to persuade that there are close to the same number of definitions on what comprises 'Cloud Computing' as there are major parts in the market looking to increase a new area in that promising new business field.

A Cloud is a sort of equal and circulated framework comprising of an assortment of interconnected and virtualized PCs that are powerfully provisioned and introduced as at least one brought together processing assets dependent on assistance level arrangements set up through exchange between the specialist organization and customers.

In the Section I, the introduction is detailed. Section II discusses on Machine Learning; Section III elaborates more on the Existing Systems. In Section IV, Problem Statement is deliberated. In Section V, the Proposed Algorithms are illustrated. Section VI briefs about the Results & Discussions, Section VII shows the Comparisons with Existing Systems. Section VIII explains the Conclusion & Section IX illustrates the Future Scope.

II. CLOUD COMPUTING

Cloud computing alludes to both the applications conveyed as administrations over the Internet and the equipment and frameworks programming in the server farms that offer those types of assistance (Software as a Service - SaaS).

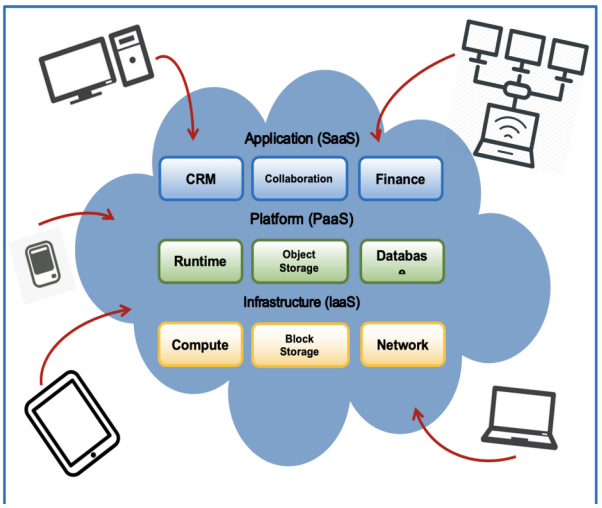

Figure 1: Service Model of Cloud Computing

IT organizations are offering types of assistance to the overall population for a charge on-request. This kind of administration is called Public Cloud. Then again if the administration is exclusively utilized inside an association and not imparted to individuals outside of the association it is called Private Cloud. There is additionally a third kind, a mix of public and private cloud. It is alluded to as Hybrid [8].

Model: eBay. Picking which one to convey simply relies upon the requirements.

Foundation as a Service (IaaS) — IaaS suppliers are organizations that give the most fundamental IT needs – workers, systems administration, and capacity – on a utilization-based instalment model. They normally make hefty interests in server farms and other foundation, and afterwards lease it out, permitting customers to keep away from speculations of their own.

As indicated by the creator [10] Current Cloud Computing contributions with virtual machines depend on correspondence utilizing Ethernet. InfiniBand as an interconnect innovation would be a superior decision because of its boss presentation. They present a novel design for HPC IaaS mists supporting InfiniBand that permits an individual organization arrangement with QoS systems. Clients can designate virtual groups which are segregated from different inhabitants in the actual organization.

Stage as a Service (PaaS) — It alludes to giving stage layer assets, including working framework backing and programming advancement structures. Instances of PaaS suppliers incorporate Google App Engine, Microsoft Windows Azure and Force.com.

Programming as a Service (SaaS) — The term SaaS dates from the 1990s and in this way originates before Cloud Computing. SaaS is additionally referred to regularly as "Web administrations." While numerous marginally various meanings of SaaS are conceivable, a straightforward and usable definition has just been figured:

Advantages of SaaS — Compared with customary registering and programming appropriation arrangements, SaaS mists give versatility and move huge weights from supporters of suppliers, bringing about various open doors for more prominent proficiency and, sometimes, execution.

The following sections describe five key benefits of SaaS clouds.

- Very Modest Software Tool Footprint
- Efficient Use of Software Licenses
- Centralized Management and Data
- Platform Responsibilities Managed by Providers
- Savings in Up-front Costs

In a general sense, Cloud Computing gives advantageous rental of registering assets. These assets, which are normally gotten to by endorsers over an organization, must be quantifiable in units that can be exclusively designated to explicit supporters, and paid for dependent on elements, for example, how long the units are held, who approaches them, how they are utilized, and so on account of SaaS, what is being leased is admittance to an application [6].

To ensure application information traded between the supporter's program and the cloud supplier over the organization, cryptography is required. Normally, the supporter's program and the cloud supplier's worker start a meeting by first arranging a mutual key utilizing one of a few standard key trade conventions (e.g., TLS [7] or SSL [8]). The HPC frameworks are right now positioned on the TOP500 list [10]. This rundown position the most impressive supercomputers from around the globe. Superior LINPAC of the LINKPACK [9] benchmark assesses and measure the outstanding tasks at hand. LINPACK is a product that gives answers for complex straight conditions on PCs, supercomputers being one of them.

Dormancy is the slack in correspondence between the frameworks. In arranged frameworks, this inertness can add up rapidly and hence affecting the general execution. It tends to be estimated by a straightforward ping test to check how lengthy timespan it takes for a message to send and get. Equipment necessities incorporate superior CPUs with multi-stringing [6].

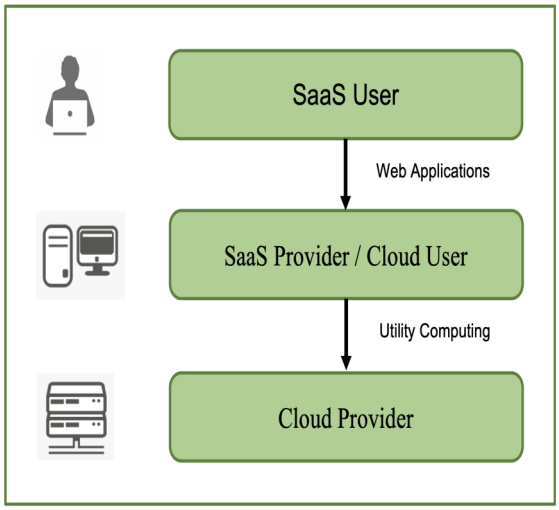

Figure 2 : Users and Providers of Cloud Computing

As an ever-increasing number of utilizations and substance is being facilitated and upheld in the Cloud there is a consistently developing requirement for supporting superior applications as well. So scientists are moving their HPC applications to the cloud to comprehend the bottlenecks and difficulties it presents. Models are planned and created to anticipate the disappointments so they can be fixed before they can adversely affect the exhibition. One such trial was performed utilizing Microsoft's Azure cloud with Numerical Generation of Synthetic Seismograms [7].

III. Existing Systems

When we talk about cloud computing, fault tolerance becomes a crucial problem. This problem is much difficult compared to other problems because cloud computing relies on the complex splitting of layers. Most cloud computing platforms exploit virtualization meaning they split into 3 layers namely virtual machines, hosts, and applications. With the help of this paper, analysis is done

on the implementation of loud computing exploiting the fault tolerance and focusing mainly on autonomous repair. It is seen that apart from this work, much approaches handle the fault tolerance by the user, provider and customers, which are not sufficiently suffice [1].

Through this paper, we propose a hybrid scheduling algorithm for balancing load amongst a distributed environment. This is done by the combination of Divide-and-Conquer and Throttled algorithms methodologies commonly referred to as DCBT. The algorithm we propose plays an important part by efficiently distributing the load (incoming) for maximizing resource utilization while in a cloud environment. Also, the main purpose of this proposed DCBT is to minimize the total execution time and maximize the utilization of resources. Furthermore, the proposed DBT algorithm is analyzed using CloudSim simulator and also in a customized distributed environment using python. Through our experimentation, we could extrapolate, that the proposed algorithm provides better and efficient results. This work minimized execution time allocated to Request Handlers (RH) by 9.972% in comparison to the Modified Throttled algorithm [2].

The main purpose of cloud computing is to convey efficient and well-managed computing resources and utilizations. It is possible through a concept called Virtualization. In an ideal model of cloud computing service, users are offered resources to their machines including raw storage, load balancers, memory locations etc. In this paper, the investigation of the utilization of various memories is done and then evaluated. Moreover, a summary of a new concept is mentioned which is also proposed, which finds the impact effects of schedule which will reduce memory utilization in the cloud in the times of transfer of high data transfer and process execution [3].

The increased popularity of the Internet and mobile devices and led to state-of-art technology in cloud-based e-Learning. At its initial stage, an academic cloud-based e-Learning behaves like a private cloud with traces of Service Level Agreements (SLA) and limited resources. These constraints the end user as he/she can only have access to limited resources. This can be reduced by proper balancing of workload. Through this paper, mapping requested Virtual Machine (VM) of each user is configured with is host configuration with exact caution. It is done using the Three-Phase Filtering Protocol (3PFP). Also, an enhanced time-efficient algorithm is developed for distributing workload evenly amongst the machines which are the host. It is vaguely based on LVMEDT (Last Virtual Machine Executed Date and Time) [4].

Cloud computing has a broad range of cloud -based models, including SaaS: Software as a Service, PaaS: Platform as a Service IaaS: Infrastructure as a Service XaaS: Everything as a Service NaaS: Network as a Service and RaaS: Recovery as a Service. Whatever the user selects, the estimation of all the above is the same. They all encourage fast development of capacities on the VMs in the cloud and also a maximum percentage of IT modest of advanced cloud computing.

This enables a necessity for the need of advanced VM load balancer for balancing VM load level. Through this paper, we propose a various dynamic load balancing methods and enhanced algorithms for analyzing and thus developing a load balancer which will thus enable to allocate workloads to the VMs, evenly and also help in migration the excess burden from the encumbered VMs to the non-congested VMs. To do this we propose an enhanced

Throttle Load Balancer which will indeed allocate the workloads evenly to all the VMs thus minimizing the delay time and response time. It will also distribute the requests and allocate dynamic resources. Thus, the proposed work of Enhanced Throttled Load Balancer will help to minimize the delay and response time [5].

III. Problem Statement

Dynamic Scalability-As the number of client demands increment the application must have the option to help the expanding load. Simultaneously as the number of solicitations declines the application should have the option to downsize. So, accomplishing dynamic versatility is a test for HPC applications in the cloud. Windows Azure arbitrarily de-dispenses the figure hubs when downsizing and consequently follows an offbeat cycle. This adversely impacts the presentation.

The two IaaS and PaaS offer types of assistance might be utilized to assemble and send adaptable applications that can be upgraded for equal registering. If there should be an occurrence of IaaS the foundation is as of now constructed and is promptly accessible for offering types of assistance on-request. A few bottlenecks, for example, delays, support, working expenses and so forth should be eliminated.

Adaptation to non-critical failure Virtual hubs is spurred on interest to deal with the heap and to play out the registering undertakings. So, giving Virtual assets is another test for HPC in the cloud. On the off chance that a virtual hub falls flat while playing out an undertaking, it gets basic to recognize wherein the framework it fizzled and why? All together for the presentation to be not affected it is significant that the heap is moved to another hub while it is distinguished and fixed. It is likewise alluded to as adaptation to non-critical failure. This is one of the difficulties while planning load adjusting frameworks for superior applications. It is conceivable that a registered hub may share its assets in running more than one application. As the number of uses increments, the heap on the figure hub it can diminish the presentation and may at some point come up short whenever came to over limit.

The idea of virtualization assumes a critical part of the usage of the HPC2 model. Other than the development of virtual on interest groups there are further advantages that could be normal, for example, live movement.

IV. Proposed Algorithms

Cloud Computing is encouraging clients around the globe for the best of the administrations accessible over the world on their machines through the web. It is gainful for both the specialist co-ops (they get gigantic customer base) and customers (they get every single accessible help).

Steps of Proposed Work

- Using CloudSim Simulator for implementation
- Create a Host and a Default Virtual Machine on the host
- Create a Data Center to execute the various processes
- Create a process generator which will generate process with Random amount of RAM and fixed CPU usage. Each process will run for a random amount of time assigned initially
- Create a Broker service which will submit the process to Data Center
- Broker Service will first check if the available virtual machines can handle the new process otherwise it will create a new virtual machine in the

data centre and will assign the new process to it.

- This step will avoid any possible fault which may occur in the data centre causing system instability or affecting other processes.
- The processes will be generated with very high speed to map the HPC over the proposed cloud.
- Various measures such as the number of processes, number of virtual machines etc. shall be counted on a time scale to show the working of the proposed system.
- The system shall be evaluated at the end:
- Using the current System
- In this, the following modules shall be used during implementation and testing:
- Datacenter
- Datacenter Broker
- Will Decide whether a new VM is required or not
- Will submit the VM & Processes to the Data Center
- Will Generate Results
- Process Generator
- Virtual Machine Generator
- The whole system shall be executed twice:
- Once without Making Decisions and generating VM for the execution of the generated processes (It will be causing the various faults and number of processes shall be failed)
- Second by running the Decision-Making Module and performing fault -tolerance on that basis so that outputs can be compared with results of the first execution.

V. PROPOSED ALGORITHMS

Cloud Computing is encouraging clients around the globe for the best of the administrations accessible over the world on their machines through the web. It is gainful for both the specialist co-ops (they get gigantic customer base) and customers (they get every single accessible help).

Steps of Proposed Work

- Using CloudSim Simulator for implementation
- Create a Host and a Default Virtual Machine on the host
- Create a Data Center to execute the various processes
- Create a process generator which will generate process with Random amount of RAM and fixed CPU usage. Each process will run for a random amount of time assigned initially
- Create a Broker service which will submit the process to Data Center
- Broker Service will first check if the available virtual machines can handle the new process otherwise it will create a new virtual machine in the data centre and will assign the new process to it.
- This step will avoid any possible fault which may occur in the data centre causing system instability or affecting other processes.
- The processes will be generated with very high speed to map the HPC over the proposed cloud.
- Various measures such as the number of processes, number of virtual machines etc. shall be counted on a time scale to show the working of the proposed system.
- The system shall be evaluated at the

end:

- Using the current System
- In this, the following modules shall be used during implementation and testing:
- Datacenter
- Datacenter Broker
- Will Decide whether a new VM is required or not
- Will submit the VM & Processes to the Data Center
- Will Generate Results
- Process Generator
- Virtual Machine Generator
- The whole system shall be executed twice:
- Once without Making Decisions and generating VM for the execution of the generated processes (It will be causing the various faults and number of processes shall be failed)
- Second by running the Decision-Making Module and performing fault -tolerance on that basis so that outputs can be compared with results of the first execution.

V. RESULTS & DISCUSSION

Inference: From the graph shown below, it can be seen that the processing time-taken for VMs is actually the average time and thus not harm the system's performance. This is because of a mechanism also called as the Pre Copy proactive fault tolerance. This also states that the time taken for the evaluation of Pre Copy Algorithm doesn't necessarily affect the cloud and it's working as well. Through various readings (execution readings), we conclude that the behaviour stays consistent for all executions.

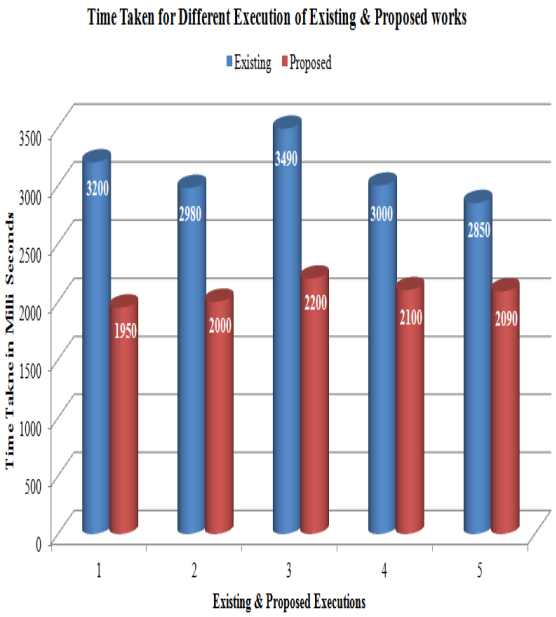

Figure 3: Readings taken from five different executions of the existing and proposed works

Inference: From the graph shown below, it is confirmed that the time required for VM processing in the proposed system is better than the existing systems. Also, from the readings and graphs as well, we can see the time-taken factor is quite less when compared to time-taken for existing systems for the application of pre-copy algorithms, indicating that the execution performance is enhanced using the application of pre-copy algorithms.

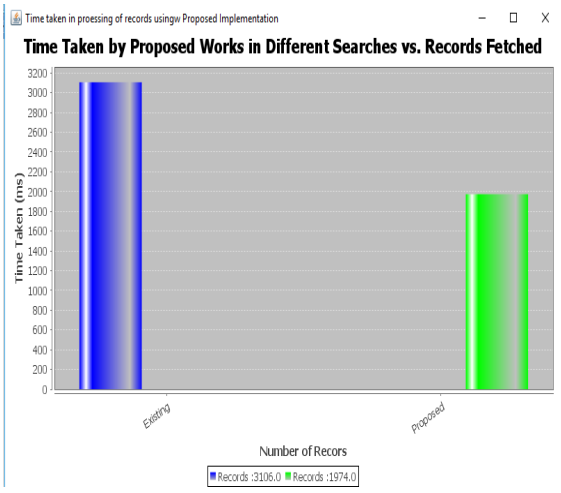

Figure 4: Readings taken from execution of the existing and proposed works

Table 1: Cloud lets and the time taken by them during the simulation

SNO	TIME TAKEN IN PROCESSING IN MILLIS SECONDS	
	EXISTING	PROPOSED
1	3200	1950
2	2980	2000
3	3490	2200
4	3000	2100
5	2850	2090

VII. Comparison With Existing Work

Table 2: Cloud lets and the time taken by them during the simulation

SNO	FEATURE	BASE-PAPER WORK	PROPOSED WORK
1.	Mechanism	reactive fault tolerance Algorithm	Proactive Fault Tolerance
2.	Algorithm	Task Replication Technique	VM Migration Technique using Best Fit Algorithm
3.	Algorithm Complexity	O(e*p) where e = Dset is the number of Task Dependencies and p is number of resources	O (n * log n) where n is number of virtual machines. Log n is the time required for searching best fit cloudlet on any virtual machines and n is number of virtual machines
4.	Resource Requirement	High as the replication will occupy double space	Low as best fit requires to migrate means only one occurrence remains in memory
5.	Results Obtained	94% success rate as the system is reactive	100% Success Rate as the system is proactive

Table 3: Cloud lets and the time taken by them during the simulation

TIME TAKEN IN PROCESSING IN MILLIS SECONDS	
EXISTING	PROPOSED
3106	1974

VIII. Conclusions

This work is providing a proactive load adjustment over the cloud for dealing with the migration of virtual machines so that it will reduce the misuse of resources without affecting the clients' operations over Cloud Computing. Proactive management of the migration over the cloud will ensure that the downtime of the VM over the cloud is mostly up and will keep away clients from unwanted delays due to migration of the VMs between servers. The proposed system has been implemented and found to be providing high performance. In comparison with the existing works it is providing better performance and proactive mechanism makes it more reliable and usable system.

The proposed work has been implemented using standard machines, in future, the equivalent can be implemented on real cloud and test it for its precision and performance.

A further improvement in the proposed calculation might be needed for real processing conditions over the cloud. At IaaS and PaaS layers changes might be useful in utilizing the proposed framework

Proactive administration of the cycles over cloud will expand the presentation of the cloud and will keep away from undesirable postponements brought about by the cycles running over the framework.

IX. FUTURE SCOPE

The proposed work is being executed on re-enactment climate utilizing standard machines, in future the equivalent can be conveyed over the genuine cloud climate and test it for its precision and execution.

A further improvement in the proposed calculation might be needed for genuine elite processing conditions over the cloud. At IaaS and PaaS layers changes might be useful in expanding the presentation of the proposed framework.

REFERENCES:

[1] A. Tchana, L. Broto and D. Hagimont, "Approaches to cloud computing fault tolerance," 2012 International Conference on Computer, Information and Telecommunication Systems (CITS), Amman, 2012, pp. 1-6, doi: 10.1109/CITS.2012.6220386.

[2] S. G. Domanal and G. R. M. Reddy, "Load Balancing in Cloud Environment Using a Novel Hybrid Scheduling Algorithm," 2015 IEEE International Conference on Cloud Computing in Emerging Markets (CCEM), Bangalore, 2015, pp. 37-42, doi: 10.1109/CCEM.2015.31.

[3] P. Lakkadwala and P. Kanungo, "Memory Utilization Techniques for Cloud Resource Management in Cloud Computing Environment: A Survey," 2018 4th International Conference on Computing Communication and Automation (ICCCA), Greater Noida, India, 2018, pp. 1-5, doi: 10.1109/CCAA.2018.8777457.

[4] C. Madhumathi and G. Ganapathy, "An effective time based load balancer for an academic cloud environment," 2015 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, 2015, pp. 1-6, doi: 10.1109/ICCCI.2015.7218106.

[5] G. J. Mirobi and L. Arockiam, "Dynamic Load Balancing Approach for Minimizing the Response Time Using An Enhanced Throttled Load Balancer in Cloud Computing," 2019 International Conference on Smart Systems and Inventive Technology (ICSSIT), Tirunelveli, India, 2019, pp. 570-575, doi: 10.1109/ICSSIT46314.2019.8987845.

[6] Viktor Mauch, Marcel Kunze and Marius Hillenbrand "High performance cloud computing", Future Generation Computer Systems, Elsevier, 2012.

[7] C. Baun, M. Kunze, T. Kurze, V. Mauch, High performance computing as a service, in: I. Foster, W. Gentzsch, L. Grandinetti, G.R. Joubert (Eds.), High Performance Computing: From Grids and Clouds to Exascale, IOS Press, 2011.

[8] Qi Zhang, Lu Cheng, Raouf Boutaba Cloud computing: state-of-the-art and research challenges Springer Published online: 20 April 2010

[9] Sanjay P. Ahuja, Sindhu Mani The State of High Performance Computing in the Cloud February 2012

[10] UC Berkeley Reliable Adaptive Distributed Systems Laboratory http://radlab.cs.berkeley.edu/ Above the Clouds: A Berkeley View of Cloud Computing February 10, 2009

[11] Amel Haji, Asma Ben Letaifa, Sami

Tabbane "Implementation of a virtualization solution: SaaS on IaaS" 2011 IEEE

[12] Software and Information Industry Association, "Strategic Backgrounder: Software as a Service," 2001, http://www.siia.net/estore/pubs/SSB-01.pdf

[13] T. Dierks and E. Rescorla, "The Transport Layer Security (TLS) Protocol Version 1.2," 2008, IETF RFC5246, http://www.ietf.org/rfc/rfc5246.txt

[14] Netscape, "The SSL Protocol: Version 3.0," Netscape/Mozilla, 1996. http://www.mozilla.org/projects/security/pki/nss/ssl/draft302.txt

[15] Napper, J. and Bientinesi, P., Can Cloud Computing Reach the TOP500?, Proceedings of the Workshop on Unconventional High Performance Computing, in conjunction with The ACM International Conference on Computing Frontiers, 18-20 May 2009

[16] Labate, B. and Korambath, P., Use of Cloud Computing Resources in an HPC Environment—IDREHPC Research Projects, 2009 International Conference on Computing Frontiers, 18-20 May 2009

[17] Subramanian, V., Ma, H., Wang, L., Lee, E. and Chen, P., Azure Use Case Highlights Challenges for HPC Applications in the Cloud, HPC in the Cloud, (Web:http://www.hpcinthecloud.com) February 21, 2011.

* * * * *