 ISSN : 0976-8491 (Online) | ISSN : 2229-4333 (Print)

Dynamic Virtual Machine Based Proactive Fault Tolerant Scheme Over Cloud

[1]Deepshikha Goutam, [2]Ashok Verma, [3]Neha Agrawal
[1,2,3]Dept. of CSE, GGITS, Jabalpur, India

Abstract
An important concerns in a cloud based environment are security, process fail rate and performance. A lot of research is currently underway to analyze how clouds can provide fault tolerance for an application. When numbers of processes are too many and any virtual machine is overloaded then the processes are failed causing lot of rework and annoyance for the users. The major cause of the failure of the processes at the virtual machine level are overloading of virtual machines, extra resource requirements of the existing processes etc.
This work introduces dynamic load balancing techniques for cloud environment which proactively decides whether the process can be applied on an existing virtual machine or it should be assigned to a different virtual machine created a fresh or any other existing virtual machine. so, In this way it can tackle the occurrence of fault.
Cloud computing load balancing of process load has been researched previously but proactive load balancing is an area where lot of work is still to be done and this paper proposes a mechanism which proactively decides the load on virtual machines and according to the requirement either creates a new virtual machine or uses an existing virtual machine for the assigning the process. Once a process completes it will update the virtual machine status on the broker service so that other processes can be assigned to it.

Keywords
Cloud, Virtual Machines, Processes, Proactive Load Balancing, Data Center, Data Center Broker, Performance of cloud computing, Fault Tolerance, Cloudsim.

I. Introduction

Cloud computing is becoming a buzz word in computer industry and everyone is looking to associate in one way or other with this brand new concept. Cloud computing is a very current topic and the term has gained a lot of traction being sported on advertisements all over the Internet from web space hosting providers, through data centers to virtualization software providers. Cutting through the hype of cloud computing is not an easy task as a simple web search suffices to convince that there are nearly as many definitions on what constitutes 'cloud computing' as there are players in the market seeking to gain new territory in that promising new business field.
A Cloud is a type of parallel and distributed system consisting of a collection of interconnected and virtualized computers that are dynamically provisioned and presented as one or more unified computing resources based on service-level agreements established through negotiation between the service provider and consumers.
Cloud Computing refers to both the applications delivered as services over the Internet and the hardware and systems software in the data centers that provide those services (SaaS).
IT companies are providing services to the general public for a fee on-demand. This type of service is called Public Cloud. On the other hand if the service is solely used within an organization and not shared with people outside of the organization it is called Private Cloud. There is also a third kind, a combination of public and private cloud. It is referred to as Hybrid.

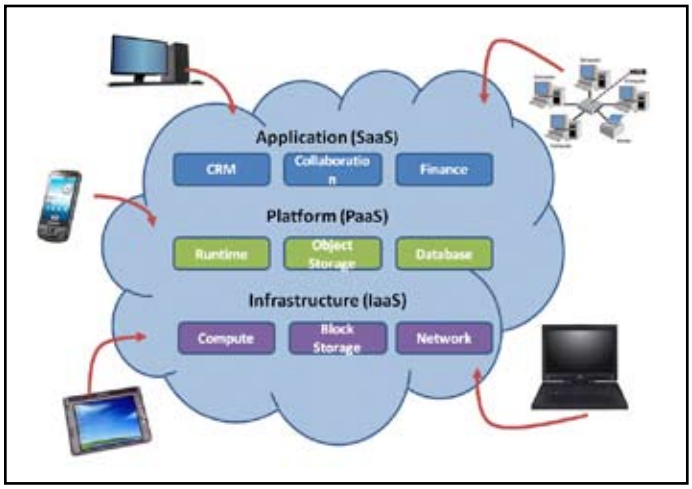

Fig. 1: Service Model of Cloud Computing

The cloud infrastructure is shared by several organizations and supports a specific community that has shared concerns (e.g., mission, security requirements, policy, and compliance considerations). It may be managed by the organizations or a third party and may exist on premise or off premise [6]. Example: eBay. Choosing which one to deploy purely depends on the needs.
Infrastructure as a Service(IaaS)- IaaS providers are companies that provide the most basic IT needs – servers, networking, and storage – on a usage-based payment model. They typically make heavy investments in data centers and other infrastructure, and then rent it out, allowing consumers to avoid investments of their own.
According to author [4] Current cloud computing offerings with virtual machines are based on communication via Ethernet. InfiniBand as an interconnect technology would be a better choice due to its superior performance. They present a novel architecture for HPC IaaS clouds supporting InfiniBand that allows an individual network configuration with QoS mechanisms. Customers are able to allocate virtual clusters which are isolated from other tenants in the physical network.
As more and more applications and content are being hosted and supported in the Cloud there is an ever growing need for supporting high-performance applications too. So researchers are moving their HPC applications to the cloud to understand the bottlenecks and challenges it presents. Models are designed and developed to predict the failures so they can be fixed before they can negatively impact the performance. One such experiment was performed using Microsoft's Azure cloud with Numerical Generation of Synthetic Seismograms [11].
According to the authors in [11] their HPC application creates "seismic waves in three dimensional complex geological media by solving some complex equations". While the results from the experiment were encouraging they also exposed some of the challenges that lie ahead of HPC in cloud. The authors in [11]

state that, "Real-time processing is a critical feature for synthetic seismogram". Massive amounts of data are used in computing and hence need the necessary infrastructure to support it. In the current state this is one of the challenges.

Platform as a Service (PaaS) It refers to providing platform layer resources, including operating system support and software development frameworks. Examples of PaaS providers include Google App Engine, Microsoft Windows Azure and Force.com.

Software as a Service(SaaS) - The term SaaS dates from the 1990s and thus predates cloud computing. SaaS is also known commonly as "Web services." While many slightly different definitions of SaaS are possible, a simple and usable definition has already been formulated:

"Software deployed as a hosted service and accessed over the Internet" [9].

Benefits of SaaS - Compared with traditional computing and software distribution solutions, SaaS clouds provide scalability and also shift significant burdens from subscribers to providers, resulting in a number of opportunities for greater efficiency and, in some cases, performance. 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

Fundamentally, cloud computing provides convenient rental of computing resources. These resources, which are typically accessed by subscribers over a network, must be measurable in units that can be individually allocated to specific subscribers, and paid for based on factors such as how long the units are retained, who has access to them, how they are used, etc. In the case of SaaS, what is being rented is access to an application [10]. Typically, access to the application is over a network connecting the SaaS provider with the subscriber. For public or outsourced SaaS, most application program logic is executed on the cloud provider's servers.

The SaaS provider's main responsibility to the subscriber is to ensure that the software that it supplies is solidly supported and tested. Another key requirement is that SaaS applications be scalable to increasingly larger subscriber workloads. Maintaining an infrastructure to carry this out in a secure environment with specified uptime for the subscriber is a critical aspect. Many subscribers may have valuable organizational data stored in the cloud and some of this information may be proprietary and business-sensitive, therefore a secure environment is paramount.

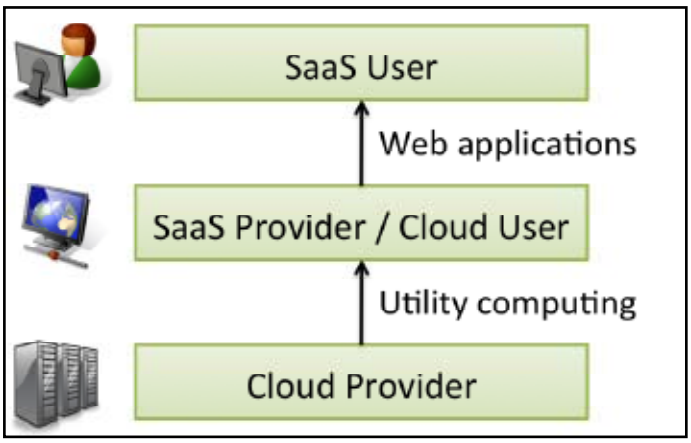

Fig. 2: Users and Providers of Cloud Computing.

II. Existing System

Author of [8], surveyed that It is difficult to define the cloud computing. Computing is a virtual pool of computing resources. It provides computing resources in the pool for users through internet. Integrated cloud computing is a whole dynamic computing system. It provides a mandatory application program environment. It can deploy, allocate or reallocate computing resource dynamically and monitor the usage of resources at all times. Generally speaking cloud computing has a distributed foundation establishment, and monitor the distributed system, to achieve the purpose of efficient use of the system.

Cloud computing collects all the computing resources and manages them automatically through software. In the process of data analysis, it integrates the history data and present data to make the collected information more accurate and provide more intelligent service for users and enterprises [8]. The users need not care how to buy servers, software's, solutions and so on. Users can buy the computing resource through internet according to their own needs. Cloud computing does not depend on special data center, but we can look it as the inevitable product of grid computing and utility computing. However, compared with general network service, cloud computing is easy to extend, and has a simple management style. Cloud is not only simply collecting the computer resource, but also provides a management mechanism and can provide services for millions of users simultaneously. Nowadays, virtualization is entering every field of data center [9]. It has become useful tool and improved service capacity. When the storage and computing capacity of the server cluster are surplus, we need not purchase servers, all we need to is to add a virtual machine running on the server. If the cluster is large enough, the request of adding server will have marginal effect, and then we can save the money that should be used in purchasing new servers. At the same time, cloud computing provides powerful supports for SAAS (software as a service) [9]. It integrates all the companies that provide similar services in the internet in order that users can compare and select service providers. Cloud computing provides dependable and secure data storage center, provides immense possibility for internet application, provides infinite space for storing and managing data, provides powerful computing capacity for users to complete all kinds of application. Future computer may only be used for connecting internet to implement services based on cloud computing. Users will change their habit of using computer totally, from services centered by desktop to services centered by Web.

Table below previews ranked list of critical obstacles to growth of Cloud Computing. The first three concern adoption, the next five affect growth, and the last two are policy and business obstacles. Each obstacle is paired with an opportunity, ranging from product development to research projects, which can overcome that obstacle [8].

Table 1: Preview of Top 10 Obstacles to and Opportunities for Growth of Cloud Computing

S. No.	Obstacle	Opportunity
1	Availability of Service	Use Multiple Cloud Providers; Use Elasticity to Prevent DDOS
2	Data Lock-In	Standardize APIs; Compatible SW to enable Surge Computing
3	Data Confidentiality and Audit ability	Deploy Encryption, VLANs, Firewalls; Geographical Data Storage

4	Data Transfer Bottlenecks	Fed Existing Disks; Data Backup/Archival; Higher BW Switches
5	Performance Unpredictability	Improved VM Support; Flash Memory; Gang Schedule VMs
6	Scalable Storage	Invent Scalable Store
7	Bugs in Large Distributed Systems	Invent Debugger that relies on Distributed VMs
8	Scaling Quickly	Invent Auto-Scalar that relies on ML; Snapshots for Conservation

Cloud computing blueprint is coming: in the future, we only need a notebook pc or a mobile phone, then we can complete what we want through net service including the huge tasks such as supercomputing. So end-user is the true owner of cloud computing. The aim of application of cloud computing is to combine all the resources, and let anyone can use it. From the most basic significance, cloud computing is to utilize software and data of the internet.

According to [12] most cloud delivery models require the application developers to build intrinsically reliable software taking into account environment specific features.

Creating and managing fault tolerance that shades the implementation details of the reliability techniques from the users by means of a dedicated service layer [1].but this is only theoretical concept not implemented.

According to author [3] they shown that there are two visions for fault management, first one consists in leaving exclusively the responsibility of fault management to one cloud participant, second sharing the responsibility between two cloud participants.

III. Problem Statement

According to various researchers, two main factors which affect the performance of cloud computing are:

A. Dynamic Scalability

As the number of user requests increase the application must be able to support the increasing load. At the same time as the number of requests decrease the application should be able to scale down. So achieving dynamic scalability is a challenge in the cloud. Windows Azure randomly de-allocates the compute nodes when scaling down and hence follows an asynchronous process. This negatively impacts the performance.

Both IaaS and PaaS provide services may be used to build and deploy scalable applications that can be optimized for parallel computing. In case of IaaS the infrastructure is already built and is readily available for providing services on-demand. Several bottlenecks such as delays, maintenance, operating costs etc. should be removed.

B. Fault Tolerance

Virtual nodes are created on demand to handle the load and to perform the computing tasks. So providing Virtual resources is another challenge for an application or HPC application in the cloud. If a virtual node fails while performing a task it becomes imperative to identify where in the system it failed and why? In order for the performance to be not impacted it is important that the load is transferred to another node while it is identified and fixed. It is also referred to as fault tolerance. This is one of the challenges while designing load balancing systems for high-performance applications. It is possible that a compute node may share its resources in running more than one application. As the number of applications increase the load on the compute node it can decrease the performance and may sometime fail if reached over capacity.

The concept of virtualization plays a key role for the implementation. Besides the construction of virtual on demand clusters there are further benefits that could be expected, such as live migration.

IV. Proposed System

Cloud Computing is facilitating users around the world for the best of the services available across the world on their machines through web. It is beneficial for both the service providers (they get huge clientele) and clients (they get all available services).

Steps of Proposed Work

- Using Cloud Sim 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 data center and will assign the new process to it.
- This step will avoid any possible fault which may occur in data center causing system instability or affecting other processes.
- The processes will be generated with very high speed over the proposed cloud.
- Various measures such as number of processes, number of virtual machines etc. shall be counted on time scale to show the working of the proposed system.

System shall be evaluated at the end:

- Using current System

1. In this, following modules shall be used during implementation and testing:
2. Datacenter
3. Datacenter Broker
4. Will Decide whether a new VM is required or not
5. Will submit the VM & Processes to the Data Center
6. Will Generate Results
7. Process Generator
8. Virtual Machine Generator
9. The whole system shall be executed twice:
10. Once without Making Decisions and generating VM for execution of the generated processes (It will be causing the various faults and number of processes shall be failed)
11. 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.

- Using Existing work of the various researchers

Base paper and other papers have been studied thoroughly and the work done by the corresponding researchers can be compared with implemented work so that better output of proposed work can be verified.

V. Conclusion

Studies of the various papers and works done by authors have been done to find out the problem and it is found that the cloud

computing is apparently a new technology which is growing very fast and provides new horizons to the computing world. It is technique where implementations are not too many and the major players in industry are very few. The situation is so because a lot of structural, architectural and security work in various applications of the cloud is still to be done. This work is focusing on proactive load balancing over the cloud for managing the processes and virtual machines so that waste of resources can be avoided and users will not be finding delays due to their process failures over the cloud computing environment.

Proactive management of the processes over cloud will increase the performance of the cloud and will avoid unwanted delays caused by the processes running over the system.

VI. Future Work

The proposed work is being implemented on simulation environment using standard machines, in future the same can be deployed over the real cloud environment and test it for its accuracy and performance.

A further improvement in the proposed algorithm may be required for real high performance computing environments over the cloud. At IaaS and PaaS layers changes may be helpful in increasing the performance of the proposed system.

References

[1] Ravi Jhawar, Vincenzo Piuri,"Marco Santambrogioy:A Comprehensive Conceptual System-Level approach to Fault Tolerance in Cloud Computing, IEEE2012.

[2] Sheheryar Malik Fabrice Huet Research Team OASIS2011 IEEE World Congress on Services Adaptive Fault Tolerance in Real Time Cloud Computing

[3] Alain Tchana Laurent Broto, Danie Hagimont,"Approache to Cloud Computing Fault Tolerance", 2012 IEEE

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

[5] 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.

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

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

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

[9] Amel Haji, Asma Ben Letaifa, Sami Tabbane "Implementation of a virtualization solution: SaaS on IaaS", 2011 IEEE

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

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

[12] J. Barr, A. Narin, J. Varia,"Building Fault-Tolerant Applications on AWS", October 2011. [Online] Available: http://media.amazonwebservices.com/AWS-Building-Fault-Tolerant-Applications.pdf

Deepshikha Goutam received his B.E. degree in Information Technology from Jabalpur Engineering College, Jabalpur, Madhya Pradesh, India, in 2007, and currently perusing M. TECH in Computer Science and Engineering from Gyan Ganga Institute of Science & Technology, Jabalpur, Madhya Pradesh, India. Her research interests are in Cloud Computing.

Ashok Verma received his Bachelor of Engineering degree in Computer Science and Engineering from Amravati University, Maharashtra, India, in 1996, Master of Engineering (Hons.) degree in Computer Science and Engineering from Rajiv Gandhi Technical University, Madhya Pradesh, India. He was Lecturer in Department of Computer Science and Engineering, Jabalpur Engineering College, Madhya Pradesh, India, from 1997 to 2006. He is an Assistant Professor and Head of Department, in Department of Computer Science and Engineering, Gyan Ganga Institute of Technology and Sciences, Jabalpur, India, since 2006. His research interests include Computer Networks, Network Management, Operating System, System Programming, Distributed Computing, Information Storage and Big Data Management. At present, He is engaged in Cloud Computing and Security technique in virtualization.

Ms. Neha Agrawal has done her bachelor of engineering in Computer Science from Shri Ram Institute of Technology, Jabalpur (Madhya Pradesh) in 2010, the M.tech. degree in computer science from Banasthali University, Jaipur (Rajasthan) in 2012. She has done her internship project on "Deployment of Private cloud" from Centre For development of Advanced computing (C-DAC), Pune. At Present, she is working as Asst. Professor in Department of Computer Science in Gyan Ganga Institute of Technology & Sciences, Jabapur (Madhya Pradesh), India.