Cloud computing – Innovation for Cyber Infrastructure Utilization

Sudhansh Sharma

Department of Information Technology, Vishveshwarya School of Business Management (VSBM),

sudhansh1974@gmail.com

Durgansh Sharma

Department of Information Technology, Jaipuria Institute of Management, NOIDA,, India

durgansh@gmail.com

Phase-II: Empirical Article

ABSTRACT

- Every ICT application needs a model of computation, a model of storage, and a model of communication. Since funds are involved in the implementation of every model, be it related to computation or storage or communication; so, its of prime concern that "How to optimize the effective usage of the available resources?", with no further expenses. Apart from this, the recurring costs like power consumption, is also to be taken care such that the organizational profits keep on growing. Thus, the statistical multiplexing necessary to achieve elasticity and the illusion of infinite capacity is required to be explored, such that each of these resources are used effectively, and at optimum level. The virtualization concept plays a dominant role in doing all this; the environment is virtualized to hide the implementation of how they are multiplexed and shared. In this paper we explore the possibilities, how the emerging data centre "Cloud" model of computing is going to more profitable and eco friendly, in comparison to the traditional computing model. The performed work is going to address the limitations of the usual Data centre, and how the same can be over come by bringing the concept of cloud computing in to the picture.

KEYWORDS

- Cloud Computing
- .Grid Computing
- Cyber
- Platform as a Service

Introduction:

Throughout the history of computing, there have been several paradigm shifts from main-frames to mini computing to micro processing to networked computers. On track to be the next major paradigm shift is that of cloud computing [1]. The shift of computing trend toward cloud computing started in the late 1980s with the concept of grid computing when, for the .first time, a large number of systems were applied to a single problem, usually scientific in nature and requiring exceptionally high levels of parallel computation. That said, it's important to distinguish between Grid Computing and Cloud Computing [2].

Cloud computing is often confused with the term Grid Computing ("a form of distributed computing whereby a 'super and virtual computer' is composed of cluster of networked loosely coupled computers, acting in concert to perform very large tasks"), Utility Computing (the "packaging of computing resources, such as computation and storage, as a metered service similar to a traditional public utility such as electricity") and Automatic Computing ("computer systems capable of self management"). Indeed many cloud computing deployments depend on grids, have automatic characteristic and utilities – but cloud computing can be seen as a natural next step from the grid-utility model. In short, Grid computing specifically refers to leveraging several computers in parallel to solve a particular, individual problem, or to run a specific application. Cloud computing, on the other hand, refers to leveraging multiple resources, including computing resources, to deliver a "service" to the end user.

Its not like that grid computing has no role in cloud computing, and it's a new technology developed from scratch, the previous technologies have considerable contribution in the evolution of the concept of cloud computing[2], refer to Figure 1 below.

Figure 1: Evolution of Cloud Computing [2]

So, to get Cloud Computing to work, you need three things: thin clients (or clients with a thick-thin switch), grid computing, and utility computing. Grid computing links disparate computers to form one large infrastructure, harnessing unused resources. Utility computing is paying for what you use on shared servers like you pay for a public utility (such as electricity, gas, and so on).

With grid computing, you can provision computing resources as a utility that can be turned on or off. Cloud computing goes one step further with on-demand resource provisioning. This eliminates over-provisioning when used with utility pricing. It also removes the need to over-provision in order to meet the demands of millions of users[4].

II - Cloud – services and classification

What is cloud computing?

Fundamentally , cloud computing can be defined as a push in designing services where information is stored and processed on internet (i.e. "the cloud") usually via massive large scale data centers which can be assessed remotely through various clients and platforms[8] . So, far as the offerings of the services through the internet is concerned [2] , So, Cloud computing in crux 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" [9]. Some examples of emerging Cloud computing infrastructures are Microsoft Azure [2], Amazon EC2, Google App Engine, and Aneka [3].

Services offered by cloud computing

Cloud Computing has often been referred as 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". Cloud computing delivers infrastructure, platform, and software (application) as services, which are made available as subscription-based services in a *pay-as-you-go model* to consumers. These services in industry are respectively referred to as Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS). Clouds [11] aim to power the next generation data centers by architecting them as a network of virtual services (hardware, database, user-interface, application logic) so that users are able to access and deploy applications from anywhere in the world on demand at competitive costs depending on users QoS (Quality of Service) requirements [9]. Developers with innovative ideas for new Internet services are no longer required to make large capital outlays in the hardware and software infrastructures to deploy their services or human expense to operate it [12]. It offers significant benefit to IT companies by freeing them from the low level task of setting up basic hardware (servers) and software infrastructures and thus enabling more focus on innovation and creation of business values.

Cloud classification- Public, private and Hybrid

The Cloud may be *Public Cloud* or *Private 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. So, far as the profitability in an eco friendly way is concerned, we will focus on SaaS Providers (Cloud Users) and Cloud Providers, and compare the issue with the provision and usage of usual data centre. *Hybrid clouds* are a combination of Public and Private clouds such that the secured components are usually contained on premises and the compute load is usually spread over the public cloud. These are planned for the future by major vendors like Microsoft for Windows Azure. But this can be custom assembled as of now using Public and Private cloud offerings.

Centralized And Distributed Clouds.

Most of the talk about Cloud Computing deals with *Centralized Clouds*. Which means the Cloud Provider will setup virtualized data centers in a geographical spread providing for failover zones and accommodating special regional requirements continuously built upon since then. The second category is *Distributed clouds,a* distributed cloud is usually based on P2P technologies and allows global distribution and pooling of compute resources, storage and enterprise integration.

SaaS, Platform and Infrastructure Clouds.

SaaS, the simplest and earliest of Cloud Offerings were *SaaS (Software as a Service)* offerings. Which basically provided a hosted multitenant application for use using a pay as you go model. E.g. Google Apps, Office Live and Zoho. Here no customization or plugins are available which an end users or developers could add to its functionality. *Platform cloud,*Then came the Platform cloud offerings which offered a Platform to uers in a pay as you go cloud hosted model. These were usually limited to a specific domain like CRM e.g.SalesForce or Social Collaboration e.g.FaceBook but offered an API that could be used to extend their features. Mostly accompanied by a marketplace for buying third party features.*Infrastructure clouds,* Infrastructure clouds are usually those offerings which give you an infrastructure like Windows Azure or Amazon upon which you can build almost any application.

III - Cyber infrastructure Resource utilization

The customers engaging in cloud computing do not own the physical infrastructure serving as host to the software platform in question. Instead they avoid capital expenditure by renting usage from a third party provider. They, consume resources as a service, paying instead for the resources they use. Many cloud computing offerings have adopted the utility computing model, which is analogous to how traditional utilities like electricity are consumed, while other are billed on subscription basis. Sharing "perishable" and "tangible" computing power among multiple tenants can improve utilization rates, as servers are not left idle, which can reduce costs significantly while increasing the speed of the application development. A side effect of this approach is that "computer capacity rises dramatically" as customers do not have to engineer for Peak loads, adoption has been enabled by "increased high-speed bandwidth" which makes it possible to receive the response time from centralized infrastructure at other sites [3].

Since, there are many ways in which computational power and data storage facilities are provided to users, ranging from a user accessing a single laptop to the allocation of thousands of computer nodes distributed around the world. Users generally locate resources based on a variety of characteristics, including the hardware architecture, memory and storage capacity, network connectivity and, occasionally, geographic location. Usually this resource location process involves a mix of resource availability, application performance profiling, software service requirements, and administrative connections. While great strides have been made in the HPC and Grid Computing communities [16,14] toward the creation of resource provisioning standards[15,17,18,19], this process remains somewhat cumbersome for a user with complex resource requirements, for which the answer is cloud computing[7]. It's the Cloud computing which delivers infrastructure, platform, and software (application) as services, which are made available as subscription-based services in a pay-as-you-go model to consumers. These services in industry are respectively referred to as Infrastructure as a Service (Iaas), Platform as a Service (PaaS), and Software as a Service (SaaS). On demand availability of cyber infrastructure, largely involves funds, and major of them for hardware setups. Thus in cloud computing, from a hardware point of view, three aspects are new which are :-

1. The illusion of infinite computing resources available on demand, thereby eliminating

the need for Cloud Computing users to plan far ahead for provisioning.

2. The elimination of an up-front commitment by Cloud users, thereby allowing companies to start small and increase hardware resources only when there is an increase in their needs.
3. The ability to pay for use of computing resources on a short-term basis as needed (e.g., processors by the hour and storage by the day) and release them as needed, thereby rewarding conservation by letting machines and storage go when they are no longer useful.

Hence, the key necessary enabler of cloud computing is the construction and operation of extremely large-scale, commodity-computer datacenters at low cost locations, which uncovers the factors of 5 to 7 decrease in cost of electricity, network bandwidth, operations, software, and hardware available at these very large economies of scale. These factors, combined with statistical multiplexing to increase utilization compared a private cloud, meant that cloud computing could offer services below the costs of a medium-sized datacenter and yet still make a good profit.

IV - Datacenters & the technology behind cloud computing – a profitable option.

Virtualization and Automation are the technologies that work behind Cloud Computing has existed in Data Centers for well over the last decade, Virtualization allows using high end physical machines with multiple virtual machines running on them, where as the technology of Automation usually means that snap shots of these virtual machines can be kept on a network storage and these snap shots can be provisioned dynamically within minutes, further in case of any hardware failure or needs they can be migrated to other physical machines, providing failover and scalability.

This doesn't mean that Clouds Computing is just a synonym of Advanced Data Centers, Which is not very accurate because usually a cloud provider provides add-on infrastructure or platform services. Actually, the computing power in a Cloud computing environments is supplied by a collection of data centers, which are typically installed with hundreds to thousands of servers [9]. The layered architecture of a typical Cloud based data center is shown in Figure 2. At the lowest layers there exist massive physical resources (storage servers and application servers) that power the data centers. These servers are transparently managed by the higher level virtualization [10] services and toolkits that allow sharing of their capacity among virtual instances of servers. These virtual instances are isolated from each other, which aid in achieving fault tolerant behavior and isolated security context.

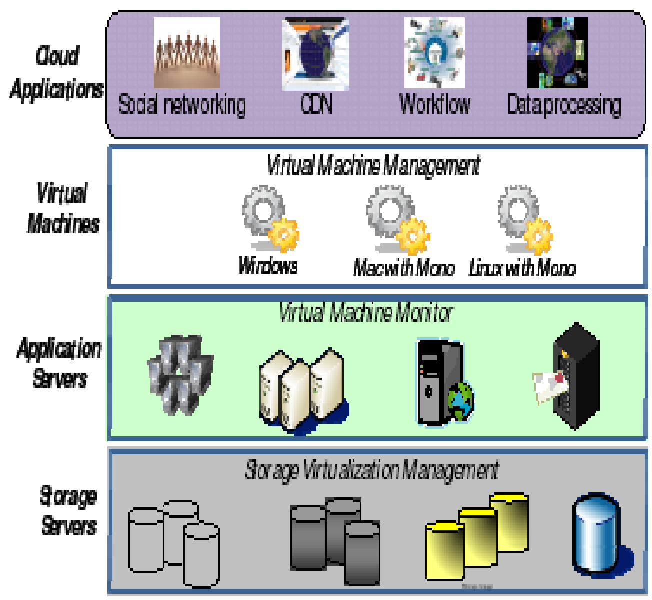

Figure 2 : The layered architecture of a typical Cloud based data center[5]

An organization of any size will have a substantial investment in its data center. That includes buying and maintaining the hardware and software, providing the facilities in which the hardware is housed and hiring the personnel who keep the data center running. An organization can streamline its data center by taking advantage of cloud technologies internally or by offloading workload into the public[5]. But without the concept of cloud computing i.e. independently, the Data Centers are known to be expensive to operate and they consume huge amounts of electric power. For example, the Google data center consumes power as much as a city such as San Francisco. As Clouds are emerging as next-generation data centers and aim to support ubiquitous service-oriented applications, it is important that they are designed to be energy efficient to reduce both their power bill and carbon footprint on the environment. To achieve this at software systems level, we need to investigate new techniques for allocation of resources to applications depending on quality of service expectations of users and service contracts established between consumers and providers [13][6].

Thus the user of this emerging technology i.e. cloud computing are going to enjoy cost effective green computing environment because the user can avoid capital expenditures on hardware, software and services, rather paying a provider only for what they use. Consumption is billed on utility of the resources (e.g. resources consumed like electricity) or subscription (e.g. time based, like a newspaper) basis with little or no upfront cost. Other benefits of this time sharing style approach are low batteries to entry, shared infrastructure and costs, low management overheads and immediate access to a broad range of applications. Users can generally terminate the contract at any time (there by avoiding return to investment risk and uncertainty) and the services are often covered by SLAs- Service Level Agreements with financial penalties[4].

However, every operation of any ICT application needs a model of computation, a model of storage, and a model of communication. The statistical multiplexing necessary to achieve elasticity and the illusion of infinite capacity requires each of these resources to be virtualized to hide the implementation of how they are multiplexed and shared. In order to achieve this, different utility computing offerings should be distinguished based on the level of abstraction presented to the programmer and the level of management of the resources. Which in turn leads to profits, but to gain one should analyze that, When is Utility Computing preferable to running a Private Cloud?

Since, profitability is generally the key Moto behind running a private cloud, to work out this key parameter, we explore the practical situation in the form of cases, which are :-

Case-1: *When demand for a service varies with time.* Provisioning of the data center for peak load may sustain for few days per month but this leads to underutilization at other times. Instead, Cloud Computing lets an organization pay by the hour for computing resources, potentially leading to cost savings even if the hourly rate to rent a machine from a cloud provider is higher than the rate to own one.

Case-2: *When demand is unknown in advance*, example, a web startup will need to support a sudden rise in demand when it becomes popular, however it might be followed potentially by a reduction in this spike, when some of the visitors turn away.

For both cases i.e. Case-1&2, where a web business is unpredictable at any instance of time, that is the resource demand is varying over time and revenue is proportional to user hours, then to capture the tradeoff, the equation (1) is derived, where the expected profits made by using cloud computing are on the LHS whereas the profits made by a fixed capacity Datacenter are on the RHS of equation (1) given below.

$$\left[\text{UserHours\_Cloud} \times \left(revenue - Cost\_Cloud\right)\right] \geq \left[\text{UserHours\_Datacenter} \times \left(revenue - \frac{Cost\_Datacenter}{avg\_Utilization}\right)\right]$$

Assumptions made while deriving equation (1) are :

a) It is assumed that the Cloud Computing vendor employs usage-based pricing model, in which customers pay proportionally to the amount of time and the amount of resources they use.

   Although more sophisticated pricing models for infrastructure services utilization [28, 6, 40] are available, but we worked with usage based pricing model because it will persist as a consequence of its simplicity and transparency, as demonstrated by its wide usage by real life utilities for electricity and gas companies.

b) Further, it is assumed that the customer's revenue is directly proportional to the total number of user-hours.

Analyzing equation (1), the left-hand side multiplies the net revenue per user-hour (revenue realized per user-hour minus cost of paying Cloud Computing per user-hour) by the number of user-hours, giving the expected profit from using Cloud Computing. The right-hand side performs the same calculation for a fixed-capacity datacenter by factoring in the average utilization, including nonpeak workloads, of the datacenter. *Whichever side is greater represents the opportunity for higher profit.*

Apparently, if Utilization = 1.0 (the datacenter equipment is 100% utilized), the two sides of the equation look the same. However, basic Queueing theory tells us that as utilization approaches 1.0, system response time approaches infinity. In practice, the usable capacity of a datacenter (without compromising service) is typically 0.6 to 0.8.Whereas a datacenter must necessarily overprovision to account for this "overhead," the cloud vendor can simply factor it into Cost cloud. (This overhead explains why we use the phrase "pay-as-you-go" rather than rent or lease for utility computing. The latter phrases include this unusable overhead, while the former doesn't. Hence, even if you lease a 100 Mbits/second Internet link, you can likely use only 60 to 80 Mbits/second in practice.) The equation makes clear that the cloud computing environment has better ability to control the cost per user hour of operating the service.

V – Result

Comparison of the factor of profitability for both Datacenter and Cloud computing, technologies for cyber infrastructure utilization is presented. It is found that cloud computing is a better option to manage the environment where the demand of the cyber infrastructure changes unpredictably.

VI – Conclusion

In this paper details, related to the evolution of cloud computing and its current classification is discussed, apart from this the confusions related to cloud computing and other technologies are presented. However from the discussion made in this paper ,it is concluded that the emerging trend of cloud computing is not only novel but also an innovative way to effectively use the cyber infrastructure and achieve cost effective, profitable and eco friendly environment.

VII – References

[1] Justin M. Grimes, Paul T. Jaeger, Jimmy Lin; "Weathering the Storm: The Policy Implications of Cloud Computing" available at nora.lis.uiuc.edu/images/iConferences/CloudAbstract13109FINAL.pdf

[2] Seeding the clouds : Key infrastructure elements for cloud computing by IBM , available at ftp://ftp.software.ibm.com/common/ssi/sa/wh/n/oiw03022usen/OIW03022USEN.PDF

[3]Cloud computing article available at http://en.wikipedia.org/wiki/Cloud_computing

[4]http://www.ibm.com/developerworks/web/library/wa-cloudgrid/

[5] Rodrigo N. Calheiros, Rajiv Ranjan, César A. F. De Rose, and Rajkumar Buyya; "CloudSim: A Novel Framework for Modeling and Simulation of Cloud Computing Infrastructures and Services"CloudSim-ICCP2009 available at http://www.gridbus.org/reports/CloudSim-ICPP2009.pdf

[6]Rajkumar Buyya, Chee Shin Yeo, Srikumar Venugopal, James Broberg, Ivona Brandic "Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility" available at http://www.gridbus.org/papers/Cloud-FGCS2009.pdf

[7] Daniel Nurmi, Rich Wolski, Chris Grzegorczyk Graziano Obertelli, Sunil Soman, Lamia Youseff, Dmitrii Zagorodnov, "The Eucalyptus Open-source Cloud-computing System"available at open.eucalyptus.com/documents/nurmi_et_al-eucalyptus_open_source_cloud_computing_system-cca_2008.pdf

[8] A. Weiss. Computing in the clouds. netWorker, 11(4):16{25, December 2007.

[9] R. Buyya, C. S. Yeo, and S. Venugopal. Market-oriented cloud computing: Vision, hype, and reality for delivering IT services as computing utilities.*In Proceedings of the 10th IEEE International Conference on High Performance Computing and Communications*, 2008.

[10] J. E. Smith and R. Nair. *Virtual Machines: Versatile platforms for systems and processes*. Morgan Kauffmann, 2005.

[11] A. Weiss. Computing in the clouds. *NetWorker*, 11(4):16–25, Dec. 2007.

[12] M.Armbrust,A.Fox,R.Griffith,A.Joseph, R. Katz, A.Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica,M. Zaharia. *Above the Clouds: A Berkeley View of Cloud computing*. Technical Report No. UCB/EECS-2009-28, University of California at Berkley, USA, Feb. 10, 2009.

[13] K.H. Kim, R. Buyya, J. Kim, Power aware scheduling of bag-of-tasks applications with deadline constraints on DVS-enabled clusters, in: Proc. 7th IEEE Int. Symposium

on Cluster Computing and the Grid, CCGrid 2007, Rio de Janeiro, Brazil, May 2007.
[14] F. Berman, G. Fox, and T. Hey. Grid Computing: Making the Global Infrastructure a Reality. Wiley and Sons, 2003.
[15] I. Foster and C. Kesselman. Globus: A metacomputing infrastructure toolkit. International Journal of Supercomputer Applications, 1997.
[16] I. Foster and C. Kesselman, editors. The Grid – Blueprint for a New Computing Infrastructure. Morgan Kaufmann, 1998.
[17] D. Gannon. Programming the grid: Distributed software components, 2002.
[18] NSF TeraGrid Project. http://www.teragrid.org/.
[19] T. Tannenbaum and M. Litzkow. The condor distributed processing system. Dr. Dobbs Journal, February 1995.

Volume-1 Issue-2
July 2009-December 2009
Phase-II: Empirical Article