A Survey on Cloud Computing

Subhadra Bose Shaw
1st year Ph.D Scholar
CSED,MNNIT,Allahabad
subhadra.shaw@gmail.com

Dr. A.K.Singh
Associate Professor
CSED,MNNIT,Allahabad
asbhadoria@gmail.com

Abstract- The long-dreamed vision of "computing as a utility" has finally taken shape in the form of cloud computing. This paradigm shift is the biggest buzz in today's computer world. It can be thought of an evolution of existing technologies rather than a revolution. The basic requirements of cloud are to meet the various needs of enterprise organizations such as scalability, adaptability, extensibility and manageability. The biggest visible change the cloud computing has made in the scenario of computing is moving the data center offsite to a third party and buying services rather than maintaining onsite applications directly. By its use companies can save money on operational costs and they can focus on their core business instead of being worried about different IT obstacles. Moreover true portability is provided by cloud by enabling access to applications and documents anywhere in the world via the internet through commodity hardware only. In this paper we have described the basic concepts of cloud computing in a nutshell so that the readers can get an essence of this newly emerging technology.

Keywords - Cloud Computing, Data Lock-in, Green IT, Grid Computing, Utility Computing, Virtualization.

1 Introduction

The journey to the cloud marks a momentous evolution in IT industry. It is changing the way we all work and manage our systems regardless of industry or size. One of the formal definitions of cloud computing as given by NIST is [1] - *"Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. This cloud model is composed of five essential characteristics, three service models and four deployment models."*

So cloud computing is a way for small organization to compete with much larger ones, it's a way to save a lot of money and to utilize energy efficiently. The users only have to worry about their bandwidth connection and the rest will be taken care of by the cloud provider for which they will pay a nominal amount according to their usage.

The paper proceeds as follows: In the second section we have given a brief history of cloud computing. Section three describes the cloud components, its characteristics, service models and deployment models. Benefits of cloud are discussed in section four which is followed by the obstacles in cloud. Finally the last section concludes the topic.

2 Evolution of Cloud Computing

The concept of cloud computing is not new. The idea was first suggested by John Mc Carthy in 1960s when he proposed the concept of "Utility Computing" [2]. Like, utility computing, at its most basic level, cloud treats computing as a utility rather than a specific product or technology. The former is defined as a packaging of computing resources such as computation, storage and devices as a metered service like other normal household services such as water, electricity, gas and telephone[2]. The model has the advantage of having low initial cost to acquire resources instead computational resources are essentially rented on demand basis, thereby avoiding under-utilization and/or over-utilization of resources.

Cloud computing evolves gradually after grid computing. Cloud computing is also defined as the user-friendly version of grid computing [4]. One of the popular definitions for Grids is [3] *''A Grid is a type of parallel and distributed system that enables the sharing, selection, and aggregation of geographically distributed 'autonomous' resources dynamically at runtime depending on their availability, capability, performance, cost, and users' quality-of-service requirements.''*

The motivation behind grid computing was to solve large-scale, resource-intensive applications that require more resources than a single PC[3]. At a glance, cloud may be confused with grid computing. The distinctions are not clear maybe because Clouds and grids share similar visions: reduce computing costs and increase flexibility and reliability by using third-party operated hardware [5].
In grid computing, a large project is divided among multiple computers to make use of their resources. Cloud computing does just the opposite. It allows multiple smaller applications to run at the same time. Clouds are more popular than grids due to its user-friendly, virtualized and automatically scalable utilities. Pay-as-you-go model makes cloud more attractive as compare to the fixed payment scale of grid [5]. Unlike Grids, scalability of Cloud resources allows real-time provisioning of resources to meet application requirements as opposed to the traditional approach that required advance reservation of resources in global multi-user Grid environments. Moreover each grid has a different configuration, which results in extra effort each time an application needs to be ported to a new site. But in cloud virtual machines allow the application

developer to create a customized, portable execution environment configured specifically for their application.

So cloud is a much bigger concept than grid and utility computing and is not restricted to a specific network. The two key technologies behind cloud are Virtualization and Clustering.

Virtualization refers to the abstraction of logical resources away from their underlying physical resources for achieving agility, flexibility and energy-efficient resource utilization. Virtualizations in cloud are of different types such as server, storage and network virtualization [11]. Server virtualization is the mapping of single physical resources to multiple logical representations or partitions. Storage virtualization attempts to maximize the efficiency of storage devices. By virtualization a network, multiple networks can be combined into a single network, or a single network can be separated logically into multiple parts. So virtualization [12] is extremely well suited to a dynamic cloud infrastructure.

Cluster computing can be described as the integration of more than one commodity computers and resources incorporated through hardware, networks, and software to create a single system image [13]. The advantage of a cluster system lies in the ability on handling large and extremely complex computations (HPC) on more than one computer, working on the same problem or part thereof, simultaneously [14].

"Cloud is a type of parallel and distributed system consisting of a collection of inter-connected and virtualized computers that are dynamically provisioned and presented as one or more unified computing resource(s) based on service-level agreements established through negotiation between the service provider and consumers" [3].

3 Cloud Computing Overview

3.1 Components of Cloud – Clients, data center and distributed servers are the main three components of cloud computing [6].

Clients – They are the devices with which end users interact to manage their information on the cloud. Client can be smart phones or computers without hard disk (thin client) or a regular computer (thick client). A thin client is generally a web browser such as Mozilla Firefox or Microsoft Internet Explorer or Google's Chrome.

applications include: Customer Relationship Management (offered by Salesforce.com), video conferencing, mail services and data sharing, Google docs, Accounting, Web analytics, web content management. Apart from cost cutting the other benefits of SAAS include:

- Less IT staffs are required.
- SAAS applications are easier to customize.
- Marketing will be more efficient for a software provider.

Data Center – It is a collection of servers containing the application subscribed by the customers who can access it via internet. Many virtual machines can simultaneously execute on a single physical server also known as host. The number of virtual servers that can exist on a physical server depends on the size and speed of the physical server and what applications will be running on the virtual server [6].

Distributed Servers – To provide more reliability and availability servers are placed in geographically dispersed locations. So if one server fails service can be continued by the other. They also increase scalability. If the cloud needs more hardware, it is not essential to deploy more servers in the data center only. They can be added at another site and can be simply made a part of the cloud [6].

3.2 Essential Characteristics - The NIST definition describes five essential characteristics of cloud computing [7].

Rapid Elasticity: It is defined as the ability to scale resources both up and down as needed.

Measured Service: Cloud services are controlled and monitored by the cloud provider. This is crucial for billing, access control, resource optimization, capacity planning and other tasks.

On-Demand Self-Service: It means that a consumer can use cloud services as needed without any human interaction with the cloud provider.

Ubiquitous Network Access: It means that the cloud provider's capabilities are available over the network and can be accessed through standard mechanisms by both thick and thin clients.

Resource Pooling: It allows a cloud provider to serve its consumers via a multi-tenant model. Physical and virtual resources are assigned and reassigned according to consumer demand. The customer generally has no control or knowledge over the exact location of the provided resources but may be able to specify location at a higher level of abstraction (e.g., country, state, or datacenter).

3.3 Service Models – The NIST definition of cloud computing defines three service models, widely known as "as a service" [6].

1. Software as a Service (SAAS) – In this model an application is hosted as a service to customer who accesses it via the internet. So customer remains free from the headache of updation and maintenance of software. Some of these

- More reliable since if one server fails another server will take off.
- Secure Socket Layer (SSL) allows customer to reach their application securely.
- Due to increase in bandwidth organization can access their application with less latencies and high speed.

Some of the limitations of SAAS are: an organization may not find very much computational specific application

through SAAS. Moreover if a customer wants to move his application to a new vendor then the current provider may not allow it or may charge a huge amount of money. This problem is well known as lock-in problem.

2. Platform as a Service (PAAS) – It supplies all the resources required to build applications. There is no need to download or install software. It is also known as cloudware. Services provided by PAAS are application design, development, testing, deployment and hosting. Google AppEngine and Microsoft Azure are two examples in this category. Some benefits of PAAS include: the ability of geographically isolated development teams to work together, to merge web services from multiple sources, to realize cost savings from using built-in infrastructure services for security, scalability, and failover, rather than having to obtain and test them separately, to realize cost savings from using higher-level programming abstractions.

The main reason of downfall of PAAS is that platforms provided by different vendors are typically not compatible. Lack of interoperability and portability among providers prevent customers to move from one vendor to another.

3. Infrastructure as a Service (IAAS) – It provides the hardware on which application provided by SAAS and PAAS can work on. So it is also known as Hardware as a Service (HAAS). The physical assets provided by IAAS are storage space, network equipment and computing power. The infrastructure provided can be scaled up and down based on the possible resource needs. Further, multiple tenants can be on the equipment at the same time. Resources are typically billed based on a utility computing basis, so providers charge by how many resources are consumed. Amazon EC2 (Elastic Cloud Compute), Amazon S3 (Simple Storage Service) are examples of IAAS. Several components of IAAS are :

• Service Level Agreements (SLA) - This is an agreement between the provider and client, guaranteeing a certain level of performance from the system.

• Computer hardware - These are the components whose resources will be rented out. Service providers often have this set up as a grid for easier scalability.

• Network - This includes hardware for firewalls, routers, load balancing, and so on.

• Internet connectivity - This allows clients to access the hardware from their own organizations.

• Platform virtualization environment - This allows the clients to run the virtual machines they want.

• Utility computing - Billing typically set up to bill customers based on how many system resources they use.

3.4 Deployment Model - The NIST definition defines four deployment models [7]:

• Public Cloud: Public cloud services are available to clients from a third party service provider via the Internet. It can be free or fairly inexpensive to use. A public cloud does not mean that a user's data is publically visible; public cloud vendors typically provide an access control mechanism for their users, e.g. Google and Amazon, which expose their services to companies and consumers via the Internet [15].

• Private Cloud: It is also elastic and service based but unlike public cloud data and processes is managed within the organization without the restrictions of network bandwidth, security exposures and legal requirements. In addition, private cloud services offer the provider and the user greater control of the cloud infrastructure, improving security and resiliency. It is not necessarily managed and hosted by the organization that uses it and can be physically located off premises but always retains full control over the infrastructure [16].

• Community Cloud: It is controlled and used by a group of organizations that have shared interests, such as specific security requirements or a common mission. The members of the community share access to the data and applications in the cloud.

• Hybrid Cloud: It is a combination of a public and private cloud. In this model users typically outsource non-business-critical information and processing to the public cloud, while keeping business-critical services and data in their control. It may add complexity regarding the distribution of applications across different environments, monitoring of the internal and external infrastructure involved, security and privacy, and may therefore not be suited for applications requiring complex databases or synchronization [15].

4 Reason behind popularity of cloud services

Cloud services are popular because they can reduce the cost and complexity of owning and operating computers and networks. Since cloud users do not have to invest in information technology infrastructure, purchase hardware, or buy software licenses, the benefits are low up-front costs, rapid return on investment, rapid deployment, customization, flexible use, and solutions that can make use of new innovations. In addition, cloud providers that have specialized in a particular area (such as e-mail) can bring advanced services that a single company might not be able to afford or develop [8].

Some other benefits to users include scalability, reliability, and efficiency. Scalability means that cloud computing offers unlimited processing and storage capacity. The cloud is reliable in that it enables access to applications and documents anywhere in the world via the Internet. Cloud computing is often considered efficient because it allows organizations to free up resources to focus on innovation and product development.

Another potential benefit is that personal information may be better protected in the cloud. Specifically, cloud computing may improve efforts to build privacy protection

into technology from the start and the use of better security mechanisms. Cloud computing will enable more flexible IT acquisition and improvements, which may permit adjustments to procedures based on the sensitivity of the data. Widespread use of the cloud may also encourage open standards for cloud computing that will establish baseline data security features common across different services and providers. Cloud computing may also allow for better audit trails. In addition, information in the cloud is not as easily lost (when compared to the paper documents or hard drives, for example).

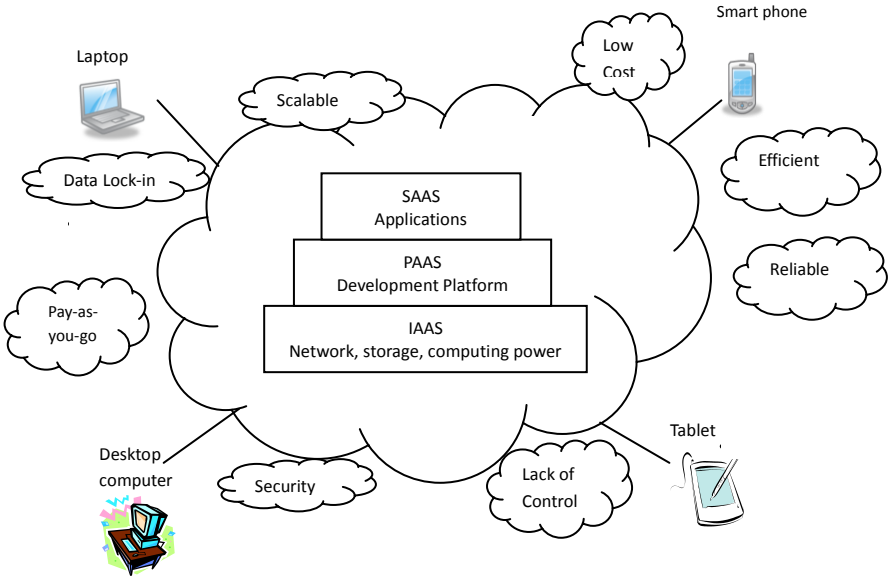

Fig 1: Overview of Cloud Computing

5 Obstacles and Opportunities for cloud computing

As we have already mentioned the benefits associated with cloud computing now we are going to discuss obstacles in its growth [9]. Each obstacle is paired with an opportunity – our thought on how to overcome the obstacle.

- Availability of Service – Most of the organizations worry about whether cloud computing services will have adequate availability or not. To achieve this there must be no single point of failure. Though the service provider has multiple data centers spread all over the world, it may have common software infrastructure and accounting systems or the company may even go out of business. So large companies are reluctant to migrate to cloud computing.

Another availability obstacle is Distributed Denial of Service (DDOS) attacks which prevent the customers from getting the required service in time.

- Data Lock-in – Cloud computing does not allow portability of application from one cloud provider to another or it may be very expensive to afford. Customer lock-in may be very attractive to cloud providers but users are vulnerable to price increment and reliability problem. Even if the provider goes out of business then the client may loss all or some of its crucial data. So the standardization of cloud computing is necessary which will make different cloud platforms interoperable and remove the fear of customers related to data lock-in. Cloud Computing Interoperability Forum (CCIF) [17] was formed to define an organization that would enable interoperable enterprise-class cloud computing platforms through application integration and stakeholder cooperation.

- Security – This is one of the major issue due to which customers are still reluctant to deploy their business in the cloud in spite of its benefits. Since customers' data is stored outside their premises and they are unaware of its location, so they are unable to protect it from unauthorized access. It is the duty of the cloud provider to ensure data security and integrity. It involves the use of strong encryption techniques for data security and fine-grained authorization to control access to data.

- Load Balancing - The cloud computing platform needs to dynamically balance the load among the servers in order to avoid hotspot and improve resource utility. Therefore, how to dynamically and efficiently manage resources and to meet the needs of subscribers is a challenge for the researchers [18].

Virtualization technology provides an effective solution to manage the dynamic resources on cloud computing

platform. All the requests are accepted by the virtual machines and are mapped to the appropriate physical servers by the cloudlet scheduler. For proper load balancing and resource utilization VMs must be migrated [19] from one physical server to another. So which VM to move, where to move and when to move are all important research issues.

- Scalable Storage – As discussed earlier the most appealing property of cloud are short term usage, no upfront cost and infinite storage and processing capacity on-demand. Here the opportunity for researchers is to create a storage system that would not only meet these needs but also combine them with the cloud advantages of scaling arbitrarily up and down as per demand.

- Bugs in Large-Scale Distributed Systems - One of the difficult challenges in Cloud Computing is removing errors in these very large scale distributed systems. A common occurrence is that these bugs cannot be reproduced in smaller configurations, so the debugging must occur at scale in the production data centers. One opportunity may be the reliance on virtual machines in Cloud Computing. Many traditional SAAS providers developed their infrastructure without using VMs, either because they preceded the recent popularity of VMs or because they felt they could not afford the performance hit of VMs. The level of virtualization may make utility computing possible to capture valuable information in ways that are implausible without VMs.

- Energy Consumption – Large-scale computing data centers are established for supporting high performance computation. In spite of using energy-efficient hardware, overall energy consumption continues to grow due to increasing requirements of computing resources [21]. Moreover, there are other problems associated with high power consumption like emission of carbon dioxide contributing to the green house effect, insufficient or malfunctioning cooling system which can cause overheating of the resources etc. So it's a challenge for the researchers to achieve energy efficiency in cloud for developing Green IT.

Conclusion – Cloud computing has gained considerable attention in recent times. It basically refers to an arrangement under which a user relies on another party to provide access to remote computer and software whose whereabouts, including their locations are neither known nor controllable by the user.

From the cloud provider's view, the construction of very large datacenters at low cost sites using commodity computing, storage, and networking uncovered the possibility of selling those resources on a pay-as-you-go model below the costs of many medium-sized datacenters, while making a profit by statistically multiplexing among a large group of customers.

From the customers' view they are free to concentrate on their work only instead of spending their time and money in the initial setup and regular maintenance of IT infrastructure of their organization. As Cloud Computing users, we were relieved of dealing with the twin dangers of over-provisioning and under-provisioning our internal data centers.

The ultimate goal of cloud computing is to provide application, platform, infrastructure as the fifth public utility after water, gas, electricity and telephone. There are still many challenges in this field including computing models, datacenters, security and privacy, virtualization etc.

Improving the resource utilization and reduce power consumption are key challenges to the success of operating a cloud computing environment. To address such challenges, GreenCloud architecture is designed [20]. It reduces unnecessary power consumption in a cloud computing environment.

In spite of all these obstacles, we believe that over the long run providers will successfully navigate these challenges and set an example for others to follow, perhaps by successfully exploiting the opportunities that correspond to those obstacles.

References:

[1] Peter Mell, Timothy Grance. The NIST Definition of Cloud Computing (Draft). NIST. 2011.

[2] Cloud Computing – The complete cornerstone guide to cloud computing best practices.pp-18.

[3] 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, Future Generation Computer Systems, Elsevier,2008.

[4] Jeremy Geelan. Twenty one experts define cloud computing. virtualization, August 2008. Electronic Magazine, article available at http://virtualization.syscon.com/node/612375.

[5] Luis M. Vaquero , Luis Rodero Merino , Juan Caceres , Maik Lindner: A Break in the Clouds: Towards a Cloud Definition. ACM SIGCOMM Computer Communication Review. Vol 39, Jan 2009.

[6] Anthony T. Velte, Toby J. Velte, Robert Elsenpeter : Cloud Computing – A Practical Approach. McGraw-Hill.

[7] A white paper produced by the Cloud Computing Use Case Discussion Group Cloud Computing Use Cases. Version 3.0, February 2010.

[8] Introduction to Cloud Computing. https://www.priv.gc.ca/resource/fsfi/02_05_d_51_cc_e.pdf.

[9] Michael Armbrust, Armando Fox, Rean Griffith, Anthony D. Joseph, Randy H. Katz, Andrew Konwinski, Gunho Lee, David A. Patterson, Ariel Rabkin, Ion Stoica, Matei Zaharia: Above the Clouds: A Berkeley View of Cloud Computing.

[10] Anuj Bala, Dr.Inderveer Chana : A Survey of Various

Workflow Scheduling Algorithms in Cloud Environment. 2nd National Conference on Information and Communication Technology (NCICT) 2011

[11] Bhaskar Prasad Rimal, Ian Lumb, Eunmi Choi : A Taxonomy and Survey of Cloud Computing Systems. Fifth International Joint Conference on INC, IMS and IDC 2009.

[12] Sun Microsystems, "Virtualization for Dummies".

[13] Giorgio Luigi Valentini, Walter Lassonde , Samee Ullah Khan, Sajjad A. Madani, Juan Li Pascal Bouvry : An overview of energy efficiency techniques in cluster computing systems. Springer Science+Business Media, LLC 2011.

[14] R. Buyya, T. Cortes, H. Jin : Single system image. Int. J. High Perform. Comput. Appl. 15(2), 124–135 (2001).

[15] Sun (2009a) A Guide to Getting Started with Cloud Computing. Sun white paper. https://www.sun.com/offers/docs/cloud_comp-uting.

[16] Shyam Patidar, Dheeraj Rane, Pritesh Jain : A Survey Paper on Cloud Computing. Second International Conference on Advanced Computing & Communication Technologies 2012.

[17] The Cloud Computing Interoperability Forum (CCIF). Available from: http://www.cloudforum.org/

[18] Jinhua Hu, Jianhua Gu, Guofei Sun, Tianhai Zhao : A Scheduling Strategy on Load Balancing of Virtual Machine Resources in Cloud Computing Environment. 3rd International Symposium on Parallel Architectures, Algorithms and Programming 2010.

[19] Clark C, Fraser K, Hand S, "Live Migration of Virtual Machines[C]," Proceedings of the 2nd Int'l Conference on Networked Systems Design & Implementation. Berkeley, CA, USA, 2005.

[20] Liang Liu, Hao Wang, Xue Liu, Xing Jin, WenBo He, QingBo Wang, Ying Chen: GreenCloud: A New Architecture for Green Data Center. ICAC-INDST'09, June 16, 2009, Barcelona, Spain.

[21] Anton Beloglazov, Rajkumar Buyya "Energy Efficient Resource Management In Virtualized Cloud Data Centers," in 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing, 2010, pp. 826-831.