An Advanced Survey on Cloud Computing and State-of-the-art Research Issues

Mohiuddin Ahmed[1], Abu Sina Md. Raju Chowdhury[2], Mustaq Ahmed[3], Md. Mahmudul Hasan Rafee[4]

**[1,2,3,4] Department of Computer Science and Information Technology, Islamic University of Technology
Board Bazar, Gazipur-1704, Bangladesh**

Abstract

Cloud Computing is considered as one of the emerging arenas of computer science in recent times. It is providing excellent facilities to business entrepreneurs by flexible infrastructure. Although, cloud computing is facilitating the Information Technology industry, the research and development in this arena is yet to be satisfactory. Our contribution in this paper is an advanced survey focusing on cloud computing concept and most advanced research issues. This paper provides a better understanding of the cloud computing and identifies important research issues in this burgeoning area of computer science.

Keywords: *Cloud Computing; Virtualization; Data Center; Server Consolidation.*

1. Introduction

Due to the unprecedented success of internet in last few years, computing resources is now more ubiquitously available. And it enabled the realization of a new computing concept called Cloud Computing. Cloud Computing environment requires the traditional service providers to have two different ways. These are infrastructure and service providers. Infrastructure providers manage cloud platforms and lease resources according to usage. Service providers rent resources from infrastructure providers to serve the end users. Cloud Computing has attracted the giant companies like Google, Microsoft, and Amazon and considered as a great influence in today's Information Technology industry. Business owners are attracted to cloud computing concept because of several features [1, 2].

These are as follows:

- Lower initial investment
- Easier to manage
- Scalability
- Deploy faster
- Location independent
- Device independent
- Reliability
- Security

Although cloud computing has shown considerable opportunities to the IT industry of today's world, but still there are number of challenges that requires to be carefully addressed. In our paper, we present a survey of cloud computing and state-of-the-art research challenges. Our aim is to provide a better understanding of cloud computing and focus on the research ongoing in this tremendously flourishing arena of computer science.

In the section 2, we provide an overview of cloud computing, section 3 contains the state-of-the-art and section 4 will focus on the research issues. We conclude the paper on section 5 along with references.

2. Cloud Computing Overview

2.1 Definitions

What is Cloud Computing?

Cloud computing is a way of leveraging the Internet to consume software or other IT services on demand. Users share processing power, storage space, bandwidth, memory, and software. With cloud computing, the resources are shared and so are the costs. Users can pay as they go and only use what they need at any given time, keeping cost to the user down. Cloud computing is very much a business model as well. Providers of cloud computing solutions, whether they are software, hardware, platform, or storage providers, deliver their offerings over the Internet. There are no shrink wrapped boxes containing discs or hardware for you to buy and set up yourself. Cloud providers typically charge monthly recurring fees based on your usage. [1]

Understanding Cloud Computing Applications [2]

Cloud computing, at its simplest, is a collection of computing software and services available from a decentralized network of servers. The term “cloud” has long been used as a metaphor for the Internet, and there are many popular services and Web sites which you may already be enjoying, without being aware that they are cloud-based. Social networking sites, Web-based email clients like Yahoo! and Gmail, Wikipedia and YouTube, and even peer-to-peer networks like Skype or Bit Torrent are all applications that run in the cloud. In other words, there is no one centralized location or organization that controls them, and nothing is required to utilize them besides a Web browser and an Internet connection. Enterprise cloud computing is cloud computing for the business world. Instead of purchasing and installing the physical infrastructure necessary to run software programs, a business instead consumes resources on a software-as-a-service (SaaS) basis. Running individual applications such as Microsoft, SAP, or Oracle will require hardware and an extensive infrastructure to support it: office space, power, networks, servers, storage, cooling, and bandwidth, not to mention the experts needed to install and run them. Cloud computing offers a streamlined, simplified solution to this complexity and the capital expenditure it necessitates.

Cloud computing [3] according to Wikipedia

Cloud computing is the delivery of computing as a service rather than a product, whereby shared resources, software, and information are provided to computers and other devices as a utility (like the electricity grid) over a network (typically the Internet).

Cloud computing [4] is a technology that uses the internet and central remote servers to maintain data and applications. Cloud computing allows consumers and businesses to use applications without installation and access their personal files at any computer with internet access. This technology allows for much more inefficient computing by centralizing storage, memory, processing and bandwidth.

A simple example of cloud computing is Yahoo email, Gmail, or Hotmail etc. You don’t need software or a server to use them. All a consumer would need is just an internet connection and you can start sending emails. The server and email management software is all on the cloud (internet) and is totally managed by the cloud service provider Yeah, Google etc. The consumer gets to use the software alone and enjoy the benefits. The analogy is, 'If you need milk, would you buy a cow?' All the users or consumers need is to get the benefits of using the software or hardware of the computer like sending emails etc. Just to get this benefit (milk) why should a consumer buy a (cow) software /hardware?

Cloud computing is broken down into three segments: "application" "storage" and "connectivity." Each segment serves a different purpose and offers different products for businesses and individuals around the world. In June 2011, a study conducted by VersionOne found that 91% of senior IT professionals actually don't know what cloud computing is and two-thirds of senior finance professionals are clear by the concept,[5] highlighting the young nature of the technology. In Sept 2011, an Aberdeen Group study found that disciplined companies achieved on average an 68% increase in their IT expense because cloud computing and only a 10% reduction in data center power costs[6].

Final Version of NIST Cloud Computing Definition[7]

Cloud computing is a relatively new business model in the computing world. According to the official NIST definition, "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."

The NIST definition lists five essential characteristics of cloud computing: on-demand self-service, broad network access, resource pooling, rapid elasticity or expansion, and measured service. It also lists three "service models" (software, platform and infrastructure), and four "deployment models" (private, community, public and

hybrid) that together categorize ways to deliver cloud services. The definition is intended to serve as a means for broad comparisons of cloud services and deployment strategies, and to provide a baseline for discussion from what is cloud computing to how to best use cloud computing.

2.2 Types of Cloud

Cloud computing is typically classified in two ways [9]:

1. Location of the cloud computing
2. Type of services offered

Location of the cloud
Cloud computing is typically classified in the following ways:

1. **Public cloud**: In Public cloud the computing infrastructure is hosted by the cloud vendor at the vendor's premises. The customer has no visibility and control over where the computing infrastructure is hosted. The computing infrastructure is shared between any organizations.

2. **Private cloud**: The computing infrastructure is dedicated to a particular organization and not shared with other organizations. Some experts consider that private clouds are not real examples of cloud computing. Private clouds are more expensive and more secure when compared to public clouds.

Private clouds are of two types: On-premise private clouds and externally hosted private clouds. Externally hosted private clouds are also exclusively used by one organization, but are hosted by a third party specializing in cloud infrastructure. Externally hosted private clouds are cheaper than On-premise private clouds.

3. **Hybrid cloud** Organizations may host critical applications on private clouds and applications with relatively less security concerns on the public cloud. The usage of both private and public clouds together is called hybrid cloud. A related term is Cloud Bursting. In Cloud bursting organization use their own computing infrastructure for normal usage, but access the cloud for high/peak load requirements. This ensures that a sudden increase in computing requirement is handled gracefully.

4. **Community cloud** involves sharing of computing infrastructure in between organizations of the same community. For example all Government organizations within the state of California may share computing infrastructure on the cloud to manage data related to citizens residing in California.

Classification based upon service provided

Based upon the services offered, clouds are classified in the following ways:

1. Infrastructure as a service (IaaS) involves offering hardware related services using the principles of cloud computing. These could include some kind of storage services (database or disk storage) or virtual servers. Leading vendors that provide Infrastructure as a service are Amazon EC2, Amazon S3, Rackspace Cloud Servers and Flexiscale.

2. Platform as a Service (PaaS) involves offering a development platform on the cloud. Platforms provided by different vendors are typically not compatible. Typical players in PaaS are Google's Application Engine, Microsofts Azure, Salesforce.com's force.com .

3. Software as a service (SaaS) includes a complete software offering on the cloud. Users can access a software application hosted by the cloud vendor on pay-per-use basis. This is a well-established sector. The pioneer in this field has been Salesforce.coms offering in the online Customer Relationship Management (CRM) space. Other examples are online email providers like Googles gmail and Microsofts hotmail, Google docs and Microsofts online version of office called BPOS (Business Productivity Online Standard Suite).

The above classification is well accepted in the industry. David Linthicum describes a more granular classification on the basis of service provided. These are listed below:

1. Storage-as-a-service
2. Database-as-a-service
3. Information-as-a-service
4. Process-as-a-service
5. Application-as-a-service

6. Platform-as-a-service
7. Integration-as-a-service
8. Security-as-a-service
9. Management/Governance-as-a-service
10. Testing-as-a-service
11. Infrastructure-as-a-service

2.3 Related Technologies

Cloud computing typically has characteristics of all these technologies [8]:

a. Grid computing
b. Virtualization
c. Utility Computing
d. Autonomic Computing

A quick overview of these technologies is given here.

Grid Computing

Grid Computing involves a network of computers that are utilized together to gain large supercomputing type computing resources. Using this network of computers large and complex computing operations can be performed. In grid computing these network of computers may be present in different locations.
A famous Grid Computing project is Folding@Home. The project involves utilizing unused computing powers of thousands of computers to perform a complex scientific problem. The goal of the project is "to understand protein folding, misfolding, and related diseases".

Virtualization

Virtualization introduces a layer between Hardware and operating system. During the sixties mainframe started supporting many users using virtual machines. These virtual machines simulated behavior of an operating system for each user. VMWare launched a product called VMware Workstation in 1999 that allows multiple operating systems to run on personal computers.
The virtualization forms the foundation of cloud technology. Using virtualization, users can access servers or storage without knowing specific server or storage details. The virtualization layer will execute user request for computing resources by accessing appropriate resources.
Typically server utilization in data centers can be as low as 10%. Virtualization can help in significantly improving server utilization.

Utility Computing

Utility Computing defines a "pay-per-use" model for using computing services. In utility computing, billing model of computing resources is similar to how utilities like electricity are traditionally billed. When we procure electricity from a vendor, the initial cost required is minimal. Based upon the usage of electricity, electricity companies bills the customer (typically monthly). In utility computing billing is done using a similar protocol.
Various billing models are being explored. A few common ones are:

1. Billing per user count. As an example if an organization of 100 people uses Google's gmail or Microsoft Live as their internal email system with email residing on servers in the cloud, Google/Microsoft may bill the organization on per user basis.
2. Billing per Gigabyte. If an organization is using Amazon to host their data on the cloud, Amazon may bill the organization on the disk space usage.
3. Billing per hour/day. As an example a user may pay for usage of virtual servers by time utilized in hours.

3. State-of-the-art Technologies

In this section, the state-of-the-art implementations of cloud computing is presented. Technologies used for cloud computing are describes here.

3.1 Architectural Design of Data Centre

A **data center [10]** (or **data centre** or **datacenter** or **datacenter**) is a facility used to house computer systems and associated components, such as telecommunications and storage systems. It generally includes redundant or backup power supplies, redundant data communications connections, environmental controls (e.g., air conditioning, fire suppression) and security devices.

Key Design Areas

- Resilience - ensuring maximum uptime without compromising on performance.

- Availability - business continuity is especially important.
- Performance - the faster the better; in a predictable manner.
- Security - ensuring data separation and controlled access to all Data centre resources.
- Effective architecture for data separation - a common infrastructure that provides facilities for network-based backup and efficient back-end network access.
- Predictable Failover - for maximum service availability.

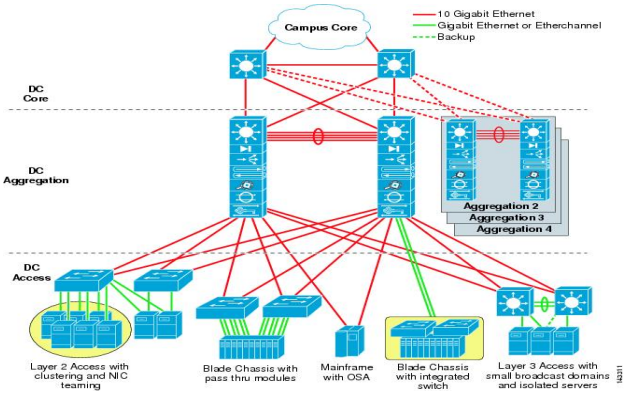

Figure : Data Centre Design

3.2 Distributed File System over Cloud

This In this paper [11], they focused on Google File System (GFS)[12] which is a proprietary distributed file system developed by Google and specially designed to provide efficient, reliable access to data using large clusters of commodity servers. Files are divided into chunks of 64 megabytes and are usually appended to a read and only extremely rarely overwritten or shrunk. Comparing traditional file systems, GFS is designed and optimized to run on data centers to provide extremely high data throughputs, low latency and survive individual server failures. Inspired by GFS, the open source Hadoop Distributed File System (HDFS)[13] stores large files across multiple machines. It achieves reliability by replacing the data across multiple servers.

3.3 Distributed Application Framework over Clouds

MapReduce[14] is a software framework introduced by Google to support distributed computing on large data sets on clusters of computers. HTTP-based applications are usually conform to some web application framework such as Java EE. MapReduce consists of one Master to which client applications submit MapReduce jobs. The Master pushes work out to available task nodes in the data centre striving to keep the tasks as close to the data as possible. The open source Hadoop MapReduce project[15] is inspired by Google's work. Currently, many organizations are using Hadoop MapReduce to run large data intensive computations.

4. Research Issues

Cloud Computing has been widely practiced by IT industry as well as business enterprises in recent times. But , research on it is still at an immature stage. Many existing issues have not been fully addressed, where as newer problems are arising due to its extensive usage. In this section, we summarize some of the crucial research issues in cloud computing.

4.1 Virtual Machine Migration

Some vendors have implemented VM migration in their virtualization solution—a big advantage for application uptime in a data center[16]. What is VM migration? Consider the case of a server with a hypervisor and several VMs, each running an OS and applications. If you need to bring down the server for maintenance (say, adding more memory to the server), you have to shut down the software components and restart them after the maintenance window—significantly affecting application availability. VM migration allows you to move an entire VM (with its contained operating system and applications) from one machine to another and continue operation of the VM on the second machine. This advantage is unique to virtualized environments because you can take down physical servers for maintenance with minimal effect on running applications. You can perform this migration after suspending the VM on the source machine, moving its attendant information to the target machine and starting it on the target machine. To lower the downtime, you can perform this migration while the VM is running (hence the name "live migration") and resuming its operation on the target machine after all the state is migrated. The following are some of the benefits of virtualization in a cloud-computing environment:

- *Elasticity and scalability:* Firing up and shutting down VMs involves less effort as opposed to bringing servers up or down.

- *Workload migration:* Through facilities such as live VM migration, you can carry out workload migration with much less effort as compared to

workload migration across physical servers at different locations.

- *Resiliency:* You can isolate physical-server failure from user services through migration of VMs.

It must be clarified that virtualization is not a prerequisite for cloud computing. In fact, there are examples of large cloud service providers using only commodity hardware servers (with no virtualization) to realize their infrastructure. However, virtualization provides a valuable toolkit and enables significant flexibility in cloud-computing deployments.

4.2 Server Consolidation

Server consolidation is an effective approach to maximize resource utilization while minimizing energy consumption in a cloud computing environment. Live VMmigration technology is often used to consolidate VMs residing on multiple under-utilized servers onto a single server, so that the remaining servers can be set to an energy-saving state. The problem of optimally consolidating servers in a data center is often formulated as a variant of the vector bin-packing problem [17], which is an NP-hard optimization problem. Various heuristics have been proposed for this problem [18,19]. Additionally, dependencies among VMs, such as communication requirements, have also been considered recently [20].

4.3 Energy Management

Improving energy efficiency is another major issue in cloud computing. It has been estimated that the cost of powering and cooling accounts for 53% of the total operational expenditure of data centers [21]. In 2006, data centers in the US consumed more than 1.5% of the total energy generated in that year, and the percentage is projected to grow 18% annually [18]. Hence infrastructure providers are under enormous pressure to reduce energy consumption. The goal is not only to cut down energy cost in data centers, but also to meet government regulations and environmental standards. Designing energy-efficient data centers has recently received considerable attention. This problem can be approached from several directions. For example, energy efficient hardware architecture that enables slowing down CPU speeds and turning off partial hardware components has become commonplace. Energy-aware job scheduling and server consolidation are two other ways to reduce power consumption by turning off unused machines. Recent research has also begun to study energy-efficient network protocols and infrastructures . A key challenge in all the above methods is to achieve a good trade-off between energy savings and application performance. In this respect, few researchers have recently started to investigate coordinated solutions for performance and power management in a dynamic cloud environment.

4.4 Information Security

Information security is another important research topic in cloud computing. Since service providers typically do not have access to the physical security system of data centers, they must rely on the infrastructure provider to achieve full data security. Even for a virtual private cloud, the service provider can only specify the security setting remotely, without knowing whether it is fully implemented. The infrastructure provider, in this context, must achieve the following objectives: (1) *confidentiality*, for secure data access and transfer, and (2) *audit ability*, for attesting whether security setting of applications has been tampered or not. Confidentiality is usually achieved using cryptographic protocols, whereas audit ability can be achieved using remote attestation techniques. Remote attestation typically requires a trusted platform module (TPM) to generate non-forgeable system summary (i.e. system state encrypted using TPM's private key) as the proof of system security. However, in a virtualized environment like the clouds, VMs can dynamically migrate from one location to another, hence directly using remote attestation is not sufficient. In this case, it is critical to build trust mechanisms at every architectural layer of the cloud. Firstly, the hardware layer must be trusted using hardware TPM. Secondly, the virtualization platform must be trusted using secure virtual machine monitors. VM migration should only be allowed if both source and destination servers are trusted. Recent work has been devoted to designing efficient protocols for trust establishment and management.

4.5 Novel Cloud Architecture

In recent times, commercial clouds are implemented in mammoth data centers and operated centrally. Although, there are economical pros but it too comes along with limitations like high energy expenses and initial investment for constructing data centers. In this paper[22], small sized data centers are suggested to be more

advantageous than large data centers. Small sized data centers eliminates the problems of energy,power,cooling system and, more over economical & geographically better distributed.

Another research trend on using voluntary resources for hosting cloud application. Voluntary resources help to build clouds which are cheaper and more suitable for non-profit applications such as scientific computing. Although it is providing some benefits but management of such heterogeneous resources is a challenge and creating such architecture is a research problem.

5. Conclusion

The advancement of cloud computing is dramatically changing the horizon of information technology and ultimately turns the utility computing into a reality. However, it provides a large array of benefits, but many challenges in this domain, including automatic resource positioning, energy management, information security are only attracted the research community. There are still so many issues to be explored. Opportunities are enough in this arena for some groundbreaking contribution and bring significant development in the industry.

In our paper, we have presented an overview of cloud computing and focused on the state-of-the-art research and future issues to be handled by the research community. Cloud computing is at an early stage of research and development, we believe our paper will provide a better understanding of the cloud computing and different research issues, thereby bolstering further research in this arena.

References

[1] http://www.thesmallbusiness.org/software/benefits-of-cloud-computing.html

[2] http://www.allthingscrm.com/cloud-computing/understanding-cloud-computing-applications.html

[3] http://en.wikipedia.org/wiki/Cloud_computing

[4] http://www.wikinvest.com/concept/Cloud_Computing

[5] Cloud Confusion Amongst IT Professionals. VersionOne (June 6, 2011).

[6] Business Adoption of Cloud Computing. AberdeenGroup (Sept 9, 2011).

[7] http://www.nist.gov/itl/csd/cloud-102511.cfm

[8] http://thecloudtutorial.com/related.html

[9] http://thecloudtutorial.com/cloudtypes.html

[10] http://en.wikipedia.org/wiki/Data_center

[11] Qi Zhang,Lu Cheng, Raouf Boutaba."Cloud Computing: State-of-the-art and research challenges". The Brazilian Computer Society 2010.

[12] GhemawatS, Gobioff H, Leung S-T(2003)
The Google file system. In: Proc of SOSP, October 2003.

[13] Hadoop Distributed File System, hadoop.apache.org/hdfs

[14] Dean J, Ghemawat S(2004) MapReduce:simplified data processing on large clusters. In: Proc of OSDI

[15] Hadoop MapReduce, hadoop.apache.org/mapreduce

[16] http://www.cisco.com/web/about/ac123/ac147/archived_issues/ipj_12-3/123_cloud1.html

[17] Chekuri C, Khanna S (2004) On multi-dimensional packing problems.
SIAM J Comput 33(4):837–851

[18]Li B et al (2009) EnaCloud: an energy-saving application live placement approach for cloud computing environments. In: Proc of international conf on cloud computing

[19] Srikantaiah S et al (2008) Energy aware consolidation for cloud computing. In: Proc of HotPower

[20] Meng X et al (2010) Improving the scalability of data center networks
with traffic-aware virtual machine placement. In: Proc INFOCOM

[21] Hamilton J (2009) Cooperative expendable micro-slice servers (CEMS): low cost, low power servers for Internet-scale services
In: Proc of CIDR

[22] Valancius V, Laoutaris N et al (2009) Greening the Internet with nano data centers. In: Proc of CoNext

Mohiuddin Ahmed has completed Bachelor of Science in Computer Science and Information Technology from Islamic University of Technology, OIC. Research interest includes Human Computer Interaction, Cloud Computing, Artificial Intelligence and Wireless Network.

Abu Sina Md. Raju Chowdhury has completed Bachelor of Science in Computer Science and Information Technology from Islamic University of Technology, OIC. Research interest includes Cloud Computing, Human Computer Interaction, Peer to Peer and Overlay Network, Data Mining and Web Engineering.

Mustaq Ahmed has completed Bachelor of Science in Computer Science and Information Technology from Islamic University of Technology, OIC. Research interest includes Cloud Computing, Human Computer Interaction, Peer to Peer and Overlay Network, Open Source Computing.

Md. Mahmudul Hasan Rafee has completed Bachelor of Science in Computer Science and Information Technology from Islamic University of Technology, OIC. Research interest includes Cloud Computing, Human Computer Interaction, Web Engineering.