

Cloud Computing: Concepts and Trends

Engr. Tariq Haleem

Abstract—This paper presents an introduction to cloud computing and its applications. Cloud computing is a model where users can gain access to their applications from anywhere through their connected devices. A survey of its architecture and latest trends is also presented

Index Terms—cloud computing, architecture

I. INTRODUCTION

ALTHOUGH cloud computing is an emerging field of computer science, the idea has been around for a few years. The trend toward cloud computing started in the late 1980s with the concepts of grid computing when, for the first time, a large number of systems were applied to a single problem, usually scientific methods. The primary difference between grid computing and cloud computing is how each provides the needed resources to a workload? It's called cloud computing because the data and applications exist on a "cloud" of Web servers. In a cloud computing system, there's a significant workload shift. Local computers no longer have to do all the heavy lifting when it comes to running applications. The network of computers that make up the cloud handles them instead. Hardware and software demands on the user's side decrease. The only thing the user's computer needs to be able to run is the cloud computing system's interface software, which can be as simple as a Web browser, and the cloud's network takes care of the rest.

Cloud computing is a model where users can gain access to their applications from anywhere through their connected devices. A simplified user interface makes the infrastructure supporting the applications transparent to users. The applications reside in massively-scalable data centers where compute resources can be dynamically provisioned and shared to achieve significant economies of scale. A strong service management platform results in near-zero incremental management costs when more IT resources are added to the cloud. The proliferation of smart mobile devices, high speed wireless connectivity, and rich browser-based Web 2.0 interfaces has made the network-based cloud computing model not only practical but also a source of reduced IT complexity.

Cloud computing comes into focus only when we think about what IT always needs: a way to increase capacity or add capabilities on the fly without investing in new infrastructure, training new personnel, or licensing new software. Cloud computing encompasses any subscription-based or pay-per-use service that, in real time over the Internet, extends IT's existing capabilities.

The term Cloud Computing derives from the common depiction in most technology architecture diagrams, of the Internet or IP

availability, using an illustration of a cloud. The computing resources being accessed are typically owned and operated by a third-party provider on a consolidated basis in Data Center locations. Target consumers are not concerned with the underlying technologies used to achieve the increase in server capability, and is sold simply as a service available on demand. Grid computing is a technology approach to managing a cloud. In effect, all clouds are managed by a grid but not all grids manage a cloud. More specifically, a compute grid and a cloud are synonymous, while a data grid and a cloud can be different. Cloud computing describes both a platform and a type of application. A cloud computing platform dynamically provisions, configures, reconfigures, and de provisions servers as needed. Cloud applications are those that are extended to be accessible through the Internet. These cloud applications use large data centers and powerful servers that host Web applications and Web services.

Cloud computing is an emerging approach to shared infrastructure in which large pools of systems are linked together to provide IT services. Cloud Computing will allow corporate data centers to operate more like the Internet by enabling computing across a distributed, globally accessible fabric of resources, rather than on local machines or remote server systems. Organizations can use them as much as they want and as wireless broadband connection options grow, wherever they need them. It describes how computer programs are hosted and operated over the Internet. The key feature of cloud computing is that both the software and the information held in it live on centrally located servers rather than on an end-user's computer. "This means people can access the information that they need from any device with an Internet connection—including mobile and handheld phones—rather than being chained to the desktop. It also means lower costs, since there is no need to install software or hardware." The architecture behind cloud computing is a massive network of 'cloud servers' interconnected as if in a grid running in parallel, sometimes using the technique of virtualization to maximize compute power per server.

The term cloud computing is used for both the platform and the type of application. Cloud computing makes use of servers, physical or virtual, dynamically when the need arises. This type of computing even makes use of other resources such as SANs, network equipment, and security devices. It can also support applications that are accessible through the Internet. These applications make use of large data centers and powerful servers that host Web applications and Web services.

II. ARCHITECTURE

A cloud computing system, can be divided into two sections: the front end and the back end. They connect to each other through a network, usually the Internet. The front end is the side the computer user, or client, sees. The back end is the "cloud" section of the system. The front end includes the client's computer (or computer network) and the application required to access the cloud computing system.

A. A typical Cloud Computing System

Not all cloud computing systems have the same user interface. Most of the time, servers don't run at full capacity. That means there's unused processing power going to waste. It's possible to fool a physical server into thinking it's actually multiple servers, each running with its own independent operating system. The technique is called server virtualization. By maximizing the output of individual servers, server virtualization reduces the need for more physical machines. On the back end of the system are the various computers, servers and data storage systems that create the "cloud" of computing services. In theory, a cloud computing system could include practically any computer program you can imagine, from data processing to video games. Usually, each application will have its own dedicated server. A central server administers the system, monitoring traffic and client demands to ensure everything runs smoothly. It follows a set of rules called protocols and uses a special kind of software called middleware. Middleware allows networked computers to communicate with each other.

A typical high-level architecture of a cloud computing platform consists of a data center, provisioning manager, monitoring server, application server, and virtualization components. A typical cloud computing system is shown in Figure 1.

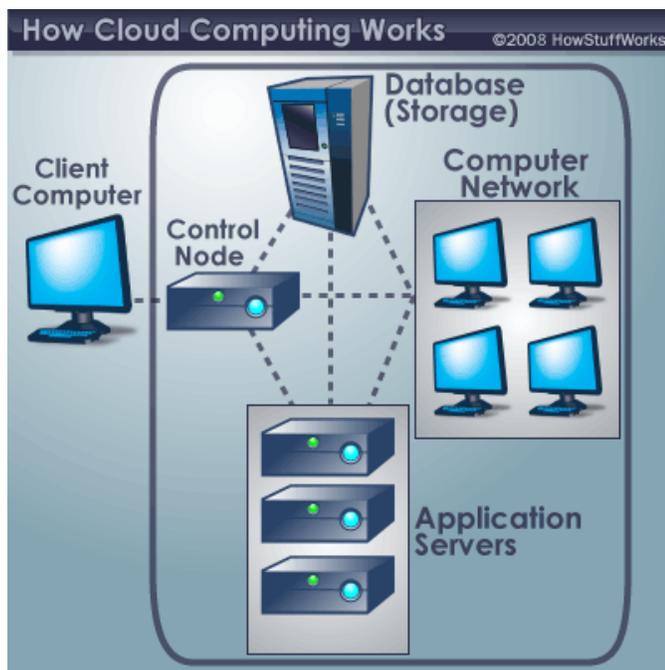


Figure 1 A typical cloud computing system

B. Architecture of Cloud Computing Platform

The architecture makes use of the following components.

User interaction interface: This is how users of the cloud interface with the cloud to request services.

Services catalog: This is the list of services that a user can request.

System management: This is the piece which manages the computer resources available.

Provisioning tool: This tool carves out the systems from the cloud to deliver on the requested service. It may also deploy the required images.

Monitoring and metering: This optional piece tracks the usage of the cloud so the resources used can be attributed to a certain user.

Data Servers: The servers are managed by the system management tool. They can be either virtual or real.

These components serve as the back-end of the cloud computing environment. Apart from this backend architecture, it even has a couple of user interfaces to provision servers, one fully loaded, and other for basic provisioning requests. All requests are handled by Web 2.0 components deployed on the application server.

In grid computing, the focus is on the ability of moving a workload to the location of the needed computing resources, which are mostly remote and are readily available for use. Usually a grid is a cluster of servers on which a large task could be divided into smaller tasks to run in parallel. From this point of view, a grid could actually be viewed as just one virtual server? In a cloud environment, computing resources, such as servers, can be dynamically shaped or carved out from its underlying hardware infrastructure and made available to a workload. In addition, while a cloud does support grid, a cloud can also support non grid environments, such as a three-tier Web architecture running traditional or Web 2.0 applications. The concept of cloud computing has evolved from the concepts of grid, utility, and SaaS. It is an emerging model where users can gain access to their applications from anywhere through their connected devices. These applications reside in massively-scalable data centers where compute resources can be dynamically provisioned and shared to achieve significant economies of scale. The proliferation of smart mobile devices, high speed wireless connectivity, and rich browser based Web 2.0 interfaces has made the network-based cloud computing model not only practical but also a source of reduced IT complexity. The strength of a cloud is its infrastructure management, enabled by the maturity and progress of virtualization technology to manage and better utilize the underlying resources through automatic provisioning, re imaging, workload rebalancing, monitoring, and systematic change request handling.

If a cloud computing company has a lot of clients, there's likely to be a high demand for a lot of storage space. Some companies require hundreds of digital storage devices. Cloud computing systems need at least twice the number of storage devices it requires to keep all its clients' information stored. That's because these devices, like all computers, occasionally break down. A cloud computing system must make a copy of all its clients' information and store it on other devices. The copies enable the central server to access backup machines to retrieve data that otherwise would be unreachable. Making copies of data as a backup is called redundancy.

Cloud storage refers to saving data to an off-site storage system maintained by a third party. Instead of storing information to your computer's hard drive or other local storage device, you save it to a remote database. The Internet provides the connection between your computer and the database.

On the surface, cloud storage has several advantages over traditional data storage. Ensure if you store your data on a cloud storage system, you'll be able to get to that data from any location that has Internet access. You wouldn't need to carry around a physical

Storage device or use the same computer to save and retrieve your information. With the right storage system, you could even allow other people to access the data, turning a personal project into a collaborative effort.

Cloud computing systems offer users access to not only storage, but also processing power and computer applications installed on a remote network. There are hundreds of different cloud storage systems. Some have a very specific focus, such as storing Web e-mail messages or digital pictures. Others are available to store all forms of digital data. Some cloud storage systems are small operations, while others are so large that the physical equipment can fill up an entire warehouse. The facilities that house cloud storage systems are called data center. Creating an effective data center requires careful planning. The three big concerns every data center must be able to address are security, electric power and cooling: Physical security is just as important as network security. Data servers are valuable not only because the machines themselves are expensive, but also because the data stored on them could include sensitive information. A single data server's power requirements aren't very taxing. But when a data center has hundreds of servers, it's crucial that the center's electric wiring can support the workload. Like all computers, data servers generate heat. Too much heat can impair or damage servers, so the data center needs an effective cooling system to prevent such problems. Cloud storage systems generally rely on hundreds of data servers.

Without redundancy, a cloud storage system couldn't ensure clients that they could access their information at any given time. Most systems store the same data on servers that use different power supplies. That way, clients can access their data even if one power supply fails.

For all the talk about cloud computing becoming a persuasive force it relies on one very simple thing - ubiquitous global broadband access. Without this global access for everyone regardless of their economic place in society this move to the cloud on a large scale could end up widening the technological divide that already exists.

Cloud computing changes the economics of IT by enabling you to pay only for the capacity that you are actually using. Compute capacity can be scaled up or down on demand to accommodate changing workloads and business requirements.

Taken to the next step, this implies that cloud computing infrastructures, and specifically their middleware and application platforms, should ideally have these characteristics:

Self-healing: In case of failure, there will be a hot backup instance of the application ready to take over without disruption (known as failover). It also means that when a policy is set that says everything should always have a backup, when such a fail occurs and backup becomes the primary, the system launches a new backup, maintaining reliability policies.

SLA-driven: The system is dynamically managed by service-level agreements that define policies such as how quickly responses to requests need to be delivered. If the system is experiencing peaks

in load, it will create additional instances of the application on more servers in order to comply with the committed service levels — even at the expense of a low-priority application.

Multi-tenancy: The system is built in a way that allows several customers to share infrastructure, without the customers being aware of it and without compromising the privacy and security of each customer's data.

Service-oriented: The system allows composing applications out of discrete services that are loosely coupled (independent of each other). Changes to or failure of one service will not disrupt other services. It also means services can be reused.

Virtualized: Applications are decoupled from the underlying hardware. Multiple applications can run on one computer or multiple computers can be used to run one application (grid computing).

Linearly Scalable: Perhaps the biggest challenge. The system will be predictable and efficient in growing the application. If one server can process 1,000 transactions per second, two servers should be able to process 2,000 transactions per second, and so forth.

C. Evolution of Cloud Computing

In a grid computing system, networked computers are able to access and use the resources of every other computer on the network. In cloud computing systems, that usually only applies to the back end. Utility computing is a business model where one company pays another company for access to computer applications or data storage. Utility computing relates to the business model in which application infrastructure resources — hardware and/or software — are delivered. While cloud computing relates to the way we design, build, deploy and run applications that operate in an a virtualized environment, sharing resources and boasting the ability to dynamically grow, shrink and self-heal Cloud computing is a broader concept than utility computing and relates to the underlying architecture in which the services are designed. It may be applied equally to utility services and internal corporate data centers. Grid computing is a term for either of two broad subcategories of distributed computing:

D. Cloud Computing Concerns

With cloud computing, we are at a point similar to when people started realizing that their money was safer in a bank than under their mattress. Some people envision a future in which the entire Web becomes a massive storage cloud. People will constantly upload and download data to and from the cloud. The concept of data ownership will lose its meaning. Everyone will have access to everything and secrecy will cease to exist. Other people dismiss these ideas and say that cloud storage is just a tool like any other. An issue that information experts, computer scientists and entrepreneurs debate is the concept of data ownership. Who Owns the data stored in a cloud system? Does it belong to the client who originally saved the data to the hardware? Does it belong to the company that owns the physical equipment storing the data? What happens if a client goes out of business? Can a cloud storage host delete the former client's data?

The biggest concerns about cloud storage are reliability, privacy and security. The idea of handing over important data to another company worries some people. Corporate executives might hesitate to take advantage of a cloud computing system because they can't keep their

company's information under lock and key. Clients aren't likely to entrust their data to another company without a guarantee that they'll be able to access their information whenever they want and no one else will be able to get at it.

A typical cloud storage system architecture includes a master control server and several storage servers. To secure data, most systems use a combination of techniques, including:

Encryption: which means they use a complex algorithm to encode information. To decode the encrypted files, a user needs the encryption key. While it's possible to crack encrypted information, most hackers don't have access to the amount of computer processing power they would need to decrypt information.

Authentication processes: which require to create a user name and password.

Authorization practices: The client lists the people who are authorized to access information stored on the cloud system. Many corporations have multiple levels of authorization. Even with these protective measures in place, many people worry that data saved on a remote storage system is vulnerable. There's always the possibility that a hacker will find an electronic back door and access data. Hackers could also attempt to steal the physical machines on which data are stored. A disgruntled employee could alter or destroy data using his or her authenticated user name and password. Cloud storage companies invest a lot of money in security measures in order to limit the possibility of data theft or corruption.

It's in each company's best interests to provide the most secure and reliable service possible. If a company can't meet these basic client expectations, it doesn't have much of a chance -- there are too many other options available on the market.

E. Advantages

Potential advantages of any cloud computing approach include:

- Location of infrastructure in areas with lower costs of real estate and electricity.
- Sharing of peak-load capacity among a large pool of users, improving overall utilization.
- Separation of infrastructure maintenance duties from domain-specific application development ,
- Separation of application code from physical resources.
- Ability to use external assets to handle peak loads.
- Not have to purchase assets for one-time or infrequent intensive computing tasks.

Storing data in the cloud already has some distinct advantages over client-based access. We can leverage the sheer processing power of the cloud to do things that traditional productivity applications cannot do. "For instance, users can instantly search over 25 GB worth of e-mail online, which is nearly impossible to do on a desktop. To take another example, each document created through Google Apps is easily turned into a living information source, capable of pulling the latest data from external applications, databases and the Web. This revolutionizes processes as simple as creating a Google spreadsheet to compare stock prices from vendors over time. Why would anyone want to rely on another computer system to run programs and store data? Here are just a few reasons:

Clients would be able to access their applications and data from anywhere at any time. They could access the cloud computing system using any computer linked to the Internet. Data wouldn't be confined to a hard drive on one user's computer or even a corporation's internal network.

It could bring hardware costs down. Cloud computing systems would reduce the need for advanced hardware on the client side. You wouldn't need to buy the fastest computer with the most memory, because the cloud system would take care of those needs for you. Instead, you could buy an inexpensive computer terminal. The terminal could include a monitor, input devices like a keyboard and mouse and just enough processing power to run the middleware necessary to connect to the cloud system. You wouldn't need a large hard drive because you'd store all your information on a remote computer.

Corporations that rely on computers have to make sure they have the right software in place to achieve goals. Cloud computing systems give these organizations company-wide access to computer applications. The companies don't have to buy a set of software or software licenses for every employee. Instead, the company could pay a metered fee to a cloud computing company.

Servers and digital storage devices take up space. Some companies rent physical space to store servers and databases because they don't have it available on site. Cloud computing gives these companies the option of storing data on someone else's hardware, removing the need for physical space on the front end. Corporations might save money on IT support. Streamlined hardware would, in theory, have fewer problems than a network of heterogeneous machines and operating systems.

If the cloud computing system's back end is a grid computing system, then the client could take advantage of the entire network's processing power. Often, scientists and researchers work with calculations so complex that it would take years for individual computers to complete them. On a grid computing system, the client could send the calculation to the cloud for processing. The cloud system would tap into the processing power of all available computers on the back end, significantly speeding up the calculation. By reducing the traditional costs and labor associated with deploying, maintaining and upgrading business technology, IT departments are increasingly becoming free to devote their limited resources to projects more strategic to the business. And since software lives in the cloud, it can be improved as often as needed without tying up the IT department or inconveniencing users. This "version less" software eliminates upgrade projects and helps technology keep pace with the speed of business, giving employees access to new technology early and often rather than forcing them to wait for a final, packaged product to be shipped.

F. Current Implementations

IBM is the most recent company to announce plans to tap cloud computing technologies. IBM executives in Shanghai unveiled a system, dubbed Blue Cloud, that will let banks and other customers distribute their programs across large numbers of machines to deliver faster, more sophisticated data analysis. The first Blue Cloud products were introduced in the spring of 2008. Google's search engine and productivity applications are among the early products of efforts to locate processing power on vast banks of computer servers, rather than on desktop PCs. Microsoft has released online software called Windows Live for photo-sharing, file storage, and other

applications served from new data centers. Yahoo has taken similar steps. Amazon.com (AMZN) recently broadened access for software developers to its "Elastic Compute Cloud" service, which lets small software companies pay for processing power streamed from Amazon's data centers. It enables you to increase or decrease capacity within minutes, not hours or days. You can commission one, hundreds or even thousands of server instances simultaneously

Web e-mail providers like G mail, Hotmail and Yahoo! Mail store e-mail messages on their own servers. Users can access their e-mail from computers and other devices connected to the Internet.

Sites like Flickr and Picasa host millions of digital photographs. YouTube hosts millions of user-uploaded video files. Web site hosting companies like Start Logic, Host monster and Go Daddy store the files and data for client Web sites. Social networking sites like Facebook and My Space allow members to post pictures and other content. All of that content is stored on the respective site's servers. Services like X drive, Media Max and Strong space offer storage space for any kind of digital data.

III. CONCLUSION

Cloud computing is particularly valuable to small and medium

businesses, where effective and affordable IT tools are critical to helping them become more productive without spending lots of money on in-house resources and technical equipment. "But we are seeing large businesses moving to the cloud as well, for a variety of reasons, such as cost savings, remote access, ease of availability and real-time collaboration capabilities. Cloud computing can be used in a variety of areas including internal pilots, hosting virtual worlds, e-business, social networks, and search. In the future, cloud computing would be able to solve many management related issues of organizations and help make the best use of resources available within an organization.

REFERENCES

- [1] Web Service for Estimating Capacity and Performance of Service-Oriented Architecture (SOA) Workloads, October 2006 at www.software.ibm.com/software/dw/wes/hipods/SONOM
- [2] www.ibm.com/developerworks/websphere/zones/hipods/library.html
- [3] IBM's Vision For The New Enterprise Data Center, March 2008 at www.ibm.com/systems/optimizeit/datacenter/pdf/NEDC_POV_MAR_2008_-_02.pdf
- [4] Rajkumar Buyya, Chee Shin Yeo, and Srikumar Venugopal, "Market-Oriented Cloud Computing," *Proceedings of the 10th IEEE International Conference on High Performance Computing and Communications (HPCC-08, IEEE Computer Society Press, Los Alamitos, CA, USA), Sept. 25-27, 2008, Dalian, China.*