Chapter I: Cloud Computing in Brief

I.1. Introduction

Nowadays with the advancement of the modern human life, daily needs are commonly provided in that way so everybody can use them. Utility services, such as water, electricity, gas, and telephony are deemed necessary for fulfilling daily life routines. These utility services are accessed so frequently that they need to be available whenever the consumer requires them at any time. Consumers are then able to pay service providers based on their usage of these utility services.

Therefore something similar is happening today in the world of computing. Data and programs are being swept up from desktop PCs and corporate server rooms and installed in the compute cloud. Many companies have a high dependence for the use of computer technology in order to boost their performance. Companies that adapt the newest technology may have better chance to win the competition between other companies.

As software move from PCs to Internet servers, users and developers alike go along for Cloud computing. Some substantial fraction of computing activity is migrating away from the desktop and the corporate server room. This change will affect all levels of the computational things, from casual user to software developer, IT manager, even hardware manufacturer.

However Cloud computing is often confused with grid computing, utility computing, and autonomic computing. Today's cloud computing is powered by grids. Cloud computing is a general concept that incorporates software as a service (SaaS), Web 2.0 and other technologies which have purpose to fulfil customer's computing needs. Cloud computing emerges as a new computing paradigm which aims to provide reliable, customized and QoS guaranteed computing dynamic environments for end-users. This paper reviews recent

advances of Cloud computing, identifies the concepts and characters of scientific Clouds, and finally presents an example of scientific Cloud for data centres.

I.1.1. History of Cloud Computing

Technically, cloud computing has been around for a while. The Web itself is a form of cloud computing. Users access web pages that are in the cloud. Over time, speed and reliability on the Internet improved. Improvements in computer processing and connection speed have made the Web an increasingly interactive experience. An increased number of users were able to contribute to the Web's content. When Hotmail began to make web-based email popular in the 1990's it foreshadowed such services.

I.2. Definition of Cloud Computing

First of all, we need to know the definition for the Cloud first:

"A Cloud is a type of parallel and distributed system consisting of a collection of interconnected and virtualised 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."[1]

Based on the definition above, we can know that Cloud Computing is where computing resources are accessed from an internet rather than a local desktop or data centre.[2] Cloud computing is a rapidly growing trend and is highly interlinked with the development of Web 2.0. This said, it needs to be

---

[1] Rajkumar Buyya, Chee Sin Yeo, and Srikumar Venugopal, "Market-Oriented Cloud Computing: Vision, Hype, and Reality for Delivering IT Services as Computing Utilities", *IEEE Distributed Systems Online*

[2] Christopher Barnatt, "Cloud Computing", *Explaining Computers Website,* http://www.explainingcomputers.com/cloud.html, retrieved November 28, 2008

appreciated from the outset that conceptually at least cloud computing and Web 2.0 are quite distinct. For specific, the key concept of Web 2.0 is making new forms of online connection between people, services and applications; at the same time as the key concept of cloud computing is the detachment of computing resources from any even notional location.

Whether it's called cloud computing or on-demand computing, software as a service, or the internet as platform, the common element is a shift in the geography of computation. For example when you create a spreadsheet with the Google Docs service, major components of the software reside on unseen computers possibly spread across continents.

Software for major business applications (such as customer support, sales, and marketing) has generally been run on corporate servers, but several companies now provide it as an on-demand service[3]. The first was Salesforce.com, founded in 1999, offering a suite of online programs for customer relationship management and other business oriented tasks; the company's slogan is "No software!"

[3] Bryan Hayes, "Cloud Computing", *IEEE Distributed Systems Online,* 2008

(Cloud Computing)

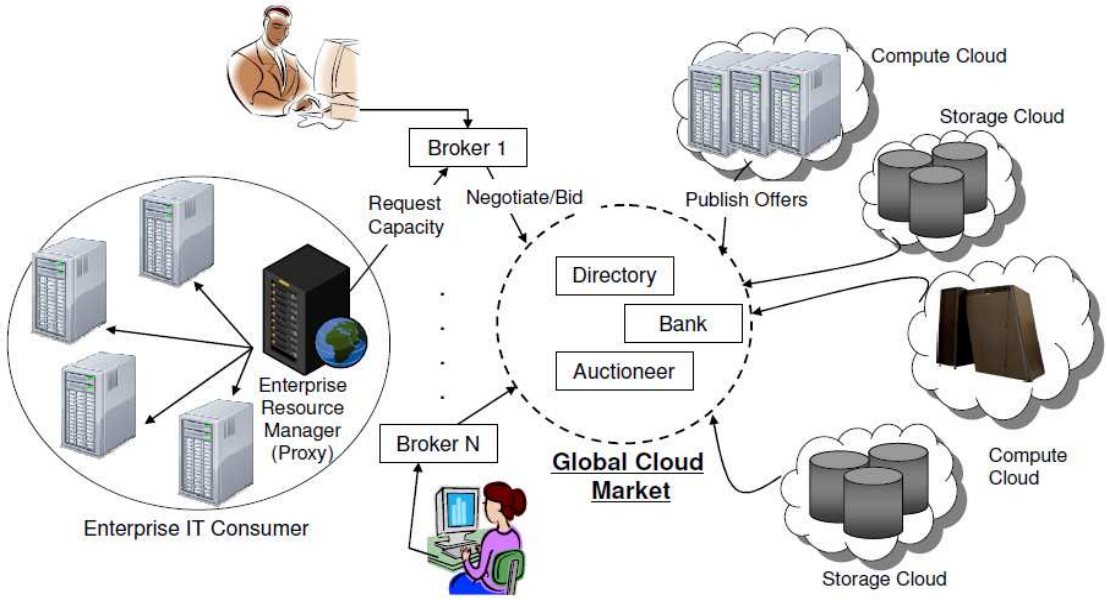

Figure I.1. Global cloud exchange and market infrastructure for trading services[4] (Market-Oriented Cloud Computing)

[4] Rajkumar Buyya, Chee Sin Yeo, and Srikumar Venugopal, "Market-Oriented Cloud Computing: Vision, Hype, and Reality for Delivering IT Services as Computing Utilities", *IEEE Distributed Systems Online,* 2008

I.3. Current State

The Web search popularity, as measured by the Google search trends during the last 12 months, for terms “cluster computing”, “Grid computing”, and “Cloud computing”, it can be observed that cluster computing was a popular term during 1990s, from early 2000 Grid computing become popular, and recently Cloud computing started gaining popularity. Here’s the current state which related to Cloud computing[5]:

- IBM Introduces 'Blue Cloud' Computing, CIO Today - Nov 15 2007
- IBM, EU Launch RESERVOIR Research Initiative for Cloud Computing, IT News Online - Feb 7 2008
- Google and Salesforce.com in Cloud computing deal, Siliconrepublic.com - Apr 14 2008
- Demystifying Cloud Computing, Intelligent Enterprise - Jun 11 2008
- Yahoo realigns to support Cloud computing, 'core strategies', San Antonio Business Journal - Jun 27 2008
- Merrill Lynch Estimates "Cloud Computing" To Be $100 Billion Market, SYS-CON Media - Jul 8 2008

[5] Rajkumar Buyya, Chee Sin Yeo, and Srikumar Venugopal, “Market-Oriented Cloud Computing: Vision, Hype, and Reality for Delivering IT Services as Computing Utilities”, *IEEE Distributed Systems Online*

System / Property	Amazon Elastic Compute Cloud (EC2)	Google App Engine	Microsoft Live Mesh	Sun Network.com (Sun Grid)	GRIDS Lab Aneka
Focus	Infrastructure	Platform	Infrastructure	Infrastructure	Software Platform for enterprise Clouds
Service Type	Compute, Storage (Amazon S3)	Web application	Storage	Compute	Compute
Virtualisation	OS Level running on a Xen hypervisor	Application container	OS level	Job management system (Sun Grid Engine)	Resource Manager and Scheduler
Dynamic Negotiation of QoS Parameters	None	None	None	None	SLA-based Resource Reservation on Aneka side.
User Access Interface	Amazon EC2 Command-line Tools	Web-based Administration Console	Web-based Live Desktop and any devices with Live Mesh installed	Job submission scripts, Sun Grid Web portal	Workbench, Web-based portal
Web APIs	Yes	Yes	Unknown	Yes	Yes
Value-added Service Providers	Yes	No	No	Yes	No
Programming Framework	Customizable Linux-based Amazon Machine Image (AMI)	Python	Not applicable	Solaris OS, Java, C, C++, FORTRAN	APIs supporting different programming models in C# and other .Net supported languages

Table I.1. Comparison of some representative Cloud Platforms[6]

I.3.1. Problems Background

One of the principle tenants I want you to take away from this subject is to be aware of, and hopefully prepared for, future technology and how best to manage it. Early in the semester, we looked at corporate structures from the STO structure to Drucker's orchestra model for knowledge based organisations through to the recent structure evolving from the "Google" model.

---

[6] Rajkumar Buyya, Chee Sin Yeo, and Srikumar Venugopal, "Market-Oriented Cloud Computing: Vision, Hype, and Reality for Delivering IT Services as Computing Utilities", *IEEE Distributed Systems Online,* 2008

One technology topic I mentioned on many occasions is “cloud computing”. This is a very important topic for the future but I didn’t expand on it in any great detail. This will be your final assignment for the subject.

Preparation

You are to form groups of three. Why three? Because there are many issues about cloud computing that are to be covered in this assignment. In groups of three, individual areas of cloud computing can be distributed to group members for research and then put together for the final submission.

As I’ve mentioned to you in the past, it is important to choose your group members to cover the tasks to be done which include:

- project (assignment) planning; who is to do what
- setting milestones with an associated schedule
- coordination and supervision of the assignment
- draft review
- compilation of the final report which should include an executive summary, a table of contents and references

To prepare yourselves for the assignment, form your groups of 3 and have your initial meeting. Send me the members of the group (to jrb17a@gmail.com).

The 3 of you work within the IT department of a large multi-national company. One of you is the CTO of the company and the other 2 are his/her most knowledgeable employees.

Your CEO knows a quite bit about technology and has heard about “cloud computing”. He wants to know what it is and the implications it will have on the business so he asks the 3 of you to provide him with a report of at least 15 pages (at least 5 pages from each member of the group) on cloud computing. He insists that you cite appropriate references to support your work.

You discuss the report with him and jointly decide that the report should include at least the following 3 sections. You can add more if you wish. The questions were included to aid the writing of the report.

Since you have a lot of other work to do, he has given you 2 weeks to do the report. Jeff Barker is a director of the company and has made himself available to answer any questions as the CEO has taken 2 weeks' vacation in Mongolia and cannot be contacted.

I.3.2. Problems Identification

Cloud Computing is all good from a company standpoint; they have no desire to buy and maintain lots of computers and software. If a company can let someone else worry about the basic technology, they can focus on the publishing. Any businessperson can see the logic of that.

The flip side, though, is that the company dependent on someone else for their technology, and that can limit the flexibility and even their creativity. In fact, the personal computer itself arose because the old model of business computing, in which companies had big mainframes (aka clouds) and everyone connected to them via "dumb" terminals, was enormously frustrating for the people sitting at those dumb terminals.

I.4. Web 2.0

Web 2.0 refers to the use of the Internet for interpersonal content sharing and online service delivery[7]. Whereas the unnamed "Web 1.0" that came before it was largely concerned with creating and viewing online content (reflected in the browser wars and in a proliferation of websites that few people ever visited), the Web 2.0 concept highlights services that allow people to find and manipulate content, coupled with those that enable all types of media and services to be published and inter-linked (or "mashed") in

---

[7] Christopher Barnatt, "Web 2.0", *Explaining Computers Website,* http://www.explainingcomputers.com/web2.html, retrieved November 27, 2008

places that lots of people actually know about. Key players in the emerging Web 2.0 marketplace therefore include Google, YouTube, MySpace and Wikipedia.

One distinction between grids and clouds is that the cloud came out of the Web 2.0 mindset. Grid came out of supercomputing teams that have a very strong habit of looking at things that are very complex by nature — and they don't fear complexity — whereas clouds come out of a mindset where there has to be simplicity. It has to be something highly participatory; you take it home, and you do it, and you put your application in the cloud. If it takes more than 10 minutes, it's just not right.

At a cursory glance, Clouds appear to be a combination of clusters and Grids. However, this is not the case. Clouds are clearly next-generation data centres with nodes "virtualized" through hypervisor technologies such as VMs, dynamically "provisioned" on demand as a personalized resource collection to meet a specific service-level agreement, which is established through a "negotiation" and accessible as a composable service via "Web 2.0" technologies.

When people think of Web 2.0, usually they think of two distinct attributes. First of all, there are the data-centric mechanisms. We think about mashups and related technologies. But what more important is the community aspect, the ability to build portals and share information and have that sense of community where the knowledge of the community is able to permeate everyone who's a member of that community. People see both of those as very powerful elements inside the enterprise. What's really interesting, though, will be this challenge for IT: as IT organizations try to become more standard and create components that are reused, people want to use Web 2.0 technologies to create new applications on the fly[8].

---

[8] Dejan Milojicic, Rich Friedrich, Dave Cohen, and Alex Dreiling, "Web 2.0 and Enterprise IT," *IEEE Distributed Systems Online*, vol. 8, no. 12, 2007, art. no. 0712-oz003.

I.4.1. Amazon Elastic Compute Cloud (Amazon EC2)

Amazon is one of the companies that use Cloud Computing. Amazon Elastic Compute Cloud (Amazon EC2) is a web service that provides resizable compute capacity in the cloud. It is designed to make web-scale computing easier for developers.

Amazon EC2 has simple web service interface allows you to obtain and configure capacity. It provides you with complete control of your computing resources and lets you run on Amazon's proven computing environment. Amazon EC2 reduces the time required to obtain and boot new server instances to minutes, allowing you to quickly scale capacity, both up and down, as your computing requirements change. Amazon EC2 changes the economics of computing by allowing you to pay only for capacity that you actually use. Amazon EC2 provides developers the tools to build failure resilient applications and isolate themselves from common failure scenarios.

I.5. Goals

In everything related to technology, there are always pros and cons. These are our findings in the Cloud Computing[9]:

- **Security** → Not a month goes by without us hearing about a security breach or leaked personal information. Which begs the question: In the cloud, where exactly is all this data being stored? Who has access to it? And how easy is it for someone who doesn't to get it? The answers to these questions just aren't as clear cut as they need to be for many SaaS businesses, especially those handling loads of personal data.

[9]Marty Abott, Michael Fisher, "The cloud isn't for anyone", New York Times Web Site, http://www.nytimes.com/external/venturebeat/2008/10/13/13venturebeat-the-cloud-isnt-for-everyone-99024.html, retrieved November 27th, 2008

- **Non-portability** → While some cloud computing platforms can work with industry standard software like VMware, they still differ from vendor to vendor. An example to illustrate why this is significant: Let's say you want to store an application or website on Amazon's EC2 cloud platform. In this case Amazon is the vendor, and you build what is called an Amazon Machine Image (AMI). Then you debug it and ensure everything works. But now that you're done, you can only use this AMI within EC2 clouds. If you want to access or use the data you've stored from another platform, you have to go about recreating your AMI all over again inefficiently.

- **Control** → When you entrust an application to the cloud, you're basically turning control over to a third party. This third party determines whether your application runs or not, whether your site loads or not and how to respond if your application or site fails to work properly. This raises the issue of service level agreements (SLAs), documents that outline the service a cloud platform provider agrees to give its users.

- **Limitations** → Even though cloud vendors are actively whittling away at the limitations of their platforms, businesses should always ask about limitations before committing, as they can lead to serious problems if not considered. Here are some things to take a close look at:

    - Persistent storage on virtual servers
    - Public and static IP addresses for virtual servers
    - Certification of databases in a virtual environment
    - Load balancing hardware or software

- **Performance** → In the cloud, many virtual servers are maintained by a smaller number of physical servers. This virtualization process requires a lot of memory and system resources and in turn can cause applications running in cloud environments to run slower than they would on actual hardware.

Whether there are still many cons about Cloud Computing, the good things that we want to say about cloud computing[10]:

1. **Cost** → We're definitely fans of the cloud computing cost structure. Paying only for actual usage makes perfect sense for growth-oriented companies, especially start-ups. Jeff Barr, web services evangelist at Amazon, notes that he can't name a single failed start-up that used the company's EC2 and S3 platforms for its core infrastructure. That's certainly something to consider when you're deciding how to host your next billion dollar idea.

2. **Speed** → Another thing to love about cloud computing is how fast and easy it is to pull in resources from other servers to improve the performance of an application or website. This is great for companies looking to scale. Let's say your site is slammed with traffic after a major announcement. With the cloud, you can simply deploy a few more servers to handle the load until the storm passes — and then draw back down later to keep costs under control.

3. **Flexibility** → ability to implement new servers at a low cost and without server procurement.

[10]Marty Abott, Michael Fisher, "The cloud isn't for anyone", New York Times Web Site, http://www.nytimes.com/external/venturebeat/2008/10/13/13venturebeat-the-cloud-isnt-for-everyone-99024.html, retrieved November 27th, 2008

I.6. Contexts Boundary

History and current state

- What is cloud computing?
- What is the current state of cloud computing?
- How is it related to Web 2.0?
- Who were the original companies who bought into cloud computing?
- What are the perceived pros and cons of cloud computing (if any)?

Changing technology

What will be its effect on the traditional computing components of:

- hardware
- operating systems
- application software
- databases
- middleware
- infrastructure
- communication facilities

Management Issues

Within an organisation, what will be the changes, if any, to technology requirements and services?

Will alignment of business processes be easier or more difficult with cloud computing?

Issues within an organisation in relation to cloud computing:

- reliability
- maintenance
- security
- risk management
- managing service provider relationships
- managing legacy systems
- managing assets
- organisational issues
- outsourcing