THE NEW CLOUD COMPUTING PARADIGM: THE WAY TO IT SEEN AS A UTILITY

Carlos Almarcha Arias
Universidad Politécnica de Madrid, Madrid, España, carlos.almarcha@upm.es

Carlos Rodrìguez Monroy
Universidad Politécnica de Madrid, Madrid, España, crmonroy@etsii.upm.es

Yilsy M. Núñez Guerrero
Universidad Politécnica de Madrid, Madrid, España, nyilsy@gmail.com

ABSTRACT

In the present competitive environment, companies are wondering how to reduce their IT costs while increasing their efficiency and agility to react when a change in the business processes is required. Cloud Computing is the latest paradigm to maximize the IT resources considering "everything as a service" and received from the Cloud (Internet) instead of owning and managing hardware and software assets. The benefits from the model are clear but there are also concerns and issues that need to be solved in order to extend the use of Cloud Computing across the industries. This model will allow a pay per use model for the IT services and many benefits like cost savings, agility to react when business demands changes and simplicity because there will not be any infrastructure to operate and administrate. It will be comparable to the well known utilities like electricity, water or gas companies. However, this paper underlines several risk factors and it will be necessary that the leading technology companies investigate on solutions to minimize the risks described in this article.

Keywords: Cloud Computing, Utility Computing, Elastic Computing, Enterprise Agility

1. INTRODUCTION

In the present competitive environment, companies are wondering how to reduce their IT costs while increasing their efficiency and agility to react when a change in the business processes is required. Cloud Computing is the latest paradigm to maximize the IT resources considering "everything as a service" and received from the Cloud (Internet) instead of owning and managing hardware and software assets. The benefits from the model are clear but there are also concerns and issues that need to be solved in order to extend the use of Cloud Computing across the industries.

Mel and Grance (2009) define the model in the following way: "Cloud computing is a pay-per-use model for enabling available, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, services) that can be rapidly provisioned and released with minimal management effort or service provider interaction".

The interest of users and organizations on Cloud Computing has grown very fast since 2004 until 2011 with some trend to flatness in 2012. This analysis suggests the maturity that this model is reaching.

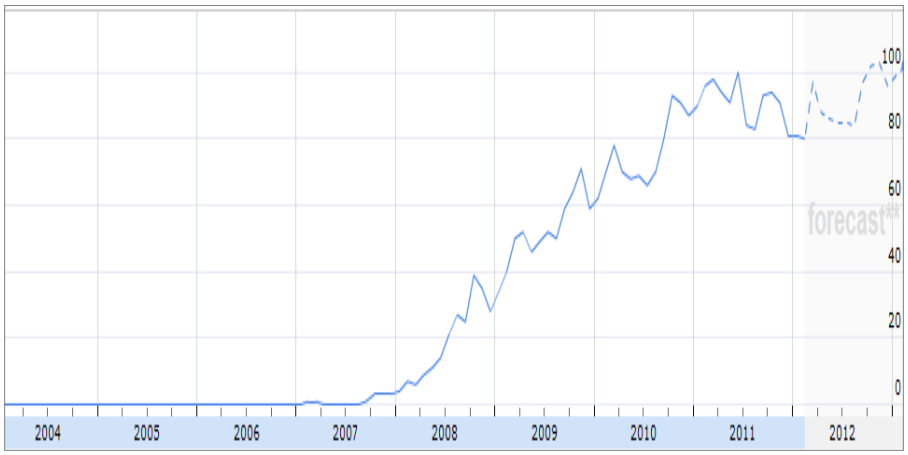

Figure 1. Average worldwide traffic of "cloud computing" in all years updated Feb 2012 (Source: Google data base)

2. Cloud Computing history and different models

Before Cloud Computing has become popular other similar models were proposed in the past like the called Network Computing in the mid of the 90's. This model was too much theoretical and did not work simply because the necessary underling technologies to sustain the services were not mature enough.

Perhaps, the first reference to Cloud Computing was made by Eric Schmidt CEO and Chairman of Google during the Search Engine Strategies Conference on August 9th, 2006. Here is a fragment of the comments:

"…What's interesting [now] is that there is an emergent new model, and you all are here because you are part of that new model. I don't think people have really understood how big this opportunity really is. It starts with the premise that the data services and architecture should be on servers. We call it cloud computing – they should be in a "cloud" somewhere. And that if you have the right kind of browser or the right kind of access, it doesn't matter whether you have a PC or a Mac or a mobile phone or a BlackBerry or what have you – or new devices still to be developed – you can get access to the cloud. There are a number of companies that have benefited from that. Obviously, Google, Yahoo!, eBay, Amazon come to mind. The computation and the data and so forth are in the servers."

After the conference, the term become generally accepted by other important players in the IT services world like Amazon, who used the same words to describe their new service EC2 launched on August 24th, just a couple of weeks after the Schmidt's speech.

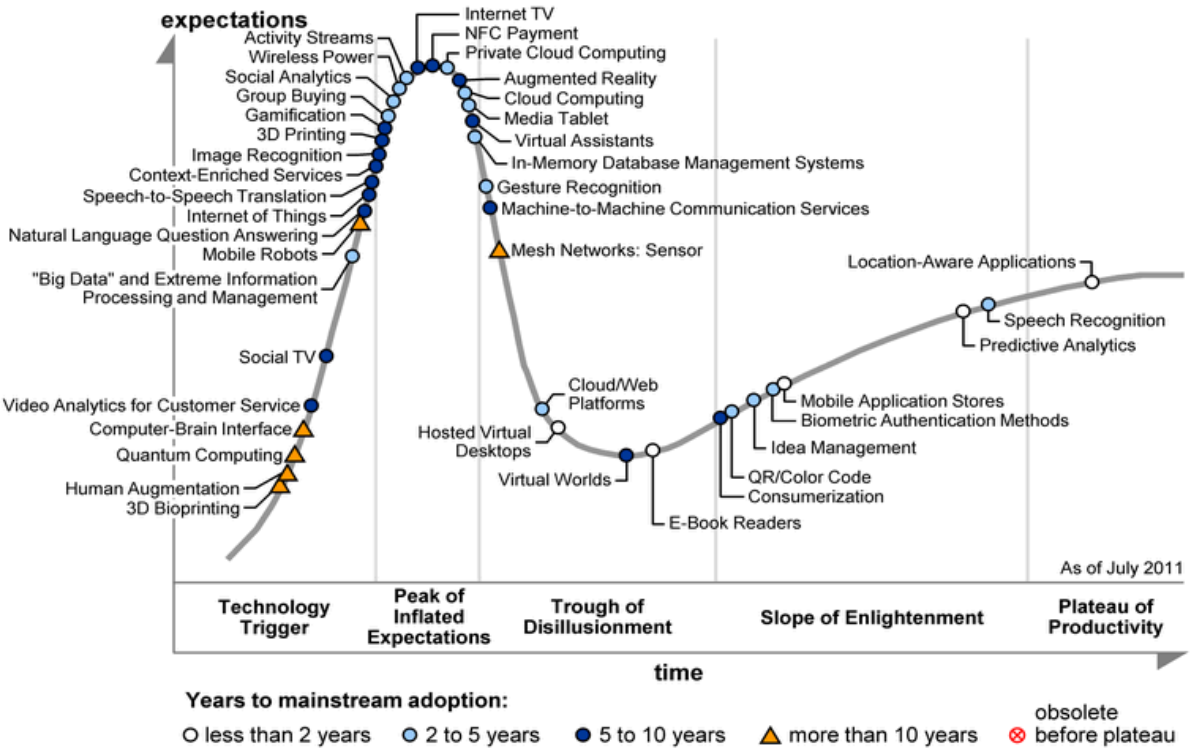

Figure 2. Hype Cycle for emerging technologies. Source: Gartner Group (2011)

According to Gartner the enthusiasm for a new technology follows a path like the shown in the figure. After the launch and introduction of a new technology there is a fast ramp where it gains visibility across the market. After that, there is an inflection point where the perception is that the expectations were too much inflated. Following there is a period of time of disillusionment until the market recognizes the real value.

Cloud Computing is still in the first ramp with growing expectations.

There are three main service types of Cloud Computing:

1. Cloud Software as a Service (SaaS). Use provider's applications over a network
2. Cloud Platform as a Service (PaaS).) Provide a computing platform and a solution stack as a service. Usually this solution stack consists of software for application development.
3. Cloud Infrastructure as a Service (IaaS). Rent processing, storage, network capacity, and other fundamental computing resources.

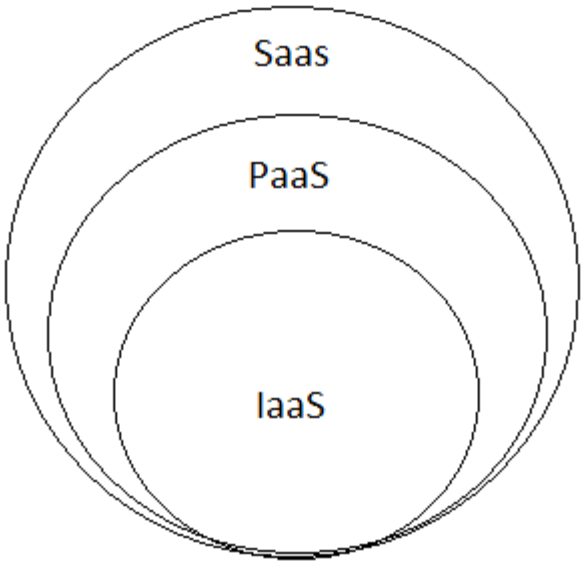

Figure 3. Types of services under the Cloud Computing model

3. Characteristics and benefits of the Cloud Computing model

Following are the main characteristics of the model:

- On-demand self-service. The services can be activated o deactivated by the users in a simple way through a web portal.

- Ubiquitous network access. The users are able to access the services from everywhere. They just need a device with a web browser and an Internet connection.

- Location independent resource pooling. It does not matter where the infrastructure to deliver the Cloud Services is physically located. The location is transparent for the users who receive the services from Internet.

- Rapid elasticity. The contracted capacity can be increased or decreased by the user according to the business needs at any time. This is a key factor to gain agility and synchronize the IT resources to the business needs.

- Pay per use. The users pay for the services based on the consumption. This is very different from the traditional model where the organizations had to invest in HW&SW assets. On top of that they had to pay for operational expenses (OPEX) like maintenance contracts and administration costs.

With Cloud Computing, many small and medium companies will be able to access IT services forbidden for them under the traditional model. For example, let's consider an application of CRM (Customer Relationship Management System). Many small companies usually have no budget enough to afford investing on the application, servers, network elements plus the expenses to maintain, operate and administrate the new hardware and software. With Cloud Computing any company, large or small, can activate a CRM service online and benefit from this type of applications without making any investment in infrastructure. This means that small and medium companies will be able to compete better with large companies by using similar resources and applications like a CRM.

Large companies will benefit also from Cloud Computing because they will reduce costs and will gain competitiveness. The cost of contracting a new IT service from a cloud service provider is much lower than investing on hardware and software assets.

Finally agility is another important benefit. When a decision to deploy a new application is made, it takes usually months finding budget, selecting the vendors for new hardware and software, negotiating prices, launching the orders and installing and testing the new systems. In Cloud Computing the service is provided almost instantaneously when it is contracted with a Cloud service provider and this improves the enterprise agility.

There are three types of clouds:

- Public clouds. These clouds are hosted by service providers offering public services to any organization who wants to contract them.
- Private clouds. Some companies, especially large ones, try to benefit from the Cloud Computing model maintaining control of data and infrastructure. They deploy private clouds in their own data centers. Deploying this type of clouds the companies have to make investments as well as expenses for maintenance, operations and administration.
- Hybrid clouds. In this case part of the services is delivered from public service providers and part of the services is provided from a local private cloud. The criteria to decide the type of cloud depends mainly on the criticism of the service. When the service is very critical or the data are sensitive then the private cloud should be the choice.

4. Examples of IT services given from the Cloud

Traditionally, companies have invested important amounts of money in hardware, software, data centres and IT personnel to support their business processes, but this model is not efficient at all. Some of the issues are the following:

1. Organizations invest in HW, SW licenses, network elements …etc, independently of the real utilization needs of these resources. Most of the times the applications are needed only eventually but they are purchased as an asset. So IT departments are investing in HW & SW resources to be available at any time but they will be used at part time only.

2. Following with the traditional model, purchasing resources for a new IT service is a time consuming task. New budget has to be approved, a purchasing process has to be carried out, and at the end the deployment and integration of the new resources also takes a long time. This can be a delaying factor for any new service that has to be launched impacting the overall business results.

3. Owning hardware, software, network elements and other assets means that specialized and skilled personnel are needed to operate and administrate the entire infrastructure. These are human resources and more headcount inside the organizations.

4. Any element of infrastructure has to be under a support contract with the vendor with the proper level of support according to the service criticality. Companies spend a significant part of their IT budget to maintain the existing infrastructure instead of spending that money to innovate and create new IT services that could improve the competitiveness and the business results.

In opposition to the traditional model owning and operating the necessary IT infrastructure the underlying idea behind the Cloud Computing model is to receive the IT resources as services from Internet, including applications, processing capacity or storage.

Some examples of real Cloud Computing services:

4.1. Cloud Computing services from Amazon

The Could Computing services from Amazon include Amazon S3 for storage , Amazon EC2 for hosting and SimpleDB for database. Amazon EC2 (Amazon Elastic Compute Cloud) 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's simple web service interface allows you to obtain and configure capacity with minimal friction. 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.

4.2. Salesforce.com

Salesforce.com is a company focused on CRM solutions (Customer Relationship Management) On-demand. Their customers do not have to purchase hardware and software. Just a connection to Internet and they have access to the application without any need of purchasing, installing, configuring and managing hardware and software. They just make a subscription and they have a quick access to the CRM application. The new service can be fully operative in days or weeks instead of the months or even years needed to deploy a CRM application inside a company. When the CRM service is not needed the subscription can be cancelled. salesforce.com was founded in 1999 by an executive of Oracle, Marc Benioff, pioneer in offering enterprise applications from a website.

4.3. Google Apps Engine

Google App Engine. This is a tool launched by Google for developers enabling them to run web applications on Google's infrastructure. This will allow start-ups to use Google's web servers, APIs, and other developer tools to build a web app on top of. Google App Engine is similar to the Amazon Web Services stack, which rolled out at the end of 2006 and has since gone on to be utilised by many startups for their infrastructure needs. But it is not a set of standalone services like Amazon's - which includes S3 for storage, EC2 for hosting and the SimpleDB database. Google App Engine is an end-to-end service and bundles everything into one package. It was launched as a beta pilot with 10,000 developers and during the preview period the resources are limited to 500MB of storage, 200M megacycles of CPU per day, and 10GB bandwidth per day.

4.4. HP Cloud Assure

On April 27, 2009 HP announced HP Cloud Assure, a new solution composed by software and services joining the concepts of Storage as a Service (SaaS) and Cloud Computing. According to IDC, the three basic pillars for success in Cloud Computing are security, performance and availability, and these are the areas addressed by HP. The solution includes three applications from the software division of HP but offered as a service through Internet instead of buying software licenses: HP Application Security Center, HP Performance Center y HP Business Availability Center:

- Security: Analyzing networks, operating systems, layers of middleware and web applications. It makes also tests to identify potential vulnerabilities and risks. So, data from providers and consumers is safe.

- Performance: Assuring the cloud computing services meet the connectivity and bandwidth requirements of the end user. This helps to control the service level agreements to improve the end user satisfaction.

- Availability: Monitoring the applications availability on the cloud to isolate potential issues and analyze the root causes of any incident as soon as possible.

4.5. Google Apps Engine

This is a tool launched by Google for developers enabling them to run web applications on Google's infrastructure. This will allow startups to use Google's web servers, APIs, and other developer tools to build a web app on top of. Google App Engine is similar to the Amazon Web Services stack, which rolled out at the end of 2006 and has since gone on to be utilised by many startups for their infrastructure needs. But it is not a set of standalone services like Amazon's - which includes S3 for storage, EC2 for hosting and the SimpleDB database. Google App Engine is an end-to-end service and bundles everything into one package. It was launched as a beta pilot with 10,000 developers and during the preview period the resources are limited to 500MB of storage, 200M megacycles of CPU per day, and 10GB bandwidth per day.

5. Inhibitors and risks for the Cloud Computing model adoption

The main inhibitors for the Cloud Computing adoption will be the risk factors associated to this model:

1. Service levels – Can the service provider meet the business needs for network and application availability?

   Owning and keeping the IT assets inside the company provides much better control of data and infrastructure. In the case of Cloud Computing the users must rely on the service level agreements signed with the service provider.

2. Workloads – Are they suitable for Cloud?

   For very heavy workloads, the networks and the service provider infrastructure could be a bottle neck to provide the right application response times.

3. Complexity – Difficult to manage for a Cloud provider?

   Again, the Colud Computing model means more complexity for the service provider. The services have to be activated online and measured for the pay per use. On the other hand the pool of hardware and software resources at the service provider side is multitenant. This means that servers, disks, applications …etc are shared for multiple users and this is an extra of complexity.

4. Network Latency – Affects end-end service experience

   A successful service based on Cloud Computing requires low network latency. If the network is slow the final user experience will not be satisfactory. At the end Cloud Computing is "everything as a service" and the services are delivered through the networks. A poor network performance will impact on the service.

5. Lock-in – How easy is it to migrate from one provider to the next?

   Sometimes it can be difficult for a company to migrate a service from one service provider to another. This could be a barrier that makes the companies feel uncomfortable.

6. Regulations – Where does the data reside?

   When a company contracts any of the possible Cloud Computing services the data are stored remotely and many times in a foreign country. In some countries there are regulations the data management forbidding moving sensitive data out of the border.

7. Security – Who will have access?

   When the data are stored at the facilities of the Cloud Computing service provider the user has no control of who has access to the data. He must trust that the Service Provider has the right processes and tools to guarantee data protection.

8. Financial stability – How stable is the provider?

   A service provider with a bad financial situation could put at risk the future service continuity forcing a migration to another service provider. This is an undesirable risk for any organization.

CONCLUSION

Cloud Computing is a new paradigm where the organizations will receive the IT services from the cloud instead of having the classic infrastructure of servers, applications, storage and networks inside the company. This model will allow a pay per use model for the IT services and many benefits like cost savings, agility to react when business demands changes and simplicity because there will not be any infrastructure to operate and administrate. It will be comparable to the well known utilities like electricity, water or gas companies.

However, this paper underlines several risk factors and it will be necessary that the leading technology companies investigate on solutions to minimize the risks described before.

Acknowledgements
Amazon, Salesforce.com, Hewlett Packard, Google

REFERENCES

Carr, Nicholas (2009). The Big Switch, from Edison to Google. W.W. Norton & Co.

Reese, George (2009). Cloud Application Architectures: Building Applications and Infrastructure in the Cloud. O'Reilly Media, Inc.

Fingar, Peter (2009). Dot Cloud: The 21st Century Business Platform Built on Cloud Computing. Meghan-Kiffer Press.

Millar, Michael (2008). Cloud Computing: Web-Based Applications That Change the Way You Work and Collaborate Online. Que.

Rittinghouse, John and Ransome, James (2009). Cloud Computing: Implementation, Management, and Security. CRC Press.

Hurwitz, Judith; Bloor, Robin; Kaufman, Marcia and Halper, Fern (2009). Cloud Computing For Dummies. Publisher: For Dummies.

Menken, Ivanka (2008). Cloud Computing - The Complete Cornerstone Guide to Cloud Computing Best Practices Concepts, Terms, and Techniques for Successfully Planning, Implementing ... Enterprise IT Cloud Computing Technology. Emereo Pty Ltd.

Menken, Ivanka and Blokdijk, Gerard (2008). Cloud Computing Foundation Complete Certification Kit - Study Guide Book and Online Course. Emereo Pty Ltd.

Velte, Toby; Velte, Anthony and Elsenpeter, Robert (2009). Cloud Computing, A Practical Approach. McGraw-Hill Osborne Media.

Wohl, Amy (2008). Succeeding at SaaS: Computing in the Cloud. Wohl Associates.

Granneman, Scott (2009). Google Apps Deciphered: Compute in the Cloud to Streamline Your Desktop. Prentice Hall PTR.

Katz, Richard N.; Oblinger, Diana G. (2008). The Tower and the Cloud: Higher Education in the Age of Cloud Computing. Educause.

Autorización y Renuncia

Los autores autorizan a LACCEI para publicar el escrito en las memorias de la conferencia. LACCEI o los editores no son responsables ni por el contenido ni por las implicaciones de lo que esta expresado en el escrito