



**Linnæus University**

School of Computer Science, Physics and Mathematics

Degree Project

# A Survey on Cloud Computing and Prospects for Information Visualization

M Hüseyin Öztürk  
2010-11-15  
Subject: Information Visualization  
Level: Bachelor  
Course code: DA3003

## **Abstract**

Today's computing vision makes users to access services, applications via lightweight portable devices instead of powerful personal computers (PC). Since today's applications and services need strong computing power and data storage, raising question will be "Who will provide these 2 attributes if users do not?" Cloud computing trend moves computing power and data storage from users' side to application infrastructure side. The services that traditionally stored in users' own computers will move into cloud computing platform and delivered by the Internet to its users. This new platform comes with its own benefits and design characteristics. Since all information data will move into another platform than individual computers, information visualization will be an opportunity field to analyze and maintain the cloud system structure as well as delivering abstract data into meaningful way to end users.

Keywords: cloud, computing power, architecture, information visualization.

## **Acknowledgments**

I would like to show my gratitude to my supervisor, Prof. Dr. Andreas Kerren, whose support and advises made me to understand and finalize the subject.

Lastly, I am grateful to Yasin Bahtiyar who supported me in any respect during my project.

# Table of Contents

<b>1</b>	<b>INTRODUCTION .....</b>	<b>1</b>
1.1	PROBLEM .....	1
1.2	GOALS AND CRITERIA .....	1
1.3	MOTIVATION .....	1
1.4	OUTLINE .....	2
<b>2</b>	<b>CLOUD COMPUTING .....</b>	<b>3</b>
2.1	INTRODUCTION .....	3
2.2	CLOUD COMPUTING HISTORY .....	4
2.2.1	<i>Client/Server Architecture Model</i> .....	6
2.2.2	<i>Peer To Peer Architecture Model</i> .....	7
2.3	WHY CLOUD COMPUTING MATTERS? .....	7
<b>3</b>	<b>ARCHITECTURES OF CLOUD SYSTEMS .....</b>	<b>9</b>
3.1	CLOUD TRENDS .....	9
3.1.1	<i>Virtual Machines</i> .....	9
3.1.2	<i>Services that are Delivered Over Networks</i> .....	9
3.1.3	<i>Open Source Development</i> .....	10
3.2	CLOUD SYSTEMS INFRASTRUCTURE .....	10
3.2.1	<i>Public Clouds</i> .....	11
3.2.2	<i>Private Clouds</i> .....	11
3.2.3	<i>Hybrid Clouds</i> .....	12
3.3	CLOUD SYSTEMS' ARCHITECTURE LAYERS .....	12
3.3.1	<i>Software as a Service (SaaS)</i> .....	13
3.3.2	<i>Infrastructure as a Service (IaaS)</i> .....	14
3.3.3	<i>Platform as a Service (IaaS)</i> .....	14
3.4	REAL WORLD EXAMPLE .....	14
3.4.1	<i>Amazon S3</i> .....	15
3.4.2	<i>Amazon EC2</i> .....	16
<b>4</b>	<b>CLOUD COMPUTING: BENEFITS AND DRAWBACKS .....</b>	<b>17</b>
4.1	BENEFITS .....	17
	<i>Usage Based Costing</i> .....	17
	<i>Processing Time Efficiency</i> .....	17
	<i>Free Space</i> .....	17
	<i>Flexibility</i> .....	17
	<i>Scalability</i> .....	17
	<i>Portability</i> .....	18
4.2	DRAWBACKS .....	18
	<i>Dependability</i> .....	18
	<i>Security</i> .....	18
	<i>Little or No Reference</i> .....	18
<b>5</b>	<b>VISUALIZATION POSSIBILITIES IN CLOUD COMPUTING .....</b>	<b>20</b>
5.1	INFORMATION VISUALIZATION .....	20
5.2	EXISTING VISUALIZATIONS .....	21
5.2.1	<i>Force.com and Google Visualization API</i> .....	21
5.2.2	<i>The World's Eyes Project</i> .....	24
5.2.3	<i>Newsmap</i> .....	25
5.2.4	<i>Fidg't</i> .....	26
5.3	VISUALIZATION POSSIBILITIES .....	26
<b>6</b>	<b>CONCLUSION AND FUTURE WORK .....</b>	<b>32</b>
6.1	CONCLUSION .....	32
6.2	FUTURE WORK .....	32
<b>7</b>	<b>REFERENCES .....</b>	<b>33</b>

# 1 Introduction

Cloud computing is defined as an “evolving paradigm” by The National Institute of Standards and Technology (NIST). This new technology platform provides computing services over the Internet connection instead of placing and running these services individually on each user’s personal computer.

Information visualization performs transformation of the data into a visual form. Thus makes huge amount of data resources easier to understand, realize and browse. The role of visualization in cloud computing is one of the big key points since cloud computing acts as an extension for users’ own personal computers. According to “The Pew Internet Project”; (1)

- 56% of the Internet users use web mail services such as Hotmail, Gmail, or Yahoo! Mail.
- 34% of them store personal photos online.
- 29% of them use online applications such as Google Documents or Adobe Photoshop Express.

As we can see in the research, cloud computing has already become an important part of our daily life. All these information data and application processes are going to need efficient and successful way to be expressed with help of information visualization.

## 1.1 Problem

As we can see in history, every incremental development of technology causes a major shift of how applications are formed, and provided to end users. After the invention of computer networking, first aim was to adapt systems according to client-server applications. Nowadays, we are witnessing powerful mobile devices, and stronger Internet connections all over the world. Today’s technology is dealing with “on demand”; customizable and shared Internet based services instead of client-server model’s inflexible and expensive environment. Every each shift in technology environment comes with its own characteristics so cloud computing. Since the deployment of software and hardware is completely different than traditional platforms, role of information visualization in cloud computing environment is standing as a big issue for both cloud computing vendors and clients.

## 1.2 Goals and Criteria

The result of the work is formed by two main goals. The first goal is to give a comprehensive view of cloud computing platform. This view contains the history of cloud computing, cloud computing architecture, architectural layers and closing with benefits and drawbacks of cloud computing usage. Second goal is to investigate the existing visualizations that already provided by cloud computing vendors and research about visualization possibilities in cloud computing.

## 1.3 Motivation

Cloud computing shift placed into technology environment with absolutely different structure than the other traditional systems. The big differences such as, maintainability of application, security issues, on demand access and device independence are general issues that needed to have background knowledge from IT crew and end users of cloud computing.

Information visualization is composed of complete visualization techniques in order to deal with data. Information visualization has been an essential technique to make more sense about data. Thus, the interaction between data and its users is an important part of cloud computing to analyze, present and explore the information.

#### **1.4 Outline**

The report structure is created as follows; Chapter 2 gives background knowledge of cloud computing with its history. Chapter 3 describes cloud computing architecture and trends. Also this chapter explains cloud architecture with a real world example. Chapter 4 clarifies advantages and disadvantages of cloud computing usage. In Chapter 5, initially, the brief description of information visualization is expressed. The investigation of existing visualizations is reported. Finally, visualization possibilities and ideas of cloud computing is explained. As a last chapter, Chapter 6 explains the conclusion and a general look of the report.

## **2 Cloud Computing**

This chapter describes the cloud computing technology. It will give a general view about cloud computing history and the attributes of this platform.

### **2.1 Introduction**

The key term of cloud computing “Cloud” is based on hiding the complexity of today’s technology framework from users. In other words, people do not deal with the large amount of data processes which is actually essential, in order to make their applications run. By using “Cloud”, users’ only task is to connect the cloud, access their data and applications in the same way as old PC centric solutions. The difference here from traditional systems is; the issues of application deployment, system requirements and storage are not a prerequisite in order to start using applications.

Cloud computing has key properties which distinguishes cloud computing platform from PC centric systems. (2) Initially, cloud computing provides user centric systems. Once the authorized user is logged in a cloud system, all information that is related to the user becomes absolutely accessible. At the same time, user can share or access authorized data with other users on the system. Sharing options can be customized by access levels. Users do not need to own shared data in order to use it. Secondly, cloud computing focuses on task centric solutions. That means; the knowledge of how to use or maintain the application is not an important issue anymore. Instead of dealing with complexity of deployment, maintenance, users spend their time focusing into what they do with applications. In other words, computer users do not need to spend their time for applications’ maintainability. Since then, the focus is kept in work instead of spending it on applications too. The power attribute of cloud computing comes from huge amount of connected computers in cloud. Cloud user’s advantage here is to process data with powerful processors instead of doing the same process with a single computer. For example; video rendering needs a lot of processing power and ram capacity. 30 minutes movie clip can take hours to render with a single computer. By using cloud computing, output video will be ready in a short time. Accessibility is another key point. Users access the information and their data from multiple servers in cloud. This feature gives instant accessibility, comparing to the single personal computer that yields single source of data. There may be lack of accessibility in case of personal computer or server crashes. And that means unavailable data session. These downtimes cannot be acceptable for enterprise businesses. If any server fails in cloud, the virtual machine that is placed in that server will be transferred to another instance and continue its service. Another attribute of cloud computing is Intelligence. Cloud computing environment stores various types of data. Intelligence of cloud computing comes into play for extracting the patterns of data information. What makes cloud computing programmable is; cloud computing systems work with automated tasks. These tasks are programmed by developers of cloud vendors. For instance, if one of the computers in the cloud fails; this might be going offline, operating system crash, hardware problems and etc.; the data on that server must be transferred or redistributed to another servers in cloud automatically. So, users of that information would be redirected to new servers during this downtime. Users will not have any problem during the server failure.

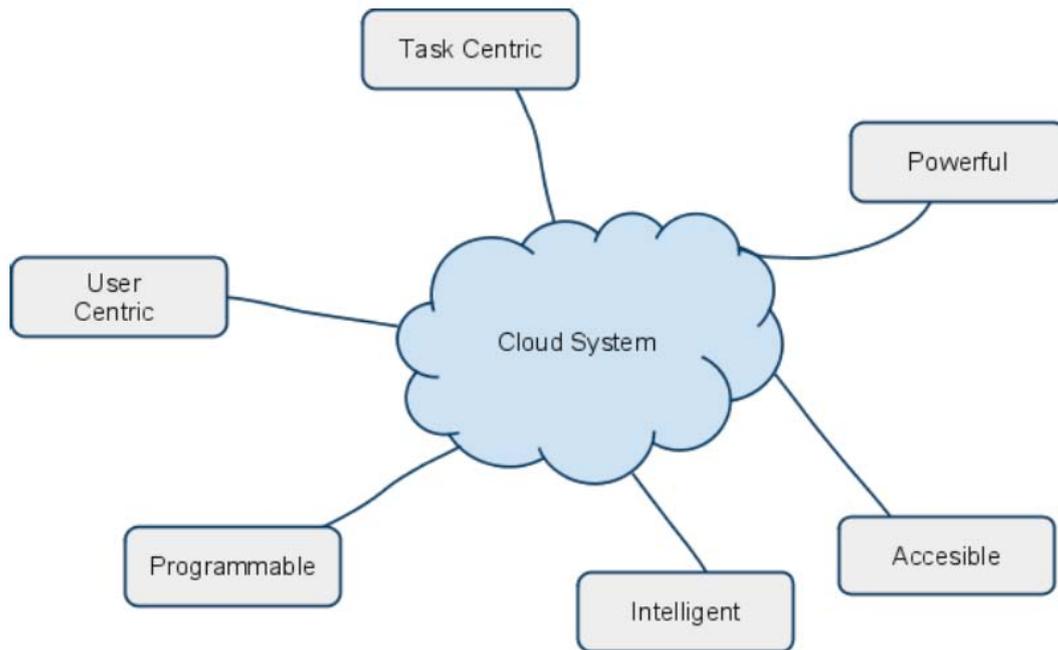


Figure 2.1: Attributes of Cloud Computing

According to attributes of cloud systems, they have mainly two perspectives. First aspect concerns the end users which connect the cloud and perform their tasks. The other aspect contains the people which manage cloud systems called IT and developers. Since there is not only one point of view in cloud computing, there are various definitions of “cloud computing” depending on its usage.

## 2.2 Cloud Computing History

The Cloud Computing as a term is not new in respect of its history. But it has evolved so fast for today’s users with support of processing and serving power. The creation of “Cloud” has still the same idea from the beginning, but this idea that coined by McCarthy, firstly named as “utility computing” in sixties. (3) His concept was to deliver information as the same way as power or electricity companies. The computing power should be centralized and everyone should use it from the central framework instead of everyone’s having individual. Since that time from the sixties, cloud computing has evolved with web 2.0. This is also because of the Internet bandwidths’ evolution in nineties. By having high bandwidth, network users have had higher level of communication layers. If we look at the communication and computing history briefly, we see two different models; Client/Server and Peer to Peer (P2P) models. In client/server models, all information and applications are stored at servers and once user request information, s/he has to connect the main computers called servers and receive data from servers via user’s terminal called client. Client/Server model was not the perfect way of computing. So the question for Client/Server model was “what is not perfect in Client/Server model?”

We can answer the question by taking a look at drawbacks of Client/Server model. Initially, users need to have granted authentication to access the main servers. Although servers are powerful enough, clients do not have instant access of information management. Another issue in this model is that the same information stored on servers cannot be accessed from several users at the same time. Only administrators and IT staff of Client/Server models have complete access of administrative actions. Clients can

only perform the actions that are allowed by administrators. No variations or flexibility is allowed. For instance, one of the clients might want to change or update any data on servers. Execution steps can be done in this ordinary; first, user requests to update a specific data on server. This can be a monthly financial report. This action waits in the queue for the permission. There is mostly big numbers of clients and few administrators. In this case, client's update will take days in order to place in servers. Briefly, a Client/Server system does not provide instant access to resources and another defect is; it does not utilize computing power of the client modules as effectively as the computing power of the server module. This is a disadvantage of current technology environment. For end users and enterprise business, time usage is as important as the rest of issues.

The other model that called Peer to Peer (P2P) computing provides communication for clients of system without depending on any server. With this model, communication and processes does not have to pass through the server first. In Peer to Peer model each computer or module has the equivalent rights and privileges. This is exactly the opposite of Client/Server model. As mentioned before, in Client/Server model, server module has all the administrative privileges and clients have to deal with them. So, how communication is executed in peer to peer computing systems? In P2P systems, all computers that have connected to system have 2 roles. They behave both as a server and client. All computers have the same rights and there is no computer that has an advantage to another one. By p2p computing, direct data exchange process is done between peers with the help of hubs.

Another characteristic of peer to peer computing is to spread the computing power between all peers. There is no centralized computing power in this model. By evolution of the Internet, peer to peer computing entered a new age. At the beginning of this model, peer to peer computing had been used for the Usenet to connect Usenet computers each other at seventies. After the evolution of the Internet peer to peer computing has been used worldwide. But that did not make all users to integrate their systems to peer to peer computing. Client/Server systems have also been in the picture. In World Wide Web, most of the systems are stand by Client/Server model. One important proceed of peer to peer computing is bringing the distributed computing idea. With this idea, many complex tasks had been succeed like; research projects in laboratories, complex encryptions, and space researches. Connecting millions of computers had provided a giant computing power.

During all these development history from Client/Server to Peer to Peer computing, people always looked for better communication ways. People have wanted to collaborate and access the information for projects, researches at real time from anywhere. And these goals have brought the cloud computing idea. What cloud computing offers different than the other models is; Firstly, cloud computing provides to work on systems in real time. People access their information and update them anytime and these changes can be seen immediately from the others who are involved in that information. For example, the people who are involved to big project works are not localized in the region even sometimes not in the same country. The solution here is to connect these people to their information without depending where they are. They do not need any special computer or additional software in order to access their data. Cloud computing creates these communication between users from different regions like in a private network. Here is a simple example of cloud computing usage. Google provides free email application to all Internet users. By using Google Mail (Gmail), there is no need to have a mail server to send and receive emails. Gmail stores all users email messages in cloud and they can access messages anytime from any computer. The infrastructure behind the email account is already provided by Google. Storage,

maintenance and accessibility are not a problem of users anymore. By the evaluation of cloud computing environment, our traditional way of data storage, information sharing and data mining will be absolutely different and easy. If we look a little bit forward, cloud computing usage will not be only limited by our computers but also any device that we use in daily life. These devices like, cars, televisions will be all compatible for cloud services to make life easier and more accessible than before.

### 2.2.1 Client/Server Architecture Model

The Client/Server architecture is mainly structured on two modules:

- A server module: Single and unique instance of system that works as a main station and all procedure is done by this module.
- A client module: Client module is formed with several client computers that are presented in system. Client modules have an interface to be able to connect, perform tasks and send requests to servers.

The client/server model can be defined as a complex system with its sub modules and components. The key attribution of this architecture is the server module. Server module is the key point for whole processes. In this architecture, there is no way for clients in order to communicate with each other. The communication is being done between server and clients as you can see in Figure 2.2. Clients that are connected to system have mostly simple components. The server has commonly the more complex part of the system.

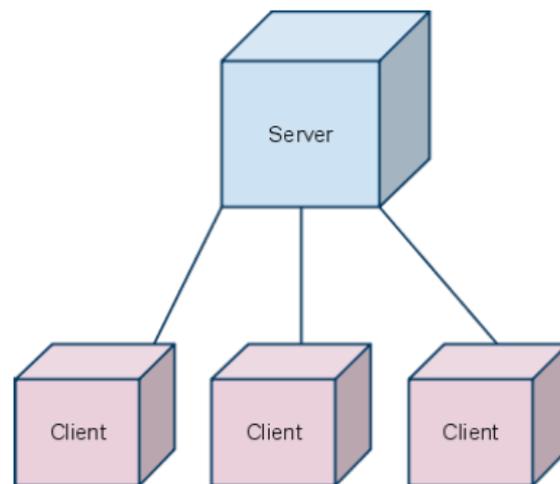


Figure 2.2: Client/Server Model

In this architecture, different components of the system need to know each other before they start to communicate. And this communication is provided by network addresses. Authentication procedure is started by clients all the time. There are no server requests to its clients in this protocol. Therefore, each client have to know network address of the server that they want to connect. Server defines specific ports and network addresses. These ports and IP addresses are known by the clients. Once the client connects to the server, then they are ready to communicate with each other. As we mentioned before server module does not need to be configured with any information about the clients.

## 2.2.2 Peer To Peer Architecture Model

Peer to Peer architecture model consists of various software modules and every module of the system generally keeps processing on different computers. These different software modules connect with each other to achieve the processing for application completion. One peer can act both as a client and server module at the same time.

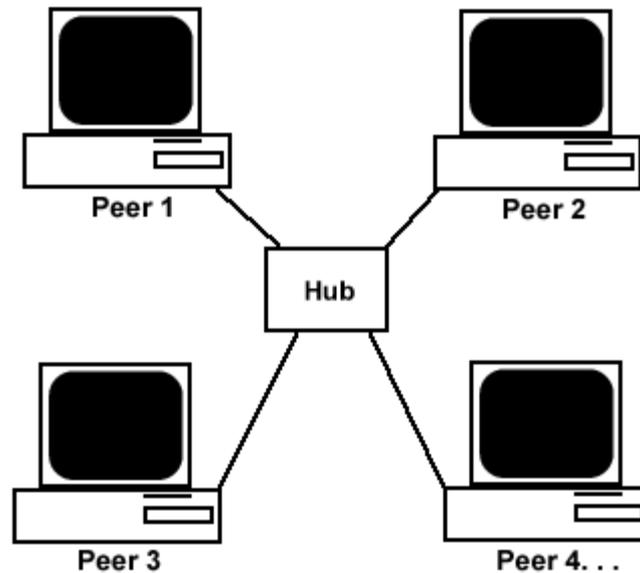


Figure 2.3: Peer To Peer Computing. Taken from (4)

Each peer (computer) accesses services from the other software modules and at the same time provide these services for other peers.

As you can see in Figure 2.3 shown above, all computers are connected to each other and complete the system. Hub connects these computers like in the same network. If we go back to Client/Server model, there is no server module in this picture because each computer behaves as a server. There is one complicated point in Peer to Peer model. Knowledge of needed peer addresses. Each user need to know network address of the application or addresses of subset users of that application. In peer to peer computing, there is no permanent server address of applications. We can see the opposite way in Client/Server architecture. Servers are stable and they do not change. Peer to peer model needs current address information in order to connect successfully.

## 2.3 Why Cloud Computing Matters?

Computing history became with mainframe computers which have so small computing capacity. Data were being transferred by floppy disks. Later, personal computers took the place of old mainframe computers. By being introduced with personal computers, computing powers moved into personal computers. Laptops and smart phones have followed individual personal computers. With every evolution, computing power has been distributed. In today's information technology shift, the trend is to distribute computing power and centralize the storage. At this point, cloud computing comes into picture. Users start to use applications and access their data in a virtual platform, cloud computing. While users can see the results on their screen, processing and storage can be done somewhere in the world. Cloud may not be a popular word in today's technology. But, we have been using this platform for a long time now. Facebook, Google search, Flickr, YouTube are some services that we use already in our daily life without knowing to be in cloud. On the other hand, enterprise companies have 2 major

issues in front of them that can make their choice to cloud platforms. Firstly, having large data centers individually becomes unsustainable for companies. Server capacities are %85 of all the time at idle position. That makes sense, since companies have big economical crisis nowadays. Second issue is IT operations' getting more complex. Technology and applications are evolving. Thus, maintenance, changes, testing takes more time than before and gets more complex. The complexity of this operations yields costs. These 2 issues can be solved by integrating systems into cloud computing platform. As we mentioned about the power usage above, there is %85 of energy that not used. If these traditional server platforms can be moved to cloud, there will be a big impact of cloud for environmental issues. That means information technology will decrease its carbon footprint in planet and will be a part of environmental solution.

Today's goal is to access anything at anytime with any devices. There will be no need to think about operating system or applications. In order to reach this goal, passive devices around us in daily life will turn out all active devices. These Internet connected active devices will be used in taxis, busses, highways and etc. Traffic management in cities, transportation, security and so many daily bases like these will get better solutions with cloud environment. Environmental consideration and time usage becomes two important factors while achieving today's information technology goals. And so far, cloud computing platform seems the best choice of today's technology.

### **3 Architectures of Cloud Systems**

One initial goal of cloud computing systems was based on faster and more effective communications. Users are oriented to focus on their task instead of spending time for application architecture and usage. Thus, time is used efficiently by both sides of cloud. So after all the simplicity that we have mentioned about usage of cloud computing, we will focus on cloud computing background in order to understand how cloud computing provides this features. In this chapter, we will touch on some trends that cloud computing is based on. (6) These trends can be counted as: virtualization, on demand deployment, Internet delivery of services, and open source software. Only looking at these trends, we can easily say that cloud computing systems are not something new, because these concepts are already known and used by other systems in today's environment. So, why is cloud computing new? Cloud computing has changed the viewpoint of how we develop, scale, maintain, invent and pay for the applications that we use.

#### **3.1 Cloud Trends**

This section gives a short summary of the computer technology trends that are used in cloud computing.

##### **3.1.1 Virtual Machines**

These days, virtual machines are a standard attribute of many systems. This technology is not only a trend for cloud systems for sure. What virtualization provides is; virtualization abstracts the hardware from users and let them to be independent from hardware. This independency yields flexibility to access systems anytime. Users do not need to depend on any physical server. By having these dynamic data centers, people access resources anytime. Meanwhile all procedures that are done by physical servers like compute, storage, and network resources are processed by virtual servers. Virtual servers have more advantages than physical servers in this case. Applications can be deployed and scaled rapidly without any preparation. Users are free from physical server procure.

Virtual servers have been general abstraction of system deployment, because, they can be named as a basic denominator interface for service providers and IT. Virtual machines as a deployment objects are beneficial enough for 80 percent of usage. Virtual machines are also satisfying the requirements of applications to scale and deploy rapidly. Putting together virtual machines and virtual appliances as deployment objects is one of the most important key points of cloud computing.

##### **3.1.2 Services that are Delivered Over Networks**

All the services that we have mentioned before make important sense when they are delivered over networks to users. By help of web based interfaces, companies provide usage of their applications. This becomes an important trend of cloud computing. The network can be whether "Internet connection" meanwhile can be "private network" applications for company employees and supplies. Here comes again, best part of the services that are delivered over network is, these services are ready to use anytime from their users. One apprehension of these services is security issue. Companies or providers of these services should be well aware of their communication security. Enterprise sectors use secure socket layer (SSL) for the connection between servers and clients. Thus the communication among cloud computing in enterprise sectors is

processed securely. Security issue of these networks should be architected carefully in order to ensure a flexible environment for its occupants.

### 3.1.3 Open Source Development

Open source development has an important place in cloud computing. Open source development lets people to access cloud environment's basic elements such as appliances and virtual software images. This provides developers to access components easily.



Figure 3.1: Open source in cloud systems. Taken from (6)

We can touch on the effect of open source development in an example. As you can see in Figure 3.1, developers can compose a database system by using MySQL onto operating system called OpenSolaris™. So the appliances are available for applications to create and scale dynamically on any request. The comfort of these open source modules is; they can be used as a component to create complex open source applications. And this makes open source development more important. In other words, tools that are created in cloud environment are reusable for the further level developments by developers.

### 3.2 Cloud Systems Infrastructure

The architecture of cloud computing has some differences comparing to standard enterprise deployment. Architects and developers need to consider when they decide to move their models from standard application models to cloud computing models. There are three different offers in cloud computing infrastructure. These are public, private and, hybrid clouds. The infrastructure of cloud computing is stand on computer hardware and the construction that keep the hardware. This infrastructure contains and processes the virtualization technology which has several different specified servers and machines. Developers of cloud computing will only consider storage, processing power and security issues. Maintenance, server status and other issues of cloud environment will be cloud vendors' problem. Companies which consider moving their system to cloud environment will not be estimating any cost of IT crew to work on hardware issues.

Developers need to consider choosing right environment when they configure their system in cloud. For instance, while public clouds are typically accessible on the Internet and shared to all computer users under basic conditions, private clouds are more for the peculiar use of specified clients.

### 3.2.1 Public Clouds

Public clouds are based on basic cloud computing model. A company provides its service via cloud and the service is available over Internet connection to public users. This service can be an application or data storage. Service provider can offer this service for free or “pay by use” model.

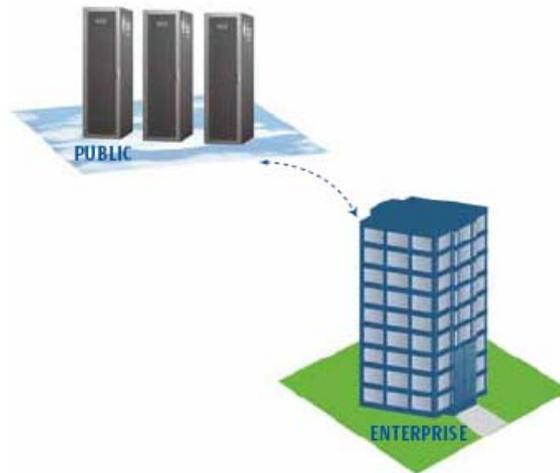


Figure 3.2: Public clouds. Taken from (6)

As you can see in Figure 3.2, companies use public cloud infrastructure to provide their service. Public clouds provide a great level to decrease clients cost with a flexible environment. They offer an efficient way for shared and common resources even they are less securing than the private clouds. There are some cases that make public clouds more beneficial than other cloud computing models. If we have an application that is used by over many people like email, public clouds are the right choice. Users that related to the company can access their emails over Internet connection. When developers want to test and improve their application codes or when they do a collaborative projects, public cloud is the right form to choose for cloud computing. Since public cloud provides this communication for both service provider and users, cloud vendors need to make assure both sides that stored information in the cloud is reliable.

### 3.2.2 Private Clouds

Private clouds are built for the exclusive use of companies that provides the utmost management for data and its security.

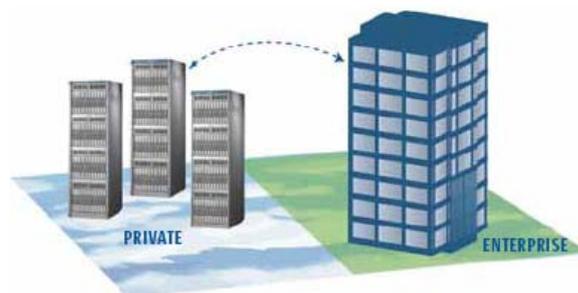


Figure 3.3: Private clouds. Taken from (6)

As you can see in Figure 3.3, by using private cloud, a company can provide and has full control of its own infrastructure. The information that is deployed on the infrastructure is completely under control of owner. So private clouds can be deployed as same place as where company's datacenters. In this private cloud model, instead of requesting a cloud service, company itself creates and manages the cloud. Thus, private clouds give high level control and management of company resources confidentially. Meanwhile, the maintenance and support issues should be taken care of IT support of the company.

### 3.2.3 Hybrid Clouds

Hybrid clouds are based on both private and public clouds.



Figure 3.4: Hybrid clouds. Taken from (6)

The goal of hybrid clouds usage is mostly to provide on demand or external usage capability. Companies use private clouds for their private resources and they use public clouds to handle rapid loads and periodic tasks. Illustration is shown in Figure 3.4. In other words, by using hybrid clouds model, companies spread the work between internal and external cloud providers according to their consideration. One issue of hybrid clouds usage comes up with determining how to distribute process between private and public cloud. For instance, transferring huge amount of data to process from company's private cloud to the public cloud is not a good idea to use hybrid cloud model. This relationship should be well made to get efficient results.

### 3.3 Cloud Systems' Architecture Layers

In this section of the research, we are going to look into architecture layers of cloud computing systems. Cloud computing systems break into three services according to its architecture layers. These are Software as a Service (SaaS), Platform as a Service and Infrastructure as a Service (IaaS) layers (7). In the terms of software services, Software as a Service (SaaS) is the only layer that refers to software in cloud computing. On the other hand, before building a system in cloud, infrastructure is a prerequisite of this layer. In one case, these two layers are connected to each other in respect of architectural layers. In this chapter, we will define Software as a Service (SaaS) and

then look into infrastructure of these services. Third layer, Platform as a Service is a development platform that development tools are already implemented. Platform as a Service model provide these development platforms that can be accessed through web browsers. Cloud computing vendors offer their services into these three categories.

### 3.3.1 Software as a Service (SaaS)

Software as a Service (SaaS) model provides “on demand” software applications as a service. In other words, Software as a Service is kind of web based deployment model. By using SaaS, applications are available through web browsers without depending on any installed software to individual computers. This also makes users of these applications not to care about where applications are hosted or which operating system is used. Users would spend all their time into their task instead of dealing with background. For instance, when a person wants to use an email program, according to his/her operating system, they can use either Apple mail (Mac) or Outlook (Microsoft) to send or receive emails. But there is another option that users do not have to have an installed program on their personal computer to use email program. Simply he/she can use the Gmail that provided by Google servers. By using Gmail, the only need from users is to point their internet browser to Gmail. After they log in, there is the complete email application for free. Gmail has the same functionality as other internal email applications.

As you can see in Figure 3.5, users connect to the cloud that provides software applications. Usages of these applications do not need any requirements. Considering traditional ways can still sound better. For example, users can install an email application to their phone to receive and send emails. But when they want to edit a video clip that they just filmed with friends, users cannot install a professional video editor to their phone. But they can easily connect to a cloud service that provides video editing service and let them to process and share the video.

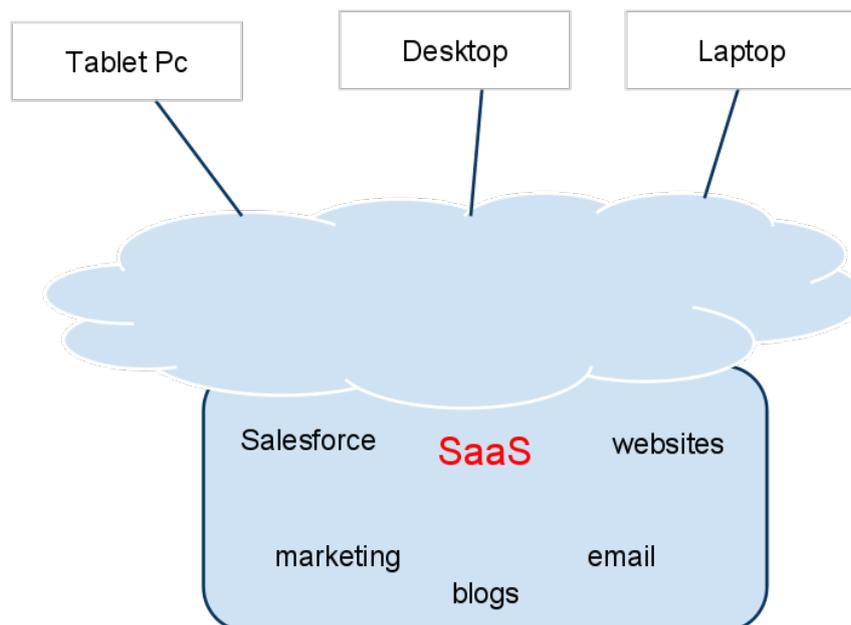


Figure 3.5: Software as a service illustration.

Nowadays, most known example of SaaS is salesforce.com. Salesforce.com is based on customer relationship management model. Salesforce.com provides its customers -

mostly sales persons- to track their expectations. Sales processes and managing their all workflows in their businesses is done by salesforce.com. Again, there is no software needed to use salesforce.com. In order to start with this service, you just need to sign up on their website and start to use this service.

When we come back to SaaS infrastructure, Software as a Service model has its own characteristics. Initially, Software as a Service (SaaS) model is available through web browsers. That means there will be no application that needs to be installed in order to use programs. This is also one characteristic of cloud computing systems. SaaS systems free users to consider about infrastructure or setup investment. SaaS requires payment for the service as long as user continues using these services. And lastly SaaS systems do not require a huge IT demands since users do not have to own or built a system network. SaaS systems require minimal IT need which makes it to use much more easily than the other systems.

### **3.3.2 Infrastructure as a Service (IaaS)**

Infrastructure as a Service model provides computing and storage features over networks as a service. These capabilities of the model are formed by different kinds of servers, routers and etc. These pooled systems are available for high performance processing of applications on demand. Usages of these systems are as same as other cloud models. IaaS vendors own all physical servers and other hardware. Users request them when they need. When users are finished with their work, they just quit using vendor's infrastructure and pay for the time that they have used. Amazon is one of the big players in IaaS platform. They base their system on virtualization. Amazon Web Services (AWS) offers their cloud computing infrastructure to different sizes of companies. Companies request to have their infrastructure and hardware in a specified term and they have full access over all systems that they request.

### **3.3.3 Platform as a Service (PaaS)**

Platform as a Service model provides development environments to code and test web applications. By using this environment, companies or individual users can deploy their system based on PaaS system vendors' framework. Advantage of this model is; users do not have to deal with infrastructure details of the platform while they are building their system. Mostly known example of Platform as a Service model (PaaS) is provided by Google. Google has a service called "Google App Engine" that allows users to code their applications rapidly without any integration needs. There are also some disadvantages of this model. For instance, we have mentioned Google's App Engine. Google allows users to code their applications only based on Python programming language. Other vendors of this model also have similar kinds of limitations. Thus, both producers and customers of this model should consider demands of the other side before they involve PaaS. Although, these limitations can be a down part of PaaS model, on the other side, it frees so many requirements comparing to have your own development environment.

## **3.4 Real World Example**

At the previous chapters, we define the general features and overview of cloud computing framework. In this chapter we will look into a real cloud computing system that Amazon provides to their customers. These technologies that Amazon provides as a cloud computing services will make clear understanding of cloud computing systems.

Amazon has a service called “Amazon Web Services (AWS)” that contains several technologies that based on cloud computing. In these services, Amazon Elastic Cloud Compute (Amazon EC2) and Amazon Simple Storage Service (Amazon S3) sections will be the main and important interest of this chapter.

### 3.4.1 Amazon S3

Amazon S3 gets its name from “Simple Storage Service” (7). As we can understand from its name, Amazon S3 provides persistent storage on demand via web. In order to use this service, Amazon offers an API to access and store objects in buckets of Amazon S3 service. We have mentioned that Amazon S3 is a storage service. The important difference here is; it does not work as a typical file system that typical users are used to. First of all, in Amazon S3, users use buckets instead of directories. By using buckets, users store data in these buckets and give them a unique name. This is because of there is a little difference between buckets and directories. A bucket cannot be put in another bucket hierarchically like classic directories. Buckets in S3 have a unique name and this name is shared for all customers in Amazon S3 system. Considering to local disks, Amazon S3 is slow if the expectations are similar to remote file systems. In fact, Amazon S3 is more than enough fast for Internet deployed systems. In fact, the goal of this system is based on “web access” not “remote file access”. Regarding to Amazon S3 systems' aim, the applications that are stored on Amazon S3 should be written in respect of this system. We will explain this matter later in Amazon EC2 service. As we understand that we do not really store files but objects in Amazon S3, we will touch on some characteristics and usage of this system. The objects that are stored in buckets cannot be larger than 5 GB in S3. Buckets have their own unique name that shared by all users in Amazon S3. So user should consider the name chaos while he/she is naming them. Users have the rights for their buckets so they can make them public for other users to access. And lastly, an Amazon S3 user would need a third party tool to access and manage S3 system since it is not a traditional file system but more primitive. In order to start using Amazon S3, registration is needed. Users should sign up for an account on Amazon's website. After they are done with registration, the system is ready to place data into Amazon S3. Amazon has its deployment either in Europe or United States. The question of which one should be picked is not really important for performance but privacy of the data. If users think that they can have some issues about their data privacy, consideration of the default location of the data can be important. Since users are ready to use S3 service, they have SOAP and REST API which makes S3 services available through. By using these mechanisms, users can manage buckets to create, delete and upload their objects. By starting with S3 service, Amazon also has a command line called “s3cmd” to access Amazon S3. Just to give simple example to use this command line:

```
“s3cmd mb s3://mybucket”
```

This command simply creates a bucket that named by the owner. Since it is an example, we could use “mybucket”. Users can create unique names by adding their domain name at the beginning.

Another feature of S3 service is BitTorrent support. As it is known, BitTorrent is a peer to peer protocol for sharing. In order to manage big amount of data by using peer to peer protocol, Amazon S3 has a good feature in this case.

### 3.4.2 Amazon EC2

Amazon EC2 provides a virtual network for all the virtual servers that users are running inside the network. At this point, interconnection between Amazon EC2 and Amazon S3 services gets into picture. While users run their servers on EC2, they will be using S3 service to store data and machine image. So it is important to know Amazon S3 service before starting with EC2. As we defined Amazon EC2 as a virtual network, EC2 service is more complex system comparing to Amazon S3 service.

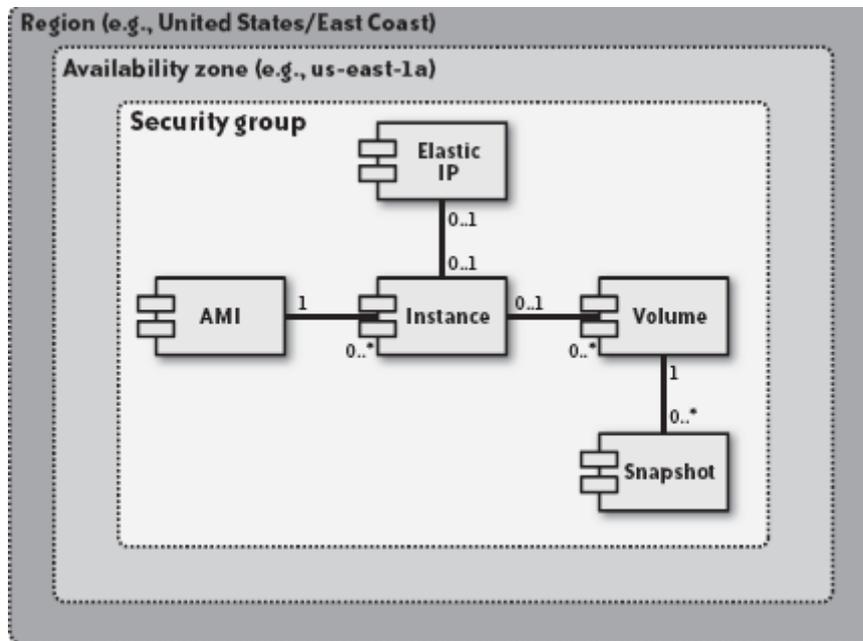


Figure 3.6: Amazon EC2 network. Taken from (7)

Complexity of EC2 service comes from its components. As you can see in Figure 3.6, EC2 modules are interconnected to each other in order to process. Now we will explain EC2 concepts that shown in the EC2 sample picture. Amazon Machine Image (AMI) keeps the copy of users' server images to run their requests. The way of this image is similar with the ghosting operation that allows computer users to create many copy of existing system. Elastic IP address is static IP address that is given users in order to access system. Elastic IP feature lets new users to get IP addresses that are no longer in use. Block storage volume supports block level storage in EC2 system. By using EC2 service users can launch their instances that stored in these volumes. There is also snapshot feature of Amazon EC2 systems. Whenever a user wants, he/she can have a copy or backup of his/her system and this backup data will be stored in their Amazon S3 account. This service is based on web. Thus, accessing Amazon EC2 service can be done through by Amazon Web Services "console" or Firefox Internet browser plug-in. there is a simple example of Amazon EC2 usage. By using command line feature of Amazon we can run this command:

```
"ec2-describe-images -o amazon "
```

This command shows all existing machine images that owned by Amazon. Users of Amazon we services can check these images and pick one. After customizing the image, anyone can build and register for his/her own image on Amazon EC2.

## **4 Cloud Computing: Benefits and Drawbacks**

Most of the web developers and users still have been working on traditional systems. The shift between traditional systems and cloud comes with some important considerations. Advantages and disadvantages of cloud computing system should be known before moving into the environment. This gives useful tips both for cloud computing vendors and clients. After all the knowledge that we have defined, it is time to take look of considerations that are surely important before moving into cloud computing environment.

### **4.1 Benefits**

#### **Usage Based Costing**

By using cloud servers, users pay exactly how much they have used. This is because, they do not own a server but use it. Similar to electricity services, you never own an electricity generator but you use it whenever and how much you want as long as you pay for it.

#### **Processing Time Efficiency**

Instead of having single server or computer for all the work to get processed, cloud computing offers parallel processing power which performs users tasks over several machines and decreases the process time.

#### **Free Space**

Traditional way of having computing infrastructure needs a big place to accommodate. After people are free from all essential hardware by moving into cloud environment, they will not need to consider having a big room of servers anymore. Traditional infrastructures usually need big room capacities in buildings.

#### **Flexibility**

We mentioned about this “renting model” feature in cloud computing instead of owning the hardware. Cloud computing gives more flexibility in terms of resource usage. Companies’ IT team can decide how big and complex systems they need. While they are using the current system, developers process their updates to system. And in case of needed, they can easily speed up the system in real time with the support of cloud computing technology. These changes are totally up to companies’ considerations and how much money they want to spend on technology support.

#### **Scalability**

Scalability is another feature of cloud computing technology. Companies and organizations do not need powerful servers all the time. At the same time these powerful servers are kindly essential requirements for some seasonal periods or specific times. For example, in a specific business sector, a company can schedule their busy times in a month. They process big amounts of data in that specified time with heavy servers and when the work is back to normal capacity, they can switch back to their

normal system. Thus, they do not have to take care of all those heavy servers all the time so cloud computing would be the perfect fit for these seasonal business sectors.

### **Portability**

Nowadays, companies and organizations need their people not only to be close to their central stations but everywhere in the world. With cloud computing, companies are able to give access to their workers and clients as long as they have the access requirements. Accessing the system from anywhere should not be still an issue for today's business environment. In other words, cloud computing abstracts the geography and time problems and this makes the environment more accessible and worthwhile.

## **4.2 Drawbacks**

### **Dependability**

Companies and organizations that are interested in cloud computing, they need to see a reliable provider from cloud computing vendors. This is a very important issue for both sides who want to provide cloud computing resources and customers. After relying on cloud computing, companies will process and place their important functions and data on cloud computing systems. This can be a risk of companies who place the all data including critical functions. At this point, keeping most critical part in companies system and place the rest of data in cloud can be a solution even this needs a well planned and architected system distribution. Another dependability issue comes with the question "What can happen to data?" Either data can be lost or damaged. So companies have to take a deep look into the cloud computing vendors that they will negotiate and how these vendors support their system. Although cloud computing providers are bit reluctant to give geographical information, this problem can be achieved while agreement process. Knowledge of where is data stored gives more confidence to the companies. As an example, there was little time period that Google application servers went down and users could not access to their information. Google is providing this service for free and people usually back up their data while they use Google services. But when cloud computing will be a big deal and several companies will involve in it, in fact, that kind of problem would make a big trouble for everyone.

### **Security**

Security of data has been always an important issue of the computer systems and this problem seems like to be the same in cloud computing area. End users' information and data can easily be hijacked unless companies keep these data carefully with secure firewalls. After data has been lost or hacked, companies would have so much trouble with lawsuit issues and this is not what they want to deal with. In cloud computing infrastructure, we have mentioned about the term "private clouds". Private clouds model can also ease these considerations about security.

### **Little or No Reference**

Cloud computing is still a new area of technology world. Also because of privacy concerns, the companies that providing cloud computing systems do not want to present their publications about cloud computing. In fact, they are still working on their own

problems. So they are unable to report their work at a large scale. This effects other companies which are interested to get involve into cloud computing environment. It seems like small companies will follow these few large companies in order to adopt in cloud computing successfully.

## 5 Visualization Possibilities in Cloud Computing

Until now, a general background about cloud computing was explained. In this section of the report, cloud computing visualizations and possible visualizations are going to be implemented. In order to have a clear understanding of the composition of information visualization and cloud computing composition, we start with a brief description of the field.

### 5.1 Information Visualization

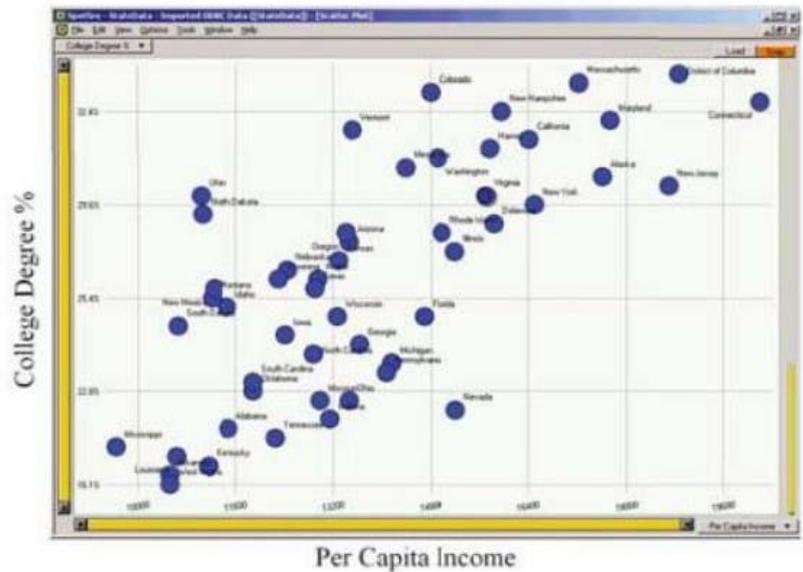
Card, S., Mackinlay, J. and Shneiderman, B describe the visualization as “the use of computer supported, interactive visual representations of data to amplify cognition” (8). With simple words, information visualization is useful for people when they want to ask more meaningful questions about the data which they are dealing with.

There is a basic and meaningful phrase that is related to information visualization topic: “A picture is worth thousand words”. The strong connection between the phrase and information visualization can be explained by a specific example. For instance, Figure 5.1 shows a spreadsheet with the data about 50 states in US together with their citizens, citizens’ college degrees and income of citizens. By just using the data table shown in the figure, to answer the question “Lowest college degree?” is not so hard and actually takes seconds to find. On the other hand, if we change the question a little bit like “Are college degree and income percentage correlated?” Obviously, we think the table is not the best presentation for this question. Or another question like “If they are correlated, is there any outlier states?” is the same issue as before. In that case, the graphical representation of the table is more helpful to answer these questions. So outlier state can be easily recognized on the graphic. Hence, we can imagine how difficult it would be to answer of these questions without this graphic.

To sum up, visualization techniques have an effect in human mind in order to make complicated data understandable at once. For more information in this field, there are many sources on the Internet. I may suggest these books that can be helpful for further reading; Kerren, A et al. *Information Visualization: Human-Centered Issues and Perspectives* (9) and Ware, C. *Information Visualization: Perception for Design* (17).

State	College Degree %	Per Capita Income
Alabama	20.6%	11486
Alaska	30.3%	17610
Arizona	27.1%	13461
Arkansas	17.0%	10620
California	31.3%	16409
Colorado	31.9%	14821
Connecticut	33.8%	20109
Delaware	27.9%	15854
District of Columbia	36.4%	18881
Florida	24.8%	14894
Georgia	24.3%	13631
Hawaii	21.2%	15770
Ideho	25.2%	11457
Illinois	26.0%	15201
Indiana	20.9%	12149
Iowa	24.5%	12422
Kansas	26.5%	13300
Kentucky	17.7%	11151
Louisiana	19.4%	10679
Maine	25.7%	12957
Maryland	31.7%	17730
Massachusetts	34.5%	17224
Michigan	24.1%	14154
Minnesota	30.4%	16101
Mississippi	19.9%	9648
Missouri	22.3%	12989
Montana	25.4%	13212
Nebraska	26.0%	12452
Nevada	21.5%	15214
New Hampshire	32.4%	15959
New Jersey	30.1%	18714
New Mexico	25.5%	11746
New York	29.8%	16501
North Carolina	24.2%	12885
North Dakota	28.1%	13981
Ohio	22.3%	13461
Oklahoma	22.0%	11893
Oregon	27.5%	13419
Pennsylvania	23.2%	14068
Rhode Island	27.5%	14801
South Carolina	23.0%	11897
South Dakota	24.8%	10861
Tennessee	20.1%	12205
Texas	25.5%	12904
Utah	20.0%	11029
Vermont	31.5%	13527
Virginia	20.0%	15712
Washington	30.9%	14923
West Virginia	16.1%	10520
Wisconsin	24.9%	13276
Wyoming	24.7%	17111

(a) A thousand words



(b) A picture

Figure 5.1: A picture and A thousand words. Taken from (9)

## 5.2 Existing Visualizations

Visualization is becoming a key challenge for cloud platform in order to make more sense of data. We have mentioned the importance of information visualization briefly in previous chapter and now, it is time to see the power of information visualization which places in cloud computing. Although the term of “cloud computing” is new in computer based technology environment, there are already a few companies that have undertaken the power of cloud computing in enterprise business. This section will focus on these companies’ cloud computing platforms and their visualizations of process.

### 5.2.1 Force.com and Google Visualization API

Force.com provides its service based on cloud computing. Users can create and start their own business applications in cloud by getting started with the Force.com. Force.com’s application platform comprehends several departments in computer science. For example, database applications with integration and user interface features, client-server system tools such as Java and .NET. Force.com provides this service in a secure and customizable way which can easily support several devices and can be integrated with other cloud applications.

Force.com uses Google's Visualization API that lets users to access structured data from different sources. While accessing these data sources, users have several visualization options that are shown in Figure 5.2. By using the Google's Visualization API, users do not need to move data from cloud platform to any other system to visualize it. Force.com is a real world example for this process. There in an important point of these examples that should be understood correctly. The visualization examples that are explained in this section are not something that you have never seen.

Probably, these charts, graphics and other tools are already been used by simple computer users, even they have not been introduced cloud computing yet. The difference here is: information data and visualization processes are being routed out from individual computers by cloud computing. Thus, there is no extra data transfer needed in order to visualize information data. Visualization process of personal data can take short time in individual computers. But, a big project with plenty of attributes can be a problem when it comes to either having a powerful processor or enough time to get results. This makes implementation of visualization more than a “need” in cloud computing environment.

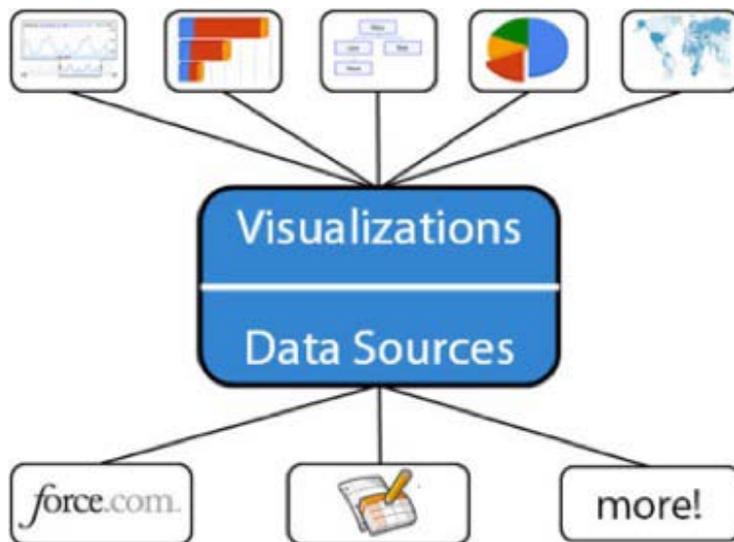


Figure 5.2: Multiple data sources and its visualizations. Taken from (10)

Now we will take look at Force.com’s visualization architecture. As the diagram is seen in Figure 5.3, user application is composed by different components. Blue colored area in the diagram is the part that users create their applications and place information data. Red colored area is the visualization part of the applications that provided by Salesforce.com (Force.com uses Salesforce.com’s system structure at background). The red area will process and query the data that users provide and display visualization outputs. The components users need to use are already provided by Force.com so there is no need to configure system to get started by users. As it mentioned previously, this advantage is in the definition of cloud computing.

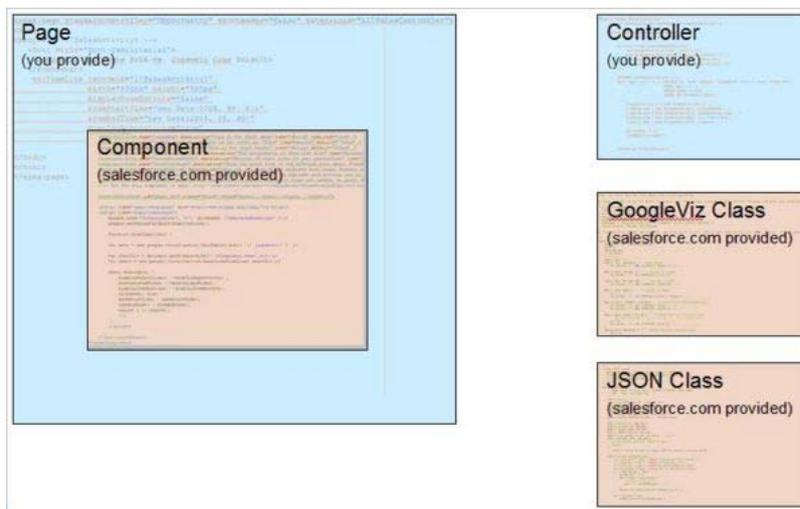


Figure 5.3: Components of Force.com platform. Taken from (11)

At this point, to interact with the third component (Google) may sound unsecure but actually users' personal data is never sent to other parties. You can see a simple demonstration of how these systems work in order to visualize data in Figure 5.4. In the first step, user requests his/her Visualforce page and gets response from Force.com. In the second step, only visualization request is sent by system to Google visualization API and data is rendered in user's browser.

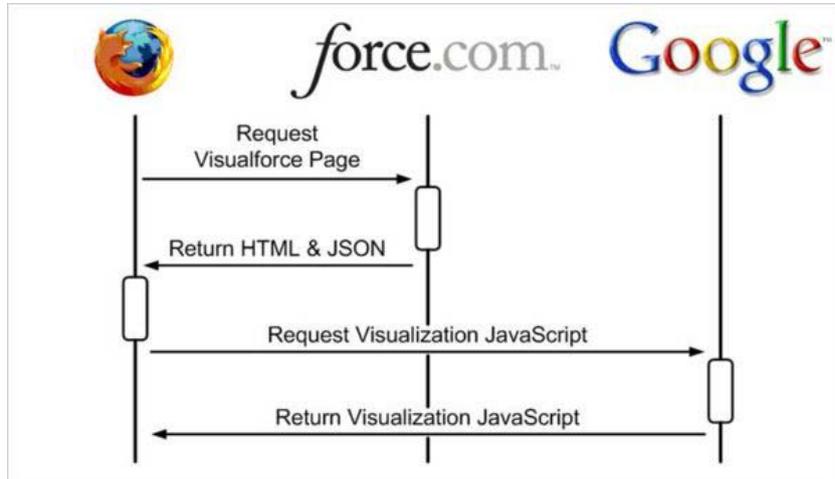


Figure 5.4 Interaction between user, Force.com and Google during the visualization.  
Taken from (11)

By having visualization outputs, user data will have more meaningful perspective than before. So that users can analyze their results and compose reports and manage their business according to these visualizations. In Figure 5.5, the final visualization of the 2 dimensional data table in Visualforce page is shown. The figure illustrates timeline chart for a user's sales activity.

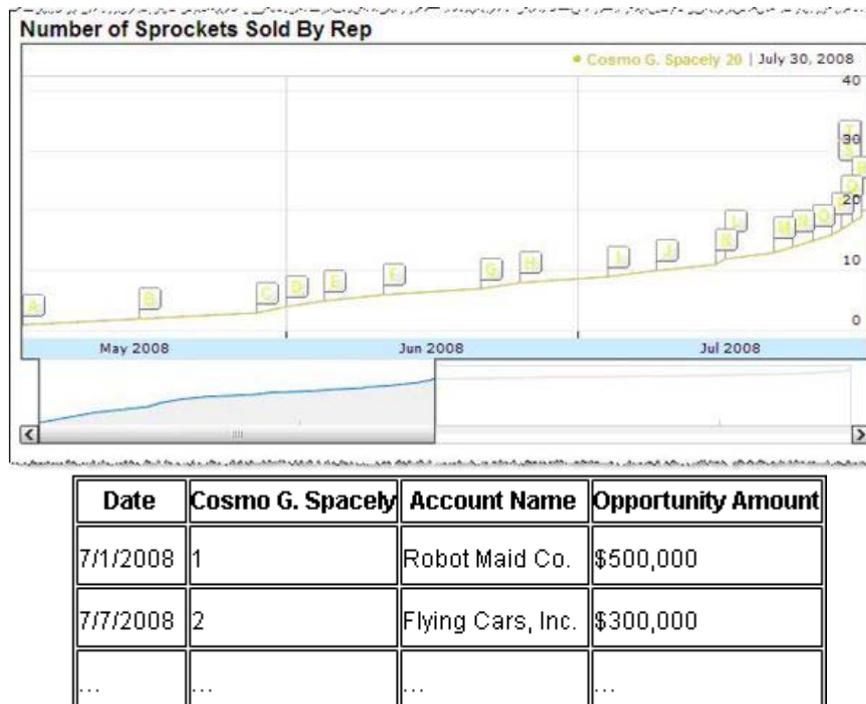


Figure 5.5: A sample of Visualforce page and its source two dimensional table  
Taken from (11)

Additional data that is added every time into user's workspace will also effect his/her visualization page so users do not need to recreate visualization outputs with the updated data.

Force.com example gets more interest by both individual and enterprise business area but information visualization in cloud computing offers a good advantage for other departments such as non-profit organizations, education, and for the people who can make their life easier with cloud computing opportunities. The next example we are going to take a look at is Flickr.

### 5.2.2 The World's Eyes Project

Flickr is another good example of cloud computing as a service. Flickr web service provides users to share and embed their photographs online. "The world's eyes" project provides a map visualization of geotagged pictures based on Flickr images. The picture files that are uploaded to Flickr service contain taken time and location information. The project analyses these information to track each Flickr user that has visited Spain. Visualization results of the project give a general idea of such as "most visited places" and "where are the most exciting places". Additionally, by using these visualizations, people can discover good perspectives of specific regions. An example visualization output can be seen in Figure 5.6.

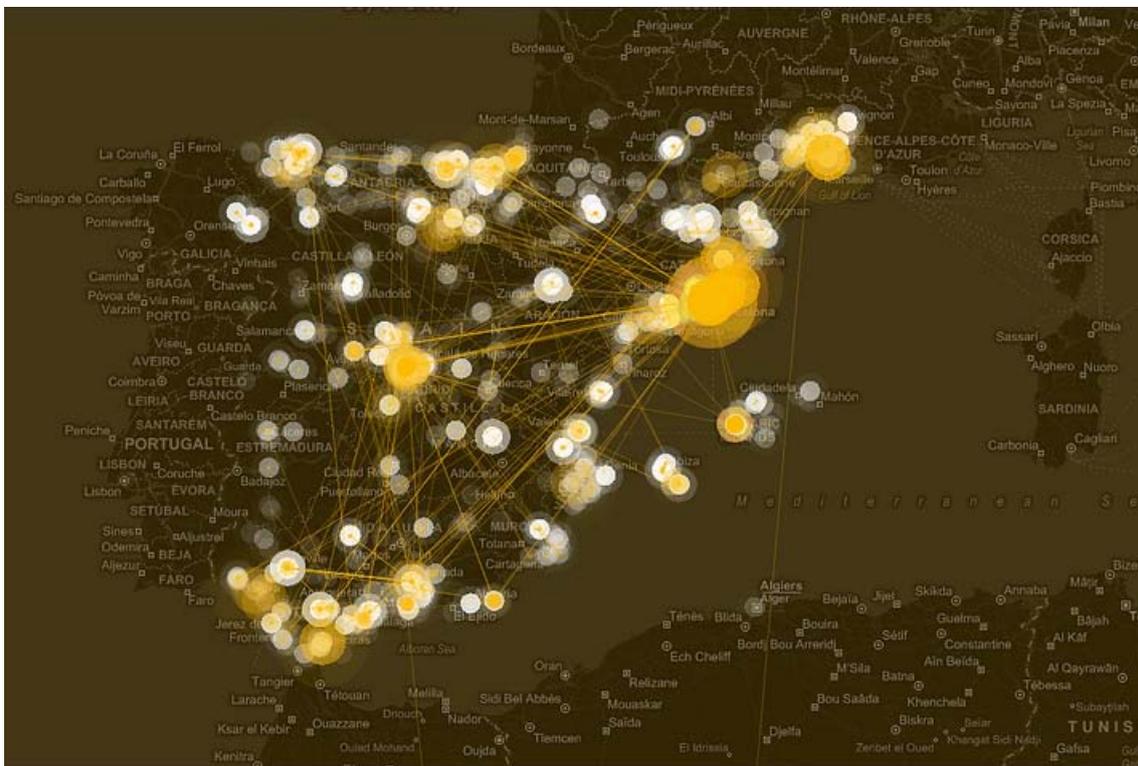


Figure 5.6: The World's Eyes project. Taken from (16)

The project makes it easier to understand and analyze regions' popularity and documentation by help of cloud computing and information visualization. This project is not only popularity map but can also be zoomed in to anywhere in the map and see the pictures that are taken from that place. Figure 5.7 shows another screenshot from the project that the pictures are falling down into places.

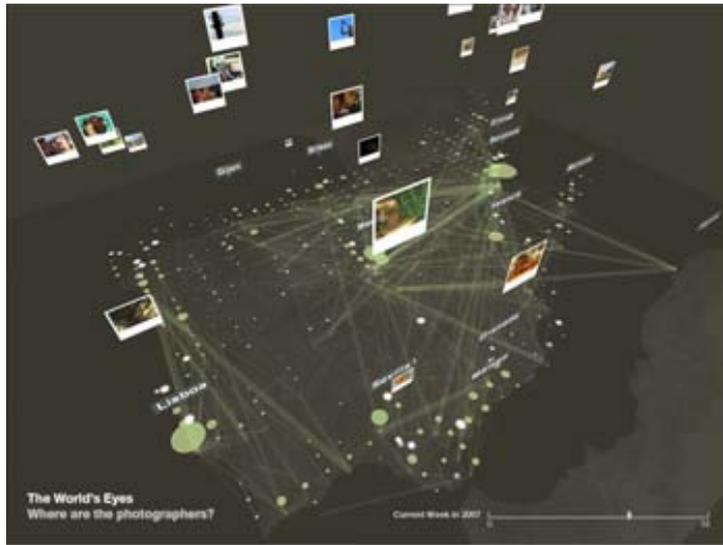


Figure 5.7: The World's Eyes Project shows where the photographers and their photos are. Taken from (16)

### 5.2.3 Newsmap

Newsmap is visualization project of relationships between data and news media. Project is based on Google news service which groups news according to their content and tags. Newsmap provides news by treemap algorithm. An example page from the project can be seen in Figure 5.8. News has its own cell with different sizes. It determines the amount of articles that are related to the title. So that users can see popularity of topic in specific times and customize settings according to countries or interests such as Germany, United Kingdom, sports, business.

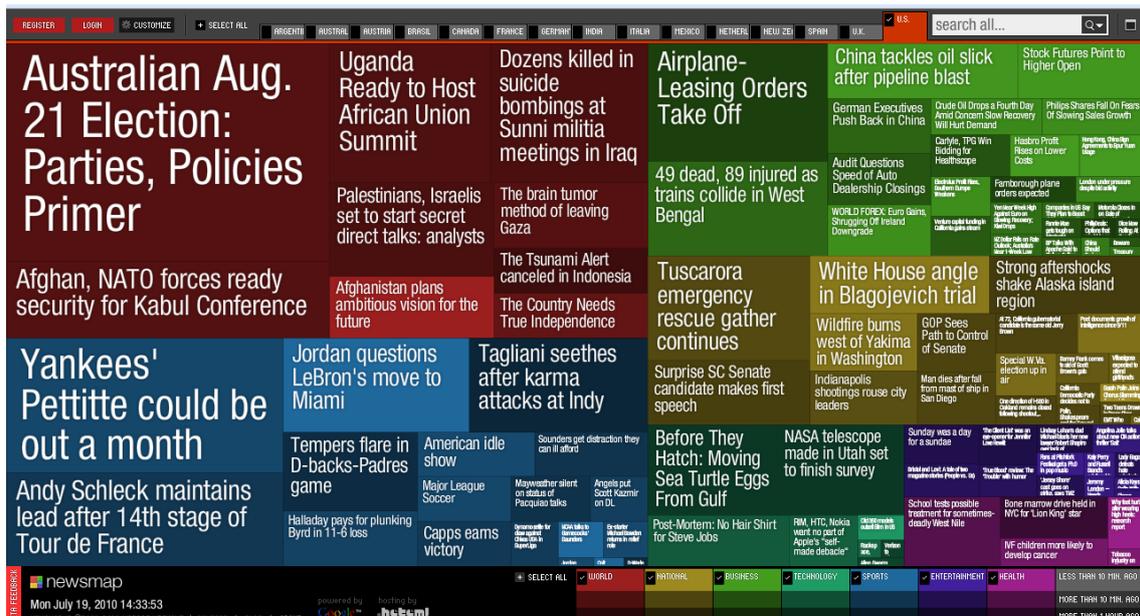


Figure 5.8: Newsmap Google News visualization.

Since we deal with bunch of overloaded information by Internet, TV and magazines every day, a tool that divides information into sensible and recognizable way like this visualization paradigm, will make life easier, simple and more organized.

### 5.2.4 Fidg't

Social networks trend is popular since we share our music, personal profiles, and pictures with all the people on planet. Since we have a virtual people network around us, it would be useful to know what the network is formed with. Fidg't software visualizes users' social network by their tagging activities based on two cloud platforms, Last.fm and Flickr services. It also uses chat networks such as Yahoo and MSN messenger. The algorithm behind the system analyses user's networks interests via their activities. So users keep track of their networks general look and ask specific question by using tag words to have better knowledge of their network. An example screenshot is shown in Figure 5.9.

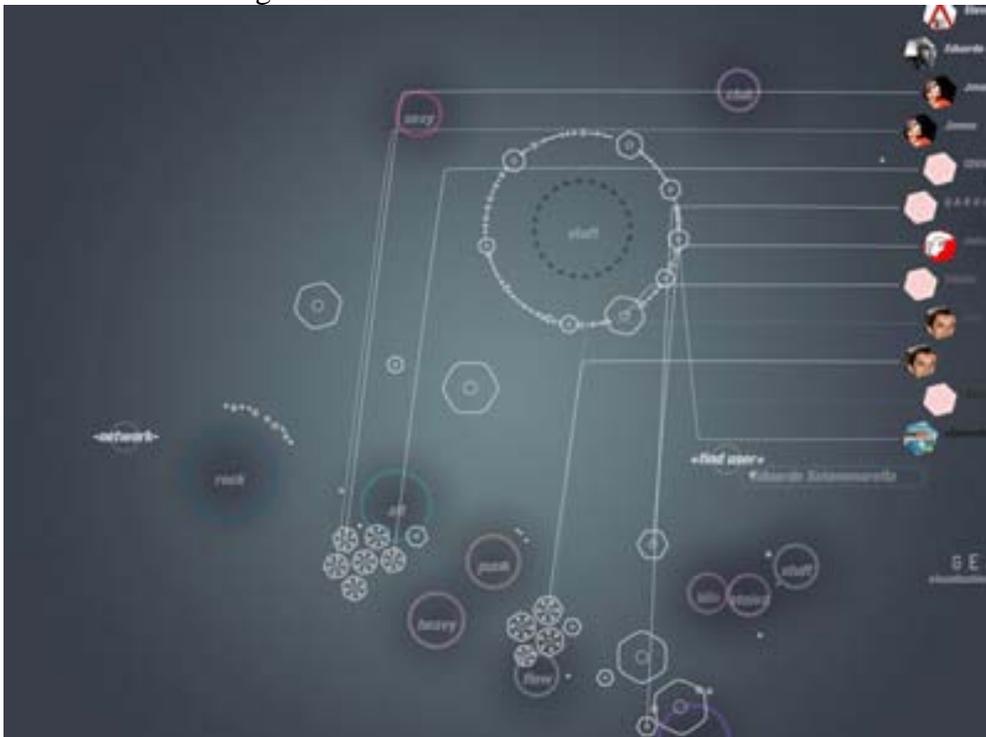


Figure 5.9: Fidg't software visualizes network users around customized magnet tags.  
Taken from (15)

This platform demonstrates not only network activities but also provides cross platform interaction. Any user can discover a new photo or music album in his/her network by using Fidg't and share or spread it by using the same platform. They do not need to jump over other software in order to do this. Additionally, users can keep track of their network with Fidg't on mobile phone since they do not need hardcore process to do all this.

### 5.3 Visualization Possibilities

Cloud computing gains its popularity in all kind of organisations with different sizes. This makes organisations' IT platform to replace both software and hardware from traditional model into cloud computing model. As a matter of fact, the new platform that based on cloud computing would have its own characteristics. Visualization possibilities of cloud computing is the topic of this section that will explore the visualization ideas which can be performed in architecture layers of cloud computing such as Infrastructure as a Service(IaaS), Platform as a Service(PaaS) and Software as a Service (PaaS).

First of all, it is better to have a look of cloud computing infrastructure in order to have a brainstorming of visualization possibilities. People who get cloud applications as a service are mostly end users. Their need of visualization possibilities will be more into software as a service (SaaS) layer. Thus, visualization needs to be integrated into their desktops, mobile devices, and browsers for cloud applications. This is end users aspect of visualization. But when we look at the whole background picture of cloud platform as it is shown in Figure 5.10, there are two other essential layers which should take attention from cloud vendors and IT. These two layers are; Platform as a Service (PaaS) and Infrastructure as a Service (IaaS).

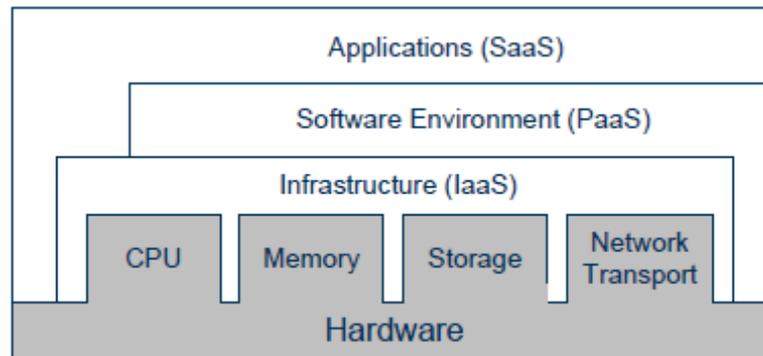


Figure 5.10: Cloud computing architecture layers. Taken from (13)

First layer of the platform is named infrastructure as a service that provides hardware for cloud platform with virtualization feature. So there is one complete infrastructure and multiple users of this infrastructure. As we mentioned before, the payment of cloud infrastructure is based on pay-by-use model. With this model, clients only pay for how much they use. Before clients start their businesses or projects by using cloud infrastructure, it is not really easy for them to presume how much storage and computing power they need. Thus, visualization of their monthly system usage from the vendor might give them useful tips. This demonstration may include monthly and daily processing power usage, storage service and users' activity. For example: a timeline graphic that shows computing usage with actual frames can be useful tools for clients. It would make sense to see utilization of the system. These visualization output would help clients to decide if they need to switch for a bigger system or not. Since cloud computing provides enormous flexibility, clients can even detect and set the specific hours when their servers are overloaded. In these time periods, they can order to get more processing power from the vendor. Visualization of the system usage will help clients with cost savings and instant system access. Animoto web site uses Amazon cloud services to run their web application. At the beginning they had 50 Ec2 instances. After 3 days of their service, they had to upgrade the system to 3500 Ec2 instances to handle the traffic. You can see the server usage in Figure 5.11. At the end of this timeline graphic, user demand goes down and then they decreased the Ec2 need to 1200 instances. However, if they would have tried to handle this uneven traffic issue with traditional ways, there would be a lot of big problems that are waiting for them. Extra servers, storage and disaster recovery plans probably would not be ready in this short time period. So their application would not be available for few days and etc.

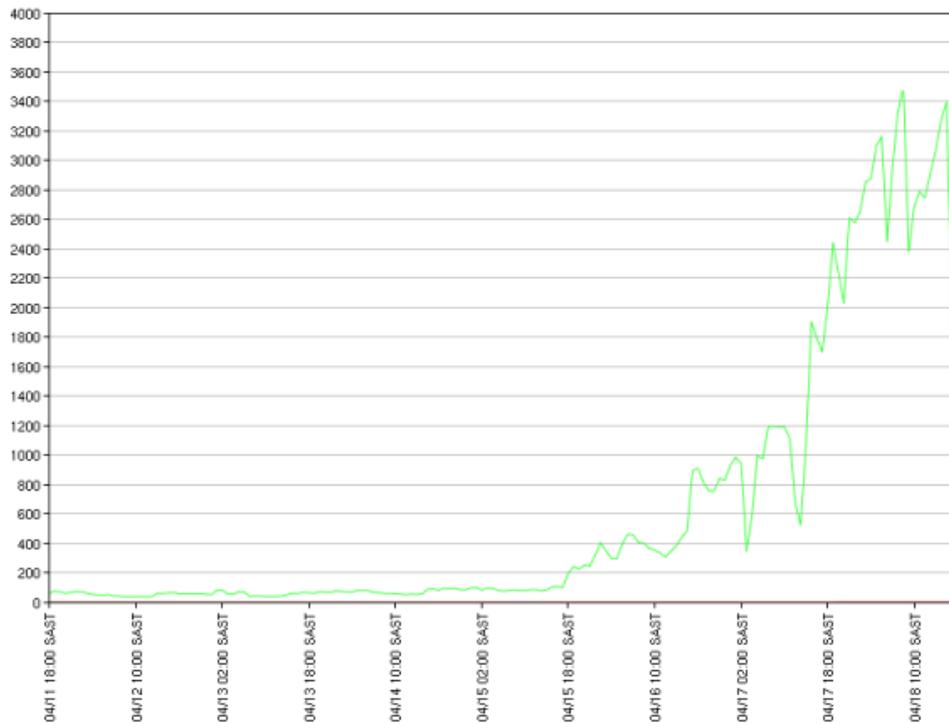


Figure 5.11: Daily Amazon Ec2 usage of Animoto web application. Taken from (12)

For example, by using timeline visualizations, a bank company that uses hybrid clouds can specify the days which clients overload online banking system such as last days of paying taxes or bills. By customizing the system according the visualization graphic results, will help the clients to have access their accounts on demand without any failure and will let company to save costs and have reliable system reputation.

Platform as a service (PaaS) provides operational environment for application development. A complete PaaS example can be given as the Google App Engine that provided by Google. Based on Google’s frameworks (Python PL), users create their application projects and run them at the same environment. By using PaaS, testing, developing and maintaining of the applications are processed at the same platform and there will be no additional cost for separate environments and additional time for moving projects to any other servers. During project development, it is important to see what has been done by project crew because software development is not only about writing lines of codes. While development process, quality of the software should be measured and outlier parts should re-design or re-written according to measurement results. As it is said by Tom DeMarco “You cannot control what you cannot measure”. At this point, information visualization tools have a good opportunity to get into play. Visualizing application development period would give useful tips for the quality of projects. How these software quality metrics can be visualized? Big project developments contain thousands lines of code, a lot of classes and packages. This kind of projects cannot handle by one person. There will be a project team from different regions and the project will be divided into parts that every team member has a part of the work. Visualization of class interaction, package interaction and project evolution will provide strong ideas to analyze and re-design of the work in case of needed. You can see a simple idea of visualization progress in Figure 5.12.

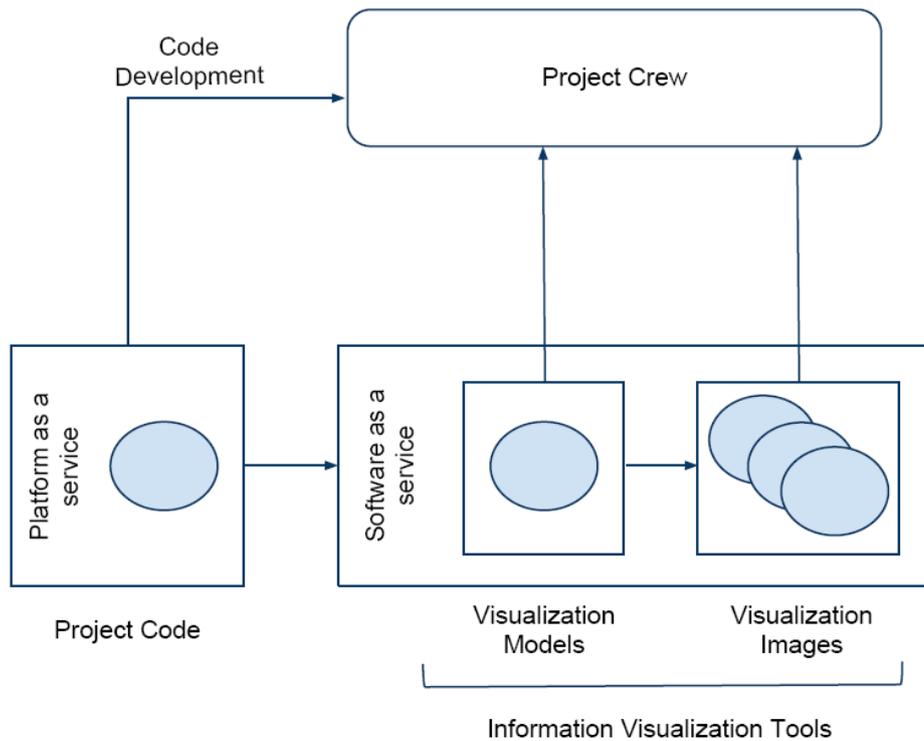


Figure 5.12: Composition idea of information visualization tools and PaaS platform

Implementation of the visualization application can be placed software as a service in cloud. The application provides visualization graphics of the project structure via information visualization tools. During the development, the project will be connected to visualization application to analyze the structure of work so they will get visualization results about complete project. There will be several different requests of visualization perspectives from developers according to their aims. For instance, outlier classes need to be discovered during the development which should be re-written in order to have well designed software. So, implemented visualization software should provide different visualization models.

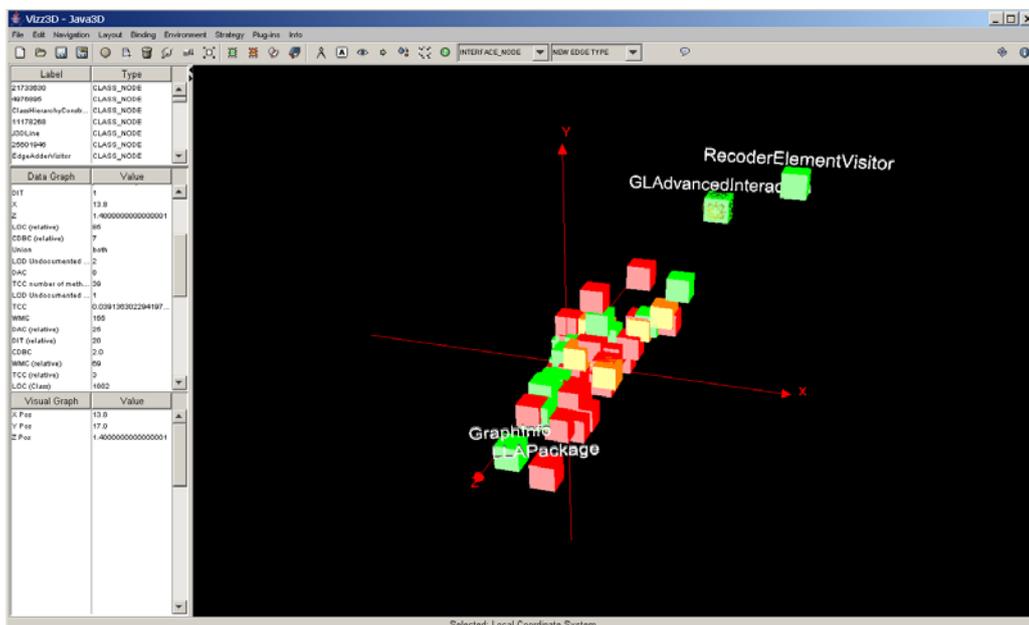


Figure 5.13: Visualizations of outlier components. Taken from (14)

Thus, project structure and development steps can be visualized by different graphics. Figure 5.13 shows outlier classes in the system for maintainability. Green classes are marked as well designed classes but the classes with red color are outlier classes.

“On demand” deployment will let workers see their teammate’s works and their effects on the system development instantly. Visualizing class diagrams of the project will be helpful to see the dependencies, coupling between classes, cohesion in methods. If any failure can be realized and re-designed during the development, the cost of the project would be decreased and time usage can be used efficiently. Using a traditional way, for example; user downloads all project files to another computer, process visualization and do this procedure several times will be leeway and slow development process. These design decisions will be updated at cloud platform instantly and let all the team about new status of the project. As a matter fact, in today’s technology environment, the developers of enterprise companies are participating projects from all over the world. Changes of dynamic visualization can be a successful development guide for coders of this platform.

Mostly known examples in cloud computing platform are based on Software as a Service (SaaS) models. The “Live Campus” idea can be perfectly used by the help of information visualization. For instance, Linnaeus University has its own campus and students have their accommodations and studies in there. They attend events in campus such as conferences, meetings, entertainments that are created by student nations, Växjö City municipality and university. Students usually get their events information via Facebook (cloud computing) service, university email, or event walls. It is possible that students can miss something because they did not get information about it. Live campus idea is structured on a map of Linnaeus University campus. The Campus Live project synthesizes information from cloud environment such as shared calendars of lectures and student nations’ event database into visualizations. These visualizations will let everyone in campus that is connected to Live Campus via their computers or mobile phones to get information about what is going on in campus right now and upcoming events. Växjö campus is a quite big area to memorize where all the buildings are placed. Campus Live will let students know when their upcoming event is and where it is placed. Sometimes lecture rooms and their time can be changed in last day due to some problems. As soon as the lecturer updates shared course calendar, related people will see where the new place is in their mobile while they are going out from apartments for instance. A blueprint of the Campus Live idea and each building that is placed in campus can be seen in Figure 5.14. Live Campus idea stores information of students’ personal calendar for instance, lecture and examination times and where they are placed. A student that is registered to the course “Database Theory” sees his/her course information for today on the campus map. Course information here is synchronized with shared calendar of the lecture created by the instructor. Any updates on this calendar will have an effect on Live Campus immediately.

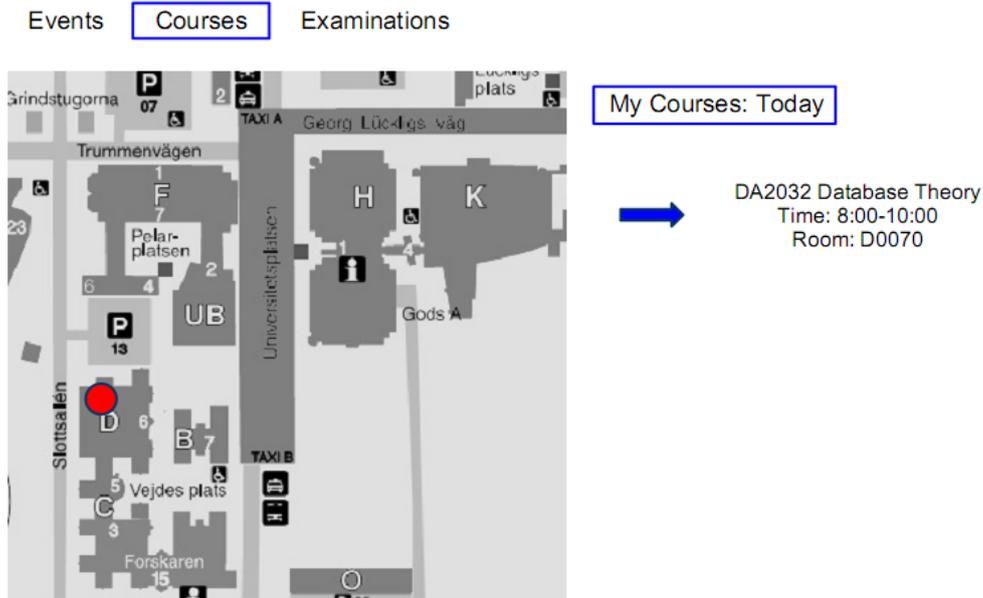


Figure 5.14: Blueprint of Campus Live idea

Course example is only one point of Campus Live idea. This idea can be developed with additional technologies in order to have better information of campus and people. For example, by including GPS information in the map, users can see where the crowd is placed in campus. This GPS information will be taken from people that use mobile phones with GPS in campus. So users can decide which place should be better to go or not, such as restaurants, student pubs, entertainment activities and so on. GPS can be also implemented in busses in Växjö City that people can keep track of transportation options. People will access Campus Live application via Internet connection. PC users can use broadband connection to access this service easily. Mobile users can use Internet connection that is provided by telephone companies or wireless connection in campus. Wireless range of campus area can be also shown on the map additionally.

Infrastructure of Campus Live will be based on cloud computing. Servers, databases and additional hardware can be provided by cloud computing vendors for instance Amazon EC2 and S3. There will be authentication levels to access and modify information in servers. Users can create their own calendar in Campus live and at the same time add courses' or student nations' public calendar to their Campus Live application.

## **6 Conclusion and Future Work**

This chapter gives a conclusion about the final result of the project and future work.

### **6.1 Conclusion**

Computing is a computer science process of improvement of software and hardware. In this report, I explained the new technology platform that called cloud computing and visualization possibilities of cloud computing. As I mentioned before in Chapter 1, cloud computing has quickly stepped into information and communication technology so it came with its questions such as what is cloud computing? What kind of architecture cloud computing has? Using this new environment good or bad? In the second and third chapters, I focused to answer these questions. After making a statement for cloud computing in second chapter, I investigated the main goal for this report “existing cloud computing visualizations and visualization possibilities”. In the fourth chapter, cloud computing visualizations are given with real world examples. Finally, visualization possibilities from different perspectives are researched.

After all goals are succeed, we have a clear understanding of cloud computing and brainstorming ideas for possibilities in order to visualize abstract data in cloud.

This work was a good experience for me to learn about cloud platform usage and the importance of information visualization for the human side of information and communication technology. I believe this experience and ideas will be so useful for my future studies.

### **6.2 Future Work**

A detailed survey with many examples was presented in this report. But when we talk about computing architecture besides a platform, the subsections that I investigated in this report can be more focused. Day after day, this new platform will be improved by developers of enterprise. So, new opportunities, new ideas can place in the main picture of cloud computing. With new attributes of cloud computing, visualization possibilities will increase and will need fresh and beneficial ideas.

## 7 References

1. Pew Internet & American Life Project. Horrigan J. Use of Cloud Computing Applications and Services. September,2008. [accessed 13 Jan 2010] Available from: <http://www.pewinternet.org/Reports/2008/Use-of-Cloud-Computing-Applications-and-Services.aspx?r=1>
2. Miller M. Cloud Computing: Web-Based Applications That Change the Way You Work and Collaborate Online. Safari books online LLC; 2008. p. 14-15.
3. Menken I; Blokdijk G. Cloud Computing - The Complete Cornerstone Guide to Cloud Computing Best Practices. The Art of service; 2009. p. 18.
4. Jelen T. and King R. Networks 101: Peer-To-Peer Networks: The basics of peer-to-peer networking. TechSoup Global:2003 [accessed 20 Dec 2009] available from: <http://www.techsoup.org/learningcenter/networks/page4772.cfm>
5. Gilberto J. Perera. 10 reasons Why Cloud Computing is the Wave of the Future;2009 [accessed 13 Jan 2010] available from: <http://laptoplogic.com/resources/10-reasons-why-cloud-computing-is-the-wave-of-the-future>
6. Sun Microsystems. Introduction to Cloud Computing Architecture. 2009:06 [accessed 13 Jan 2010]. Available from: [www.sun.com/featuredarticles/CloudComputing.pdf](http://www.sun.com/featuredarticles/CloudComputing.pdf)
7. Reese G. Cloud Computing Architectures. Safari books online, O'Reilly;2009.
8. Card, S. and Mackinlay, J. and Shneiderman, B.: Readings in Information Visualization - Using Vision to Think, Morgan Kaufmann, 1999
9. Kerren, A.; Stasko, J.; Fekete, J.-D.; North, C. (Eds.). Information Visualization: Human-Centered Issues and Perspectives. Springer; 2008.
10. Developerforce. Force.com as a Google visualization data source 2009 [accessed 3 may 2010] Available from: [http://wiki.developerforce.com/index.php/Force.com\\_as\\_a\\_Google\\_Visualization\\_Data\\_Source](http://wiki.developerforce.com/index.php/Force.com_as_a_Google_Visualization_Data_Source)
11. Developerforce. Google visualizations. 2009 [accessed 3 may 2010] Available from: [http://wiki.developerforce.com/index.php/Google\\_Visualizations](http://wiki.developerforce.com/index.php/Google_Visualizations)
12. Amazon Web Services Blog. Animoto - Scaling Through Viral Growth. 2008. [accessed 20 June 2010] Available from: <http://aws.typepad.com/aws/2008/04/animoto---scali.html>
13. Ohlman B.; Eriksson A.; Rembarz R.; What Networking of Information Can Do for Cloud Computing. 2009 18th IEEE International Workshops on Enabling Technologies: Infrastructures for Collaborative Enterprises.
14. Panas T.; Lincke R.; Löwe W. THE VIZZANALYZER HANDBOOK. October 2005. [accessed 10 may 2010] Available from: [www.arisa.se](http://www.arisa.se).
15. Fidg't Visualizer. 2007. [Accessed 1 July 2010] Available from: <http://fidgt.com/visualize>
16. MIT Senseable City Lab. The world's Eyes Project. 2009. [Accessed 20 May 2010]. Available from: <http://senseable.mit.edu/worldseyes/index.html>
17. Colin Ware. Information Visualization: Perception for design. Morgan Kaufmann Series in Interactive Technologies; 1999.



# **Linnæus University**

School of Computer Science, Physics and Mathematics

SE-351 95 Växjö / SE-391 82 Kalmar

Tel +46-772-28 80 00

dfm@lnu.se

Lnu.se