Analysis of Privacy Issues in Cloud Computing

Neetu Anand
(Research Scholar, Lingayas University)
Assistant Professor (Dept. of Computer Sc.)
Maharaja Surajmal Institute, Delhi
neetuanand@msi-ggsipu.org

Dr. Tapas Kumar
HOD, Computer Sc.Engg.
Professor of LU
Lingaya's University (LU)
Kumartapas534@gmail.com

Abstract— Recent evolutions in information technology have led to a more distributed computing environment, while also reviving the utility of centralized storage. The growth in high-speed data lines, the falling cost of storage, the advent of wireless high-speed networks, the proliferation of handheld devices that can access the web – together, these factors mean that users now can store data on a server that likely resides in a remote data center. Cloud computing premise is very similar in that it provides a virtual computing environment that's dynamically allocated to meet user needs. But How much secure is cloud computing environment is a big challenge. This paper, focused on the security issues in cloud computing and its main objectives to describe cloud computing and all major security risks and issues related with it.

Keywords- *Cloud computing, Cloud Scalability, Privacy risks , Risk mitigation, Cloud services.*

*****

I. INTRODUCTION

The term "cloud computing" has gained currency in the information technology world in the past 18 months as a way to describe the ongoing evolution in how people access and manage digital information.

The term "cloud" originates from the telecommunications world of the 1990s, when providers began using virtual private network (VPN) services for data communication. VPNs maintained the same bandwidth as fixed networks with considerably less cost: these networks supported dynamic routing, which allowed for a balanced utilization across the network and an increase in bandwidth efficiency, and led to the coining of the term "telecom cloud."

Users can then access the data from their own computer, someone else's desktop computer, a laptop that wirelessly connects to the internet, or a handheld device. Users face new challenges as they try to manage their data that might be stored in a variety of devices. This is where cloud computing enters the picture.

When talking about a cloud computing system, it's helpful to divide it into two sections: the **front end** and the **back end**. They connect to each other through a network, usually the Internet. The front end is the side the computer user, or client, sees. The back end is the "cloud" section of the system.

Cloud computing is emerging as a model in support of **"everything-as-a-service"** (XaaS).*Cloud computing* is a paradigm that focuses on sharing data and computations over a scalable network of nodes. Examples of such nodes include end user computers, data centers, and Web Services. We term such a network of nodes as a *cloud*. An application based on such clouds is taken as a *cloud application*. Below shown are the models for conventional data centre in Figure 1 and cloud centric model in Figure 2.

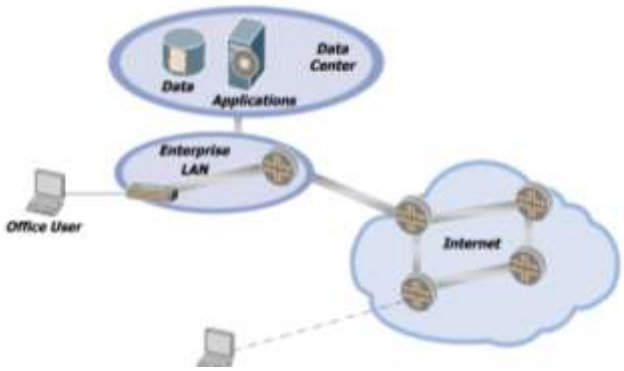

Fig 1. Conventional Data Centre

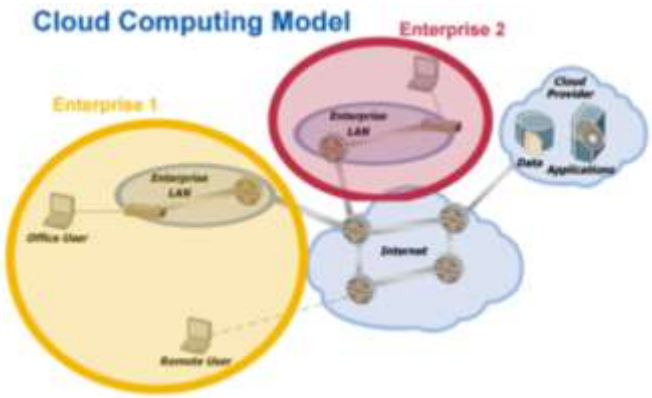

Fig 2. Cloud Centric Model

II. MAJOR CLOUD SYSTEMS

Cloud Systems are categorized as:

- Software-as-a-service

- Platform-as-a-service
- Infrastructure- as-a-service
- Human-as-a-Service

These system categories are described as follows:

SaaS

In the SaaS model, the user buys a subscription to some software product, but some or all of the data and code resides remotely. For example, Google Docs offers an alternative to Microsoft Office that stores documents on Google's server. It doesn't keep any code on the client machine, even though some code might execute on the client temporarily. For example, Google Docs relies on JavaScript, which runs in the Web browser. In this model, applications could run entirely on the network, with the user interface living on a thin client.

PaaS

From the consumer's viewpoint, PaaS software probably resembles SaaS, but instead of software developers building the program to run on their own Web infrastructure, they build it to run on someone else's. For example, Google offers Google App Engine, a service that lets development organizations write programs to run specifically on Google's infrastructure.

IaaS

Similar to PaaS, IaaS lets the development organization define its own software environment. This basically delivers virtual machine images to the IaaS provider, instead of programs, and the machines can contain whatever the developers want. The provider can automatically grow or shrink the number of virtual machines running at any given time so that programs can more easily scale to high workloads, saving money when resources aren't needed.

HuaaS

Some services rely on massive-scale aggregation and extraction of information from crowds of people. Each individual in the crowd may use whatever technology or tools he or she see fit to solve the task.

III. LAYERS IN CLOUD COMPUTING

There are several recognized layers in cloud computing. The vendors in these layers have very different service offerings and operating models. Some vendors concentrate on building and maintaining a huge data center, while others concentrate on building a user friendly and feature-rich application. The layers, from bottom to top, are: Infrastructure, storage, platform, application, services, and client [1].

Infrastructure

At the bottom is the infrastructure of the service, or the platform virtualization. We get the kind of server environment we want. This is the basic offering; customers still need to handle the server, all software installation and maintenance by themselves. The cloud computing infrastructure does differ from traditional Maintaining the Integrity of the Specifications hosting services because of scalability, and pay-as-you-go pricing. A start-up company might be very interested in getting the scalability, and in not paying for the time they're not using the service. It is convenient, especially if you're trying to grow the traffic on your Web application but don't know how soon, or how well, you'll succeed.

Storage

With the storage layer, we get a database or something similar, and pay per gigabyte per month. A storage layer is nothing new or special, except for the full stack of services. It is of course, vital. There are several possibilities for storage. Some are traditional relational databases, and some are proprietary solutions such as Google's Bigtable or Amazon's SimpleDB.

Platform

The platform layer has solution stacks such as Ruby on Rails, LAMP, or Python Django. Now things start to get interesting. That fictitious start-up company doesn't have to deal with the installation of server software, or keep their versions updated, because that comes with the service. They can focus on developing and marketing their application.

Application

The application layer contains applications that are offered as services. The most famous examples are probably Salesforce.com and Google Docs, but there are hundreds if not thousands of (real) applications that can be purchased as services. Popular Web applications such as Facebook, Flickr, and LinkedIn are cloud services. In these cases, the customer probably doesn't know if the application is run in a scalable data center, in an ordinary hosting service, or in the service providers basement. But, that isn't a concern or problem for the customer who needs to use the application. This layer is probably the most visible part of cloud computing. It emphasizes the benefits that can be seen by customers.

Services

The services layer contains interoperable machine-to-machine operations over the network. The most prevalent example of this layer is Web services. Other examples include payments systems, such as Paypal, and mapping services, such as Google Maps and Yahoo Maps.

Client

At the top of the stack is the client layer, which contains the users of the cloud systems. Clients are, for example, desktop users (thin client or thick client), and mobile users.

IV. CLOUD SCALABILITY

Scalability is a quality feature of the computing cloud. It has at least two dimensions, namely *horizontal cloud scalability* and *vertical cloud scalability [2]*.

- *Horizontal cloud scalability* is the ability to connect and integrate multiple clouds to work as one logical cloud. For instance, a cloud providing calculation services (*calculation cloud*) can access a cloud providing storage services (*storage cloud*) to keep intermediate results. Two calculation clouds can also integrate into a larger calculation cloud.
- *Vertical cloud scalability* is the ability to improve the capacity of a cloud by enhancing individual existing nodes in the cloud (such as providing a server with more physical memory) or improving the bandwidth that connects two nodes. In addition, to meet increasing market demand, a node can be gradually upgraded from a single power machine to a data center.

Scalability should be transparent to users. For instance, users may store their data in the cloud without the need to know where it keeps the data or how it accesses the data.

A. Advantages of cloud computing when it comes to scalability, are:

- Inexpensive testing - Testing can be done against a test cluster without risking the performance or integrity of the production system. You can also test the upper limits of the ideal cluster's performance by using "robot users" in the cloud to generate load.
- Reduced risk - Bring up a test instance of the cluster to prove a new code base, and roll out a new version one cluster at a time. Fall back to an older version if the new version doesn't work, without disrupting current users.
- Ability to segment the customer base – Use clusters to separate customers with varying demands, such as a large customer who wants a private instance of the application, or one who requires extensive customizations.
- Auto-scaling based on application load – With the ready availability of resources, applications can be built to recognize when they are reaching the limits of their current configuration and automatically bring up new resources.

V. CHARACTERISTICS

There are several key characteristics of cloud computing [3]:

- The customer doesn't have to know (and buy) the full capacity they might need at a peak time. Cloud computing makes it possible to scale the resources available to the application.
- Customers pay only for what they use. They don't have to buy servers or capacity for their maximum needs. Often, this is a cost savings.
- The cloud will automatically (or, in some services, with semi-manual operations) allocate and de-allocate CPU, storage, and network bandwidth on demand. When there are few users on a site, the cloud uses very little capacity to run the site, and vice versa.
- Because the data centers that run the services are huge, and share resources among a large group of users, the infrastructure costs are lower (electricity, buildings, and so on). Thus, the costs that are passed on to the customer are smaller.

Why would anyone want to rely on another computer system to run programs and store data? Here are just a few reasons:
Clients would be able to access their applications and data from anywhere at any time. They could access the cloud computing system using any computer linked to the Internet. Data wouldn't be confined to a hard drive on one user's computer or even a corporation's internal network.

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

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

Servers and digital storage devices take up space. Some companies rent physical space to store servers and databases because they don't have it available on site. Cloud computing gives these companies the option of storing data on someone else's hardware, removing the need for physical space on the front end.

Corporations might save money on IT support. Streamlined hardware would, in theory, have fewer problems than a network of heterogeneous machines and operating systems.

If the cloud computing system's back end is a grid computing system, then the client could take advantage of the entire network's processing power. Often, scientists and researchers work with calculations so complex that it would take years for individual computers to complete them. On a grid computing system, the client could send the calculation to the cloud for processing. The cloud system would tap into the processing power of all available computers on the back end, significantly speeding up the calculation.

A. *Cloud computing distinguish itself from other computing paradigms, like grid computing, global computing, Internet computing in the following aspects:*

User interfaces
cloud interface do not force users to change their working habits, e.g., developing language, compiler, OS, and so on.
-Cloud client who is required to be installed locally is lightweight e.g., Nimbus Cloud Kit client size is around 15MB.
-Cloud interfaces are location independent and can be accessed by some well established interfaces like Web Services and Internet Browser.

On-demand service provision
-Computing clouds provide resources and services for user on-demand. User can customize required computing environments later on ,for e.g., software installation, network configuration, as user normally own "root privilege".

QoS guaranteed offer
-The cloud computing environments provided by computing clouds can guarantee QoS for users e.g., hardware performance like CPU bandwidth and memory size.

Autonomous system
-The computing cloud is an autonomous system and managed transparently to users. Hardware, software and data inside clouds can be automatically reconfigured.

VI. PRIVACY ISSUES IN CLOUD SERVICES

There is a key challenge for software engineers to design cloud services in such a way as to decrease privacy risk. As with security, it is necessary to design in privacy from the outset, and not just bolt on privacy mechanisms at a later stage.[4] There is an increasing awareness for the need for design for privacy from both companies and governmental organizations Furthermore, there are opportunities for the provision of a new range of 'privacy services' that offer a cloud computing infrastructure with assurances as to the degree of privacy offered, and related opportunities for new accountability-related services to provide certification and audit for these assurances (analogous, for example, to privacy seal provision for web services and mechanisms for privacy assurance on the service provider side).

A. What is privacy?
Privacy is a fundamental human right, enshrined in the United Nations Universal Declaration of Human Rights and the European Convention on Human Rights [5]. There are various forms of privacy, including 'the right to be left alone' and 'control of information about ourselves' A taxonomy of privacy has been produced that focuses on the harms that arise from privacy violations , and this can provide a helpful basis on which to develop a risk/benefit analysis.

B. What types of information need to be protected?
Privacy sensitive information includes the following:

- *Personally identifiable information* (PII): any information that could be used to identify or locate an individual (e.g. name, address) or information that can be correlated with other information to identify an individual (e.g. credit card number, postal code, Internet Protocol (IP) address).
- *Sensitive information*: information on religion or race, health, sexual orientation, union membership or other information that is considered private. Such information requires additional safeguards. Other information that may be considered sensitive includes personal financial information and job performance information. Information considered to be sensitive PII, e.g. biometric information or collections of surveillance camera images in public places.
- *Usage data*: Usage data collected from computer devices such as printers; behavioral information such as viewing habits for digital content, users' recently visited websites or product usage history.
- *Unique device identities*: Other types of information that might be uniquely traceable to a user device, e.g. IP addresses, Radio Frequency Identity (RFID) tags, unique hardware identities.

C. The main privacy risks are:

- *for the cloud service user:* being forced or persuaded to be tracked or give personal information against their will, or in a way in which they feel uncomfortable.
- *for the organization using the cloud service:* non compliance to enterprise policies and legislation, loss of reputation and credibility.
- *for implementers of cloud platforms:* exposure of sensitive information stored on the platforms (potentially for

fraudulent purposes), legal liability, loss of reputation and credibility, lack of user trust and take-up

- *for providers of applications on top of cloud platforms:* legal non compliance, loss of reputation, 'function creep' using the personal information stored on the cloud, i.e. it might later be used for purposes other than the original cloud service intention
- *for the data subject:* exposure of personal information

D. Several security issues that one should discuss with cloud-computing vendors are[6]:

- Privileged user access—Who has specialized access to data and about the hiring and management of such administrators?
- Regulatory compliance—Is the vendor willing to undergo external audits and/or security certifications?
- Data location—Does the provider allow for any control over the location of data?
- Data segregation—Is encryption available at all stages, and were these encryption schemes designed and tested by experienced professionals?
- Recovery—What happens to data in the case of a disaster, and does the vendor offer complete restoration, and, if so, how long does that process take?
- Investigative Support—Does the vendor have the ability to investigate any inappropriate or illegal activity? Long-term viability—What happens to data if the company goes out of business, and is data returned and in what format?
- Data availability—Can the vendor move your data onto a different environment should the existing environment become compromised or unavailable?

E. There are several points to consider before signing up for cloud base service [7, 8]:

- *Vendor lock-in*
  Make sure there is an easy way to get your data out of the service. If you're using an infrastructure service, backing up the files and data should be relatively easy. If you're using a Web application, be sure to have a plan for taking your data with you in case you need to switch to another vendor. You don't always need to move all of your data to the new application if you have a way to somehow view the data. For example, you don't have to move all of the old time tracking application's data to the new one if you have viewable access to it.

- *Reliability*
  If something goes wrong with the service provider, such as servers going down, the customer can do nothing. For situations like this, it's better to choose a service provider who offers mirrored sites. Sometimes, though, even this is not enough; even the big vendors can have problems.

- *Data security*
  This is not always a risk. The security procedures and expertise of the vendor might be a lot better than those of a small start-up. The issue to consider is: who gets to see your data, and what are the vendor's policies for this. For example, if your data is sensitive for competitors to see, check your vendor's policies.

- *Going out of business*
  Investigate what would happen to your data, or to the application, if your vendor is forced to shut down. This negative aspect might be something that is seldom mentioned in marketing materials. If exporting your data is easy, then the possible shut down of the vendor shouldn't be that dangerous. You would still face the problem of finding a suitable new application (or vendor) for your business needs, though.

VII. CONCLUSION

Cloud computing represents an exciting opportunity to bring on demand applications to customers, but it is very important to take security into account when designing cloud services. As companies make plans to install applications in private or public cloud environments, new security challenges need to be addressed. Optimal cloud security practices should include encryption of sensitive or important data used by cloud-based virtual machines; centralized key management that allows the user (and not the cloud provider) to control cloud data; and ensuring that cloud data is accessible according to established enterprise policies.

ACKNOWLEDGMENT

I am thankful to my PhD thesis supervisor, Professor Tapas Kumar for guiding me in preparing this paper.

REFERENCES

[1] Kevin et.al. "Security Issues for Cloud Computing", International Journal of Information Security and Privacy, 4(2), pp.39-51, 2010.
[2] R. L Grossman, "The Case for Cloud Computing", IT Professional, vol. 11(2), pp. 23-27, 2009.
[3] Ronald L. Krutz, Russell Dean Vines "Cloud Security A Comprehensive Guide to Secure Cloud Computing", Wiley Publishing, Inc., 2010
[4] "Cloud Computing Security: Making Virtual Machines Cloud Ready", Available online at: http://www.techrepublic.com/whitepapers/cloud-computing-security-making-virtual-machines-cloudready/1728295
[5] Vouk, M. A., "Cloud Computing – Issues, Research and Implementations", In Proceedings of the 30th

International Conference on Information Technology Interfaces (ITI'08), pp. 31-40, Cavtat, Croatia, June 2008.

[6] L. Wei, H. Zhu, Z. Cao, X. Dong, W. Jia, Y. Chen, A.V. Vasilakos, "Security and privacy for storage and computation in cloud computing", Inform. Sci. pp. 371–386, 2014.

[7] Naresh Vurukonda, B.Thirumala Rao, "A Study on Data Storage Security Issues in Cloud Computing", 2nd International Conference on Intelligent Computing, Communication & Convergence ,2016

[8] "Addressing Data Security Challenges in the Cloud", A Trend Micro White Paper , July 2010.