International Journal of Advanced Technology in Engineering and Science
Vol. No.3, Issue 11, November 2015
www.ijates.com
ijates
ISSN 2348 - 7550

SECURITY CONCERNS OF DATA IN CLOUD ENVIRONMENTS

Kalpana Parsi

Assistant Professor, Dept of Computer Science, St Francis College for Women, Hyderabad, Telangana (India)

ABSTRACT

As technology advances, so issues comes up. The same is applicable to the new computing paradigm, Cloud Computing. From the era of desktop computing, there is a long journey towards this. As with any paradigm, both pros & cons too exist for Cloud Computing, Security, especially Data Security is the major obstacle in Cloud Computing which has to be resolved for its wider adoption, as data plays a vital role in any organization.

Keywords: Cloud Computing, Cloud Computing issues, Data Life Cycle, Data Security, Security.

I. INTRODUCTİON

Cloud Computing is a generic term that follows pay-as-you-go approach for delivering hosted services over the Internet. The original intention of the proposed cloud computing is to provide users with inexpensive computing and storage services, reducing the cost of computing.

According to NIST (National Institute of Standards and Technology, US), Cloud Computing provides *a convenient, on-demand network access to a shared pool of computing resources* [1].

1.1 Real time Deployment of Cloud Computing

General public have been using Cloud Computing without the awareness of the term in the form of Internet services like *Hotmail* (since 1996), *YouTube* (since 2005), *Facebook* (since 2006) and *Gmail* (since 2007). Hotmail is probably the first Cloud Computing application that allowed the general public to keep their data in the form of text and image files at remote servers, provided and managed by others. In the commercial sector, *Amazon.com* was one of the first vendors to provide storage space, computing resources and business functionality following the Cloud Computing model. In 2006, they launched *Elastic Compute Cloud* (EC2) that allowed companies and individuals to rent computers to run their own enterprise applications and services. *Salesforce.com*, founded in 1999, pioneered the concept of delivering enterprise applications as Cloud-based services to enterprises.

Thus the number of Cloud providers are increasing enormously day by day. To mention a few, the major benefits that Cloud Computing promises are:

a. Easy access to software and hardware resources available in the Cloud.

b. Reduced costs because of reduced upfront corporate investment.

c. No long term contracts with vendors (or suppliers) as services and resources are used on pay-per-usage.

d. Reduced management as major of the responsibilities are handled by Cloud provider.

Cloud computing has many potential advantages when compared with the traditional IT model. But from the consumers perspective, cloud computing security concerns remain a major barrier. As per Cloud Adoption Practices & Priorities Survey Report January 2015, Security of data is a major challenge.

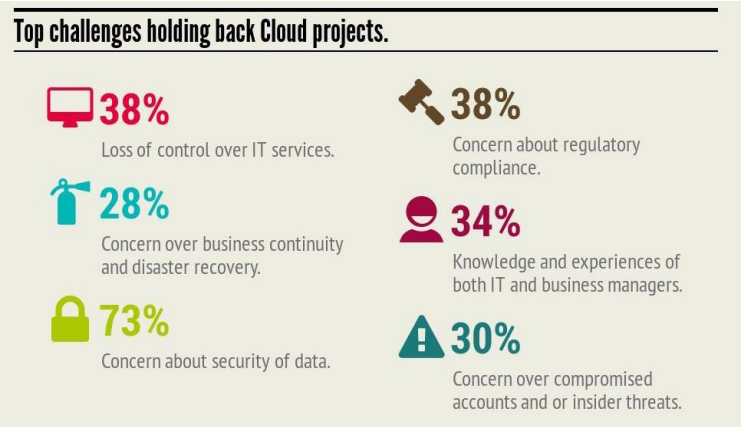


Figure 1. Cloud Adoption Practices & Priorities Survey Report January 2015.

According to a survey carried by Gartner(2009), more than 70% CTOs believed that the primary reason not to use cloud computing services is that there are data security and privacy concerns. In another survey carried by IDCI(2009), 74% IT managers and CIOs believed that the primary challenge that hinders them from using cloud computing services is cloud computing security issues [2]. Although cloud computing service providers promises the security and reliability of their services, actual deployment of cloud computing services is not as safe and reliable as they claim.

1.2 Real Time Incidents Occurred Recently

As per CSA's Cloud Computing vulnerabilities incidents March 2013, Number of incidents reported by cloud service providers is depicted in the following figure:

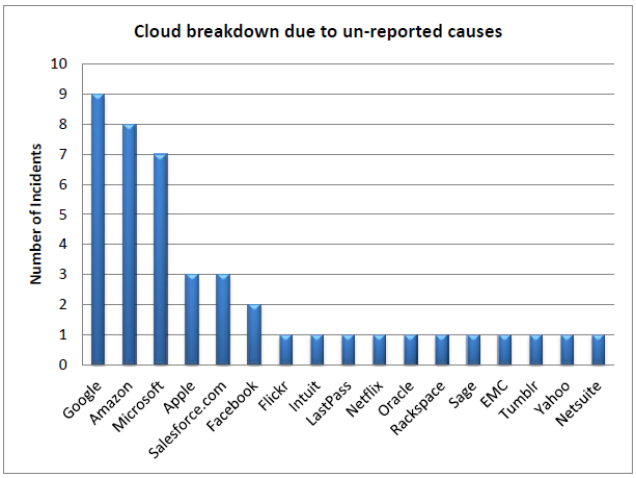


Figure 2. CSA's Cloud Computing vulnerabilities incidents March 2013 report.

Amazon's Simple Storage Service was interrupted twice in February and July 2009. In March 2009, security vulnerabilities in Google Docs even led to serious leakage of user private information. There was a global failure in Google Gmail for about 4 hours. In May 2009, there was serious security vulnerability in VMware virtualization software for Mac version. Amazon's AWS cloud computing platform twice failed in February and July 2008, including the famous micro blogging site Twitter as well as the New York Times causing great trouble among huge number of users [3]. In another scenario, cloud storage vendor LinkUp had been forced to close because there was a huge loss of 45% user data because of the administrators' misuse.

II. INHERENT ISSUES ASSOCIATED WITH THE CLOUD

Cloud Computing promises many benefits, however, there are also numerous issues and challenges for organizations embracing the Cloud technology.

Gartner [4] specifies before making a choice of cloud vendors, users should ask the vendors for seven specific safety issues: Privileged user access, regulatory compliance, data location, data segregation, recovery, investigative support and long-term viability.

In 2009, Forrester Research Inc. [5] evaluated security and privacy practices of some of the leading cloud providers (such as Salesforce.com, Amazon, Google and Microsoft) in three major aspects: Security and privacy, compliance, and legal and contractual issues.

The CSA has identified thirteen domains of concerns on cloud computing security [6].

Zhen [7] lists a number of such challenges including the following:

a. Governance, management and updating of data

b. Management of software services

c. Security of information & data

d. Monitoring of products and processes

e. Reliability and availability of systems and infrastructure

The Expert Group Report [8] mentions a number of issues including:

a. Concerns over security with respect to valuable knowledge, information and data placed on an external service

b. Concerns over availability and business continuity

c. Concerns over data transmission across anticipated broadband speeds.

In addition, another impediment is that corporate or government worries about their data security. It has been recognized by most people that data is the life of a enterprise, to a third party to manage their own lives, regardless of how they promise that there is no problem, many companies and sectors, especially large enterprises and government departments still need to re-measure the cost of savings and necessity. After all, no enterprises can let their core data be handled by others. [9] lists the top ten obstacles in the popularity of cloud computing.

3. DATA SECURITY LIFE CYCLE

The key challenge is data security as discussed. In this, the emphasis is on data security issues in cloud around the data life cycle. Data life cycle refers to the entire process from Data generation to its destruction. The basic life cycle is divided into six phases:

Figure 3. Basic phases of Data Life Cycle.

Although phases in the life cycle are shown as a linear progression, once created data can bounce between phases without any restriction & may not pass through all phases (For example, not all data is eventually destroyed)[10].

3.1 Create: This is the first phase of data life cycle which deals with the generation of new digital content or alteration/updation of existing content.

Data generation is involved in the data ownership. In the traditional IT environment, usually users or organizations own and manage the data. But if data is to be migrated into cloud, it should be considered that how to maintain the data ownership. For personal private information, data owners are entitled to know what personal information being collected & its usage.

3.2 Store: Storing is the act committing the digital data to some sort of storage repository & typically occurs simultaneously with creation.

The data stored in the cloud is similar with the ones stored in other places and needs to consider three aspects of information security: confidentiality, integrity and availability.

The common solution for data confidentiality is data encryption. As the cloud computing environment involves large amounts of data transmission, storage and handling, there also needs to consider processing speed and computational efficiency of encrypting large amounts of data.

Another key problem about data encryption is key management. Who is responsible for key management? Ideally, it's the data owners. But at present, because the users have not enough expertise to manage the keys, they usually entrust the key management to the cloud providers. As the cloud providers need to maintain keys for a large number of users, key management will become more complex and difficult.

In addition to data confidentiality, there also needs to be concerned about data integrity. When the users put several Gigabytes (or more) data into the cloud storage, how to check the integrity of the data?

3.3 Use: Data is viewed, processed or used in some sort of activity.

For the static data using a simple storage service, such as Amazon S3, data encryption is feasible. However, for the static data used by cloud-based applications in PaaS or SaaS model, data encryption in many cases is not feasible. Because data encryption will lead to problems of indexing and query, the static data used by Cloud-based applications is generally not encrypted.

Due to the multi-tenant feature of cloud computing models, the data being processed by cloud- based applications is stored together with the data of other users. Unencrypted data in the process is a serious threat to data security.

Regarding the use of private data, situations are more complicated. The owners of private data need to focus on and ensure whether the use of personal information is consistent with the purposes of information collection and whether personal information is being shared with third parties, for example, cloud service providers.

3.4 Share: Data is exchanged between users, customers and partners.

Within the enterprise boundaries, data transmission usually does not require encryption. For data transmission across enterprise boundaries, both data confidentiality and integrity should be ensured in order to prevent data from being tapped and tampered with by unauthorized users.

Data sharing is expanding the use range of the data. The data owners can authorize the data access to one party, and in turn the party can further share the data to another party without the consent of the data owners. Therefore, during data sharing, especially when shared with a third party, the data owners need to consider whether the third party continues to maintain the original protection measures and usage restrictions.

3.5 Archive: Date leaves active use and enters long-term storage.

Archiving for data focuses on the storage media, whether to provide off-site storage and storage duration. If the data is stored on portable media and if the media is out of control, the data are likely to take the risk of leakage. If the cloud service providers do not provide off-site archiving, the availability of the data will be threatened.

3.6 Destroy: Data is permanently destroyed using physical or digital means.

When the data is no longer required, whether it has been completely destroyed? Due to the physical characteristics of storage medium, the data deleted may still exist and can be restored. This may result in inadvertently disclose of sensitive information.

IV. PROTECTING DATA DURING THE ENTIRETY OF ITS CLOUD LIFE-CYCLE

In the previous section, analysis of security issues in all phases of data life cycle was done. In this, emphasis was on three areas of use that security and privacy professionals need to consider when thinking about protecting data in the cloud: data in transit, data at rest, and data in use [11].

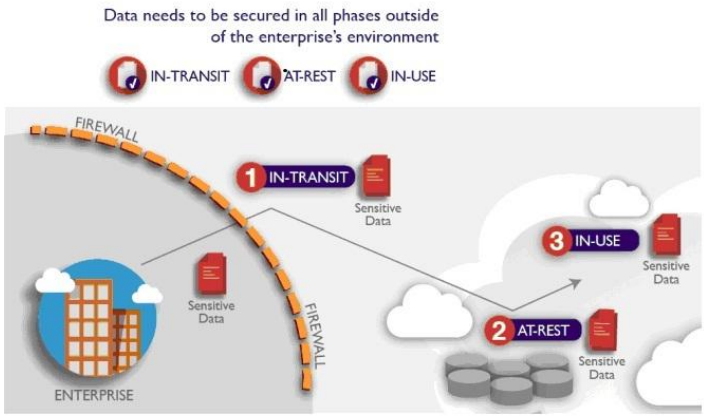


Figure 4. Protecting Data in the Cloud: Data in Transit, Data at Rest, and Data in Use

4.1 Data in transit: The goal of protecting data in motion is to prevent a third party from eavesdropping on a conversation on the wire. Cryptographic protocols, such as Secure Sockets Layer (SSL) or Transport Layer Security (TLS), are typically used for protecting data in motion by establishing an encrypted and authenticated channel, but the data inside the channel is typically transferred in an unencrypted state (e.g. the "payload" is in the clear).

4.2 Data at rest: Data at rest is essentially the data that is stored persistently in some form, as a file, in a database, etc. The goal of protecting data at rest is to prevent a third party from reading the data, should they gain access to the data in its persistent form. Database vendors provide a variety of tools to provide encryption. One key concern regarding the encryption of data at rest in a cloud environment is who owns the keys, and where do the keys reside? The benefits of data at rest protection are somewhat weakened if the data, and the key used to encrypt the data, are both stored in a less trusted security zone. In response, cloud service providers (CSPs) are innovating in this space and are developing techniques whereby the enterprise, not the cloud service provider, owns the keys.

4.3 Data in use: It is effectively, the data that has been loaded into a process and is in the memory of the program that is running. In general, this data is in the clear while being processed and is typically not protected by techniques such as the in-cloud based encryption provided by a cloud service provider. Keeping the data in use in a clear and readable form is required, by design, since the data needs to be in the clear to perform value-added functions on the data (e.g. creating reports, searching on fields, sorting lists, performing calculations, etc.). There is, however, a growing concern since new attack vectors are emerging that specifically target data in use. The recent "Heartbleed" exploit is a good example of a data in use attack. The Heartbleed attack exploited a vulnerability in OpenSSL, which allowed attackers to directly access the memory space of the affected process, leaking sensitive data in use such as usernames and passwords.

V. OTHER RELEVANT SECURITY ISSUES

Security and related issues are already big concerns. Apart from above discussed data security issues, the following are the other main issues which are acting as barrier for full deployment of cloud environments.

5.1 Data Location and Relocation

Cloud Computing offers a high degree of data mobility. Consumers do not always know the location of their data. In most cases, this does not matter. For example, emails and photographs uploaded to Facebook can reside anywhere in the world and Facebook members are generally not concerned. However, when an enterprise has some sensitive data that is kept on a storage device in the Cloud, they may want to know the location of it. They may also wish to specify a preferred location (e.g. data to be kept in the UK). This in turn requires a contractual agreement, between the Cloud provider and the consumer that data should stay in a particular location or reside on a given known server. The issue is that consumers are, sometimes, not aware of the implication of this and thus no such contract is agreed beforehand. An important factor influencing the choice of location for data centers is the cost of running a centre (by reducing the electricity bills, for example) [12].

Another issue is the movement of data. In Public Clouds, data is often regularly routed to other locations at certain times of the day or year, or when there is a huge climatic temperature fluctuation [13]. The main factor is the cost of provision.

Qureshi's [14] proposes a method of dynamically routing data that may become an attractive solution for Cloud storage providers. Dynamic routing of data is also considered to improve resource optimization. This, in turn, may also help to reduce costs by 40%.

Cloud providers have contracts with each other and they use each others' resources. Reasons for using other providers' resources are usually the following:

a. Lack of own resources due to high demand from consumers

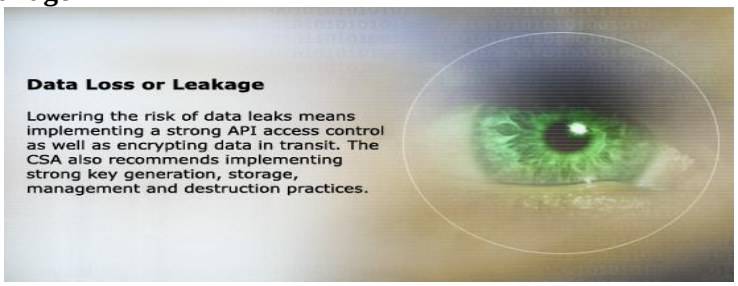
b. There is an opportunity of cost saving or better pricing policy

c. Efficiency in retrieving data

d. Resource Optimization

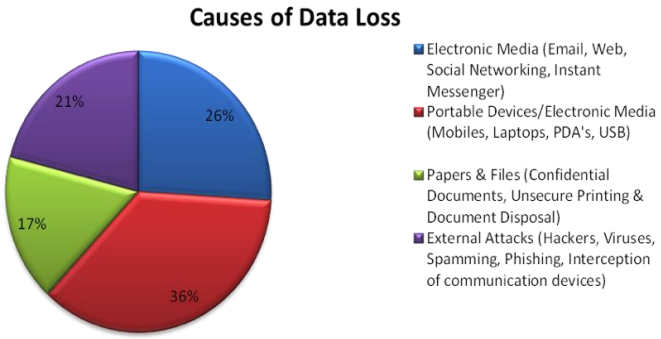
Cross border data transition (from one country to another) also happens for the same reasons. This can lead to additional potential legal risks due to different countries having varying policies, regulations and legislation. This has implications in that data protected by legislation in one country may not have the same, or even similar, protection mechanism in another country [15].

5.2 Data Loss/Leakage



This threat mainly occurs due to vulnerabilities related to data, Vulnerabilities in Virtual Machines, Virtual Machine Images & Vulnerabilities in Virtual Networks.

Data is always in danger of being lost or stolen, may be deletion without a backup, by loss of encoding key or by unauthorized access or to gain confidential information from other VMs co-located in the same server as the attacker.



5.3 Data Availability

In Cloud Computing Data availability becomes a major legitimate issue as the availability of uninterruptible and seamless provision becomes relatively difficult because customer data is normally stored in chunk on different servers often residing in different locations or in different Clouds Furthermore, data availability is such a crucial issue that it is common for Cloud providers to credit customer accounts if the system downtime duration drops below that specified in the SLA (service level agreement). The related issue is that, often, such measures are not specified in the SLAs.

The issue of data availability is exemplified by the outages suffered by Google's *Gmail* service in February 2009 which resulted in embarrassing headlines for the company [16].

5.4 Data Management Security

Data management security means that storage service providers will collect user information as little as possible and should ensure that the data will not be disclosed to any third party without the user's consent.

For most businesses, data security and data protection is the biggest obstacle for implementation of cloud computing technology on applications that contain sensitive or confidential data. As we all know, data encryption is a basic measure to ensure data security. But many applications require data processing in the cloud, and then they need decrypt the data first.

Main issue with respect to data management is dealing with replacement or deactivation of service provider. Once the agreement between the service provider & user expires, the user may disable a service, then what way will the data stored in the cloud return back to its owner?

In general, cloud computing provides huge storage capacity and special format, if copy back to the user exactly, the user is almost impossible to restore itself, even has nowhere to store. A direct solution is to use products of another service provider, and then they will deal with data seamless panning. Currently, major cloud computing service providers, such as Google, Microsoft, Amazon, all have respective form of data storage and system, how to ensure the seamless panning of data between these service providers also need to be considered.

Even though, the data in multiple cloud computing platforms can be transferred smoothly, effective means are needed to ensure the effective destruction of data in the original service provider after the expiration of the contract.

Another issue is third-party backup among service providers is necessary, that is, the data should be backed up to a provider who is not in the same network, same area or even same country, to obtain stable services that is separate from equipment failures, regional policies and even natural disaster.

5.5 Data Control by Third Party

The legal implications of data being held by a third party are complex and not well understood. There is also a potential lack of control and transparency when a third party holds the data.

5.6 Data combination and commingling

Cloud client needs to ensure whether its private data is stored separately from others or not. If they combined or commingled with those of others, then it is much more vulnerable. Ex: Viruses may be transmitted from one client to others easily.

VI. CURRENT SECURITY SOLUTIONS

With time, Cloud providers are becoming more understanding as Cloud consumers are becoming more knowledgeable. As a result, data mobility, location and relocation of data concern are being partially addressed by Cloud providers.

Cloud Security Alliance (CSA) [17] is gathering solution providers, non-profits and individuals to enter into discussion about the current and future best practices for information assurance in the cloud.

Two of the largest vendors have started offering solutions to customers, as mentioned below:

a. Amazon's *AWS* (Amazon Web Services) provides an option within its *S3* (Simple Storage Service) package to allow customers to specify the geographic regions for the storage and location of data. It also provides assurance that data will not leave the customer selected regions [18].

b. In 2009, Microsoft announced that its Windows *Azure* system would provide users with an option to specify geographic regions where the customer data is to be stored.

IBM has developed a fully homomorphic encryption scheme in June 2009. This scheme allows data to be processed without being decrypted [19].

A key problem for data encryption solutions is key management. On one hand, the users have not enough expertise to manage their keys & on the other hand, the cloud service providers need to maintain a large number of user keys. The Organization for the Advancement of Structured Information Standards (OASIS) Key Management Interoperability Protocol (KMIP) is trying to solve such issues [20].

VII. CONCLUSION

Even though Cloud Computing is becoming a hugely attractive paradigm, the issue of data security is one of the most important & primary problem to be solved. Also as mentioned, Data security issues exist in all stages of data life cycle. In the current scenario Cloud computing cannot completely replace traditional computing. It is still not being fully accepted to manage personal data by a third party, especially for large enterprises and government departments.

It is foreseeable that in the near future, the average user will shift entirely to the cloud computing model if the above mentioned security issues are resolved appropriately.

REFERENCES

[1] Peter Mell and Tim Grance, The NIST Definition of Cloud Computing, version 15, National Institute of Standards and Technology (NIST), Information Technology Laboratory, www.csrc.nist.gov, 7 Oct 2009.

[2] Sun Cloud Architecture Introduction White Paper (in Chinese).
http://developers.sun.com.cn/blog/functionalca/resource/sun_353cloudcomputing_chinese.pdf.

[3] Lu Changhai, "Cloud computing-- A beautiful cloud on Internet", Science lllustrated, 2010.6.

[4] Gartner: Seven cloud-computing security risks. InfoWorld. 2008-07-02. http://www.infoworld.com/d/security-central/gartner-seven-cloudcomputing-security-risks-853.

[5] Cloud Security Front and Center. Forrester Research. 2009-11-18. http://blogs.forrester.com/srm/2009/11/cloud-security-front-andcenter.html

[6] Cloud Security Alliance, Security Guidance for Critical Areas of Focus in Cloud Computing, V2.1, http://www.cloudsecurityalliance.org/guidance/csaguide.v2.1.pdf.

[7] Jian Zhen, Five Key Challenges of Enterprise Cloud Computing, Cloud computing journal, 16 Nov 2008.

[8] Lutz Schubert, The Future of Cloud Computing, Expert Group Report, [Online] Available at: http://cordis.europa.eu/fp7/ict/ssai/docs/executivesummaryforweb_en.pdf

[9] Michael Annbrust etc.,Above the Clouds: A Berkeley View of Cloud Computing,http://www.eecs.berkeley.edu/Pubs/TechRpts/2009/EECS-2009 28.pdf:2009.2 .

[10] Data Security Lifecycle - Securosis | https://securosis.com/tag/data+security+lifecycle

[11] Protect Data During Entire Cloud Life-Cycle | Perspecsys
perspecsys.com/.../protect-data-during-the-entirety-of-its-cloud-life-cycle

[12] Jaeger, P. T., Grimes, J. M., Lin, J. & Simmons, S, 'Cloud Computing and Information Policy: Computing in a Policy Cloud?' Journal of Information Technology & Politics, 5(3), 2008.

[13] Knight, W, Energy-Aware Internet Routing, 2009. [Online].
Available at: www.technologyreview.com/business/23248/page2/(Access: March 2011)

[14] Qureshi, A, Plugging Into Energy, 7th ACM Workshop on Hot Topics in Networks (HotNets). Calgary, Canada, October 2008.

[15] European Network and Information Security Agency, (2009) Cloud Computing, Benefits Risks and Recommendations for Information Security, [Online] Available at: http://enisa.europa.eu/

[16] BBC, Google users hit by mail blackout, BBC News, 24 February 2009. [Online]. Available at http://news.bbc.co.uk/1/hi/technology/7907583.stm (Accessed: March 2011).

Cloud Security Alliance. http://www.cloudsecurityalliance.org. Amazon Web Services, Amazon Simple Storage Service FAQs, 2009.[Online] Available at: http://aws.amazon.com/s3/faqs/#Where_is_my_data_stored (Accessed: March 2011)

[17] Dropbox.http://www.dropbox.com.2011.1.

[18] "Amazon Simple Storage Service," http://aws.amazon.com/s3 2011.1.