Cloud Computing and Data Protection Challenges

Adesh Kumar
Mewar University, Chittorgarh, Rajasthan-312901, India

Dr.Vikas Kumar
Asia Pacific Institute of Management, New Delhi

Abstract- Cloud computing is basically a new technique for outsourcing computer services. Instead of storing your data or your software on your network, you pay a company to store them for you. Instead of maintaining a technical infrastructure in order to manipulate, calculate or whatever else you do with your information, you log in through the Internet and do it on the cloud provider's systems instead. Using the cloud raises legal and other risks. In this paper we try to find out the risk and challenges to protect data in public cloud.

Keywords: Cloud Computing, Data Protection, Security

I. INTRODUCTION

A cloud computing is a computing model that makes IT resources such as servers, middleware, and applications available over the Internet as services to business organizations in a self-service manner[1]. In Cloud Computing, data owners share their outsourced data with a large number of cloud users. Each user might be interested in retrieving only a specific data file in a given session. Also it must be guaranteed that only authorized users must have the permission to view the data file. User authentication can be performed by using many scientific ways. Alphanumeric passwords and graphical passwords are both guaranteed service [2]. Cloud Computing has evolved over the past from utility computing, autonomic computing and grid computing through the sharing of resources, computation and storage capabilities. Although cloud computing promises lower costs, rapid scaling, easier maintenance, and service availability anywhere, anytime, a key challenge is how to ensure and build confidence that the cloud can handle user data securely. A recent Microsoft survey found that "58 percent of the public and 86 percent of business leaders are excited about the possibilities of cloud computing. But more than 90 percent of them are worried about security, availability, and privacy of their data as it rests in the cloud."1 This tension makes sense: users want to maintain control of their data, but they also want to benefit from the rich services that application developers can provide using that data. So far, the cloud offers little platform-level support or standardization for user data protection beyond data encryption at rest, most likely because doing so is nontrivial. Protecting user data while enabling rich computation requires both specialized expertise and resources that might not be readily available to most application developers. Building in data-protection solutions at the platform layer is an attractive option: the platform can achieve economies of scale by amortizing expertise costs and distributing sophisticated security solutions across different applications and their developers [3].

1.1 Cloud Service Models

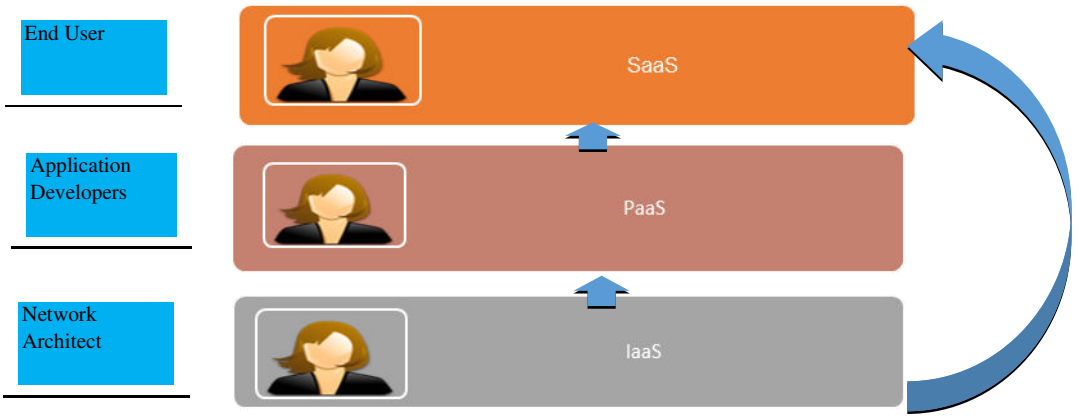

Fig 1: A typical Cloud paradigm

Cloud computing can be considered as a new computing technique that can provide services on demand at a minimal cost. The three well-known and commonly used service models in the cloud paradigm are software as a service (SaaS), platform as a service (PaaS), and infrastructure as a service (IaaS). In SaaS, software with the related data is deployed by a cloud service provider, and users can use it through the web browsers. In PaaS, a service provider facilitates services to the users with a set of software programs that can solve the specific tasks. In IaaS, the cloud service provider facilitates services to the users with virtual machines and storage to improve their business capabilities Fig 1.

1.2 Cloud Deployment Models

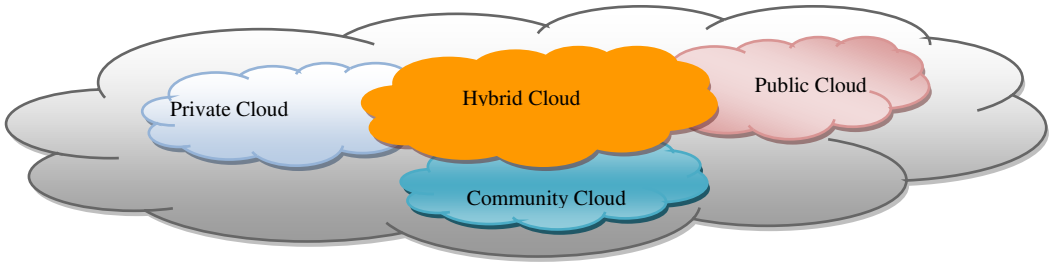

Fig.2 Cloud Deployment Models

The Cloud deployment models, which can be either internally or externally implemented, can be classified as private cloud, public cloud, community cloud and hybrid cloud as shown in fig.2 and can be described as:

1.2.1 Private Cloud

Private cloud is a new term that some vendors have recently used to describe offerings that emulate cloud computing on private networks. It is set up within an organization's internal enterprise datacenter. In the private cloud, scalable resources and virtual applications provided by the cloud vendor are pooled together and available for cloud users to share and use. It differs from the public cloud in that all the cloud resources and applications are managed by the organization itself, similar to Intranet functionality. Utilization on the private cloud can be much more secure than that of the public cloud because of its specified internal exposure. Only the organization and designated stakeholders may have access to operate on a specific Private cloud.

1.2.2 Public Cloud

A public cloud is a cloud computing deployment scheme that is generally open for use by the general public [8]. The general public is defined in this case as either individual users or corporations. The public cloud infrastructure used is owned by a cloud services vendor organization; examples of public cloud deployment vendor offerings include Amazon Web Services, Google App Engine, Salesforce.com, and Microsoft Windows Azure.

1.2.3 Community Cloud

A cloud deployment model that is being rapidly implemented is called a community cloud. Conceptually residing somewhere between a private cloud and a public cloud, community cloud describes a shared

infrastructure that is employed by and supported by multiple companies [8]. This shared cloud resource may be utilized by groups that have overlapping considerations, such as joint compliance requirements, non-competitive business goals, or a need to pool high-level security resources.

Although the physical existence of the shared cloud may reside on any member's premises, or even on a third-party site, managing the community cloud may become complicated, due to unspecified or shifting ownership and responsibility, making it somewhat technically challenging to deal with concerns over resource management, privacy, resilience, latency, and security requirements.

1.2.4 Hybrid Cloud

A hybrid cloud is any combination of the private, public and community cloud deployment models [8]. It is defined by NIST as "a composition of two or more clouds (private, community, or public) that remain unique entities but are bound together by standardized or proprietary technology that enables data and application portability An example of hybrid cloud deployment may consist of an organization deploying noncritical software applications in the public cloud, while keeping critical or sensitive apps in a private cloud, on the premises. Hybrid clouds combine both public and private cloud models, and they can be particularly effective when both types of cloud are located in the same facility.

II. THE NEED FOR DATA PROTECTION

Moving data into the cloud offers great convenience to users since they don't have to care about the complexities of direct hardware management. The pioneer of Cloud Computing vendors, Amazon Simple Storage Service (S3) and Amazon Elastic Compute Cloud (EC2) are both well-known examples. While these internet-based online services do provide huge amounts of storage space and customizable computing resources, this computing platform shift, however, is eliminating the responsibility of local machines for data maintenance at the same time. As a result, users are at the mercy of their cloud service providers for the availability and integrity of their data from the perspective of data security, which has always been an important aspect of quality of service, Cloud Computing inevitably poses new challenging security threats for number of reasons. Firstly, traditional cryptographic primitives for the purpose of data security protection can not be directly adopted due to the users' loss control of data under Cloud Computing. Therefore, verification of correct data storage in the cloud must be conducted without explicit knowledge of the whole data. Considering various kinds of data for each user stored in the cloud and the demand of long term continuous assurance of their data safety, the problem of verifying correctness of data storage in the cloud becomes even more challenging. Secondly, Cloud Computing is not just a third party data warehouse. The data stored in the cloud may be frequently updated by the users, including insertion, deletion, modification, appending, reordering, etc. To ensure storage correctness under dynamic data update is hence of paramount importance. However, this dynamic feature also makes traditional integrity insurance techniques futile and entails new solutions. Last but not the least, the deployment of Cloud Computing is powered by data centers running in a simultaneous, cooperated and distributed manner. Individual user's data is redundantly stored in multiple physical locations to further reduce the data integrity threats. Therefore, distributed protocols for storage correctness assurance will be of most importance in achieving a robust and secure cloud data storage system in the real world.

Because private data moves online from one location to another, the need to secure it properly becomes increasingly urgent. The good news is that the same forces concentrating data in enormous datacentres will also aid in using collective security expertise more effectively. Adding protections to a single cloud platform can immediately benefit hundreds of thousands of applications and, by extension, hundreds of millions of users. While we have focused here on a particular, albeit popular and privacy-sensitive, class of applications, many other applications also need solutions, and many practical questions still remain open:

- Can we standardize technology across platforms to facilitate switching among providers?
- How can we make migration to the DPaaS (Data Protection as a service) cloud as easy as possible for existing applications?
- How can we minimize the cost of application audits?
- What kinds of audits are most important for building user confidence?
- Can technologies such as TC and code attestation be made scalable in the presence of constantly evolving software?
- How can we generalize the ideas presented here to other classes of applications?

In posing these questions, we hope to provoke thought and inspire future research and development in this important direction

III. DATA PROTECTION PERSPECTIVES

It's impossible to develop a single data-protection solution for the cloud because the term means too many different things. Cloud computing can be categorised in various domains. In this paper if we consider an important class of widely used applications that includes e-mail, personal financial management, social networks, and business tools such as word processors and spreadsheets. The following criteria define this class of applications:

- Provide services to a large number of distinct end users, as opposed to bulk data processing or workflow management for a single entity.
- Use a data model consisting mostly of sharable units, where all data objects have access control lists (ACLs) with one or more users.
- Developers could run the applications on a separate computing platform that encompasses the physical infrastructure, job scheduling, user authentication, and the base software environment, rather than implementing the platform themselves.

Excessively rigid security is as detrimental to cloud service value as inadequate security. A primary challenge in designing a platform-layer solution useful to many applications is ensuring that it enables rapid development and maintenance. To ensure a practical solution, we considered the following goals relating to data protection as well as ease of development and maintenance:

- Integrity- Data integrity in the cloud system means preserving information integrity. The data should not be lost or modified by unauthorized users. Data integrity is the basis to provide cloud computing service such as SaaS, PaaS, and IaaS. Besides data storage of large-scaled data, cloud computing environment usually provides data processing service. Data integrity can be obtained by techniques such as RAID-like strategies and digital signature. Owing to the large quantity of entities and access points in a cloud environment, authorization is crucial in assuring that only authorized entities can interact with data. By avoiding the unauthorized access, organizations can achieve greater confidence in data integrity. The monitoring mechanisms offer the greater visibility into determining who or what may have altered data or system information, potentially affecting their integrity. Cloud computing providers are trusted to maintain data integrity and accuracy. However, it is necessary to build the third party supervision mechanism besides users and cloud service providers [7].
- Privacy- In the cloud, the privacy means when users visit the sensitive data, the cloud services can prevent potential adversary from inferring the user's behaviour by the user's visit model (not direct data leakage). Researchers have focused on Oblivious RAM (ORAM) technology. ORAM technology visits several copies of data to hide the real visiting aims of users. ORAM has been widely used in software protection and has been used in protecting the privacy in the cloud as a promising technology [7]. The privacy issues differ according to different cloud scenarios and can be divided into four subcategories as follows:
    - how to enable users to have control over their data when the data are stored and processed in cloud and avoid theft, nefarious use, and unauthorized resale,
    - how to guarantee data replications in a jurisdiction and consistent state, where replicating user data to multiple suitable locations is an usual choice, and avoid data loss, leakage, and unauthorized modification or fabrication,
    - which party is responsible for ensuring legal requirements for personal information,
    - to what extent cloud subcontractors are involved in processing which can be properly identified, checked, and ascertained
- Data confidentiality- Data confidentiality is important for users to store their private or confidential data in the cloud. Authentication and access control strategies are used to ensure data confidentiality. The data confidentiality, authentication, and access control issues in cloud computing could be addressed by increasing the cloud reliability and trustworthiness. The users hesitate to trust the cloud providers and cloud storage services because it is virtually impossible to eliminate potential insider threat, it is very dangerous for users to store their sensitive data in cloud storage directly. Simple encryption is faced with the key management problem and cannot support complex requirements such as query, parallel modification, and fine-grained authorization. Encryption is usually used to ensure the confidentiality of data.
- Data availability- Data availability means when accidents such as hard disk damage, IDC fire, and network failures occur, the extent that user's data can be used or recovered and how the users verify their data by techniques rather than depending on the credit guarantee by the cloud service provider alone. The issue of storing data over the trans boarder servers is a serious concern of clients because the cloud vendors are governed by the local laws and, therefore, the cloud clients should be aware of those laws. Moreover, the cloud service provider should ensure the data security, particularly data confidentiality and integrity. The cloud provider should share all such concerns with the client and

build trust relationship in this connection. The cloud vendor should provide guarantees of data safety and explain jurisdiction of local laws to the clients. The main focus of the paper is on those data issues and challenges which are associated with data storage location and its relocation, cost, availability, and security.

- Access transparency- Logs should clearly indicate who or what accessed any data.
- Ease of verification- Users should be able to easily verify what platform or application code is running, as well as whether the cloud has strictly enforced their data's privacy policies.
- Rich computation- The platform should allow efficient, rich computations on sensitive user data.
- Development and maintenance support- Because developers face a long list of challenges as bugs to find and fix, frequent software upgrades, continuous usage pattern changes, and user demand for high performance they should receive both development and maintenance support.

IV. CONCLUSION

Cloud Computing is a new concept that presents a number of benefits for its end users; however, it also raises some security problems which may slow down its use. Understanding what vulnerabilities exist in Cloud Computing will help organizations to make the necessary measures towards the Cloud. Since Cloud Computing leverages many technologies, it also inherits their security issues. Traditional web applications, data hosting, and virtualization have been looked over, but some of the solutions offered are immature or inexistent. We have presented data protection issues for cloud models: IaaS, PaaS, and IaaS, which vary depending on the model. As described in this paper, storage, virtualization, and networks are the biggest security concerns in Cloud Computing. Virtualization which allows multiple users to share a physical server is one of the major concerns for cloud users. Also, another challenge is that there are different types of virtualization technologies, and each type may approach security mechanisms in different ways. Virtual networks are the tempting target for some attacks especially when communicating with remote virtual machines. Some surveys have discussed security issues about clouds without making any difference between vulnerabilities and threats. However, new security techniques are needed as well as redesigned traditional solutions that can work with cloud architectures. Traditional security mechanisms may not work well in cloud environments because it is a complex architecture that is composed of a combination of different technologies.

REFERENCES

[1] K. Bhima and S. Suresh, "Privacy Data Control and Data Protection as a Service in Cloud Computing", International Journal of Advanced Research in Computer and Communication Engineering Vol. 2, Issue 9, pp.1-2, September 2013
[2] C. Sunumol and M. Kavitha, " Providing Data Protection as a Service in Cloud Computing", International Journal of Scientific and Research Publications, Volume 3, Issue 6, pp.1-2,June 2013
[3] S. Dawn, S. Elaine, F. Ian and S. Umesh, "Cloud Data Protection for the Masses", Published by the IEEE Computer Society, 2012
[4] W. Cong, W. Qian and R. Kui, "Ensuring Data Storage Security in Cloud Computing"
[5] Hashizume et al. "An Analysis of Security Issues for Cloud Computing", Journal of Internet Services and Applications, Springer, 2013
[6] S.G. Farhad and B. Meysam, "Evaluation of the Data Security Methods in Cloud Computing Environments", International Journal in Foundations of Computer Science & Technology (IJFCST), Vol. 3, No.2, March 2013
[7] S. Yunchuan, Z. Junsheng, X. Yongping and Z. Guangyu, "Data Security and Privacy in Cloud Computing", Review Article, Hindawi Publishing Corporation International Journal of Distributed Sensor Networks, Volume 2014, Article ID 190903
[8] R. L. Krutz and R. D. Vines, "Cloud Security: A Comprehensive Guide to Secure Cloud Computing", Wiley Publishing, pp.62-85, 2010
[9] Cloud Computing Security. A Trend Micro White Paper, May 2010
[10] G. Lin, D. Fu, J. Zhu, and G. Dasmalchi, "Cloud computing: It as aservice," IT Professional, vol. 11, 2009