International Journal of Emerging Trends & Technology in Computer Science (IJETTCS)
Web Site: www.ijettcs.org Email: editor@ijettcs.org, editorijettcs@gmail.com
Volume 2, Issue 2, March – April 2013 ISSN 2278-6856

Secure Framework for Data Storage from Single to Multi clouds in Cloud Networking

B.Sujana[1], P.Tejaswini[2], G.Srinivasulu[3], Sk.Karimulla[4]

[1,2,3,4] QUBA COLLEGE OF ENGINEERING & TECH, NELLORE

Abstract: ***Cloud computing presents a delivery model for IT services based on internet and provide a scalable service to easily consumed over the internet on an as-needed basis. A major feature is cloud service is that data processed on clouds are often outsourced, leading to number of issues related to accountability including the handling of personally identifiable information. Dealing with "single cloud" providers is predicted to become less popular with customers due to risks of service availability failure and the possibility of malicious insiders in the single cloud. A movement towards "multi-clouds", or in other words, "interclouds" or "cloud-of-clouds" has emerged recently. It is found that the research into the use of multi-cloud providers to maintain security has received less attention from the research community than has the use of single clouds. This work aims to promote the use of multi-clouds due to its ability to reduce security risks*** **that affect the cloud computing user.**

Keywords: Cloud computing, single cloud, multi-clouds, cloud storage, data integrity, data intrusion, service availability.

1. Introduction

The use of cloud computing has increased rapidly in many organizations. Cloud providers should address privacy and security issues as a matter of high and urgent priority. Dealing with "single cloud" providers is becoming less popular with customers due to potential problems such as service availability failure and the possibility that there are malicious insiders in the single cloud. In recent years, there has been a move towards "multi-clouds", "intercloud" or "cloud-of-clouds".

This paper focuses on the issues related to the data security aspect of cloud computing. As data and information will be shared with a third party, cloud computing users want to avoid an untrusted cloud provider. Protecting private and important information, such as credit card details or a patient's medical records from attackers or malicious insiders is of critical importance. In addition, the potential for migration from a single cloud to a multi-cloud environment is examined and research related to security issues in single and multi-clouds in cloud computing are surveyed.

The remainder of this paper is organized as follows. Section II describes the beginning of cloud computing and its components. In addition, it presents examples of cloud providers and the benefits of using their services. Section III discusses security risks in cloud computing. Section IV analyses the new generation of cloud computing, that is, multi -clouds and recent solutions to address the security of cloud computing, as well as examining their limitations. Section 5 presents suggestions for future work. Section 6 will conclude the paper.

2. Background

NIST describes cloud computing as "a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction".

2.1. Components of the Cloud Computing

The cloud computing model consists of five characteristics, three delivery models, and four deployment models. The five key characteristics of cloud computing are: location-independent resource pooling, on-demand self-service, rapid elasticity, broad network access, and measured service. These five characteristics represent the first layer in the cloud environment architecture (see Figure1).

The three key cloud delivery models are infrastructure as a service (IaaS), platform as a service (PaaS), and software as a service (SaaS). In IaaS, the user can benefit from networking infrastructure facilities, data storage and computing services. In other words, it is the delivery of computer infrastructure as a service. An example of IaaS is the Amazon web service. In PaaS, the user runs custom applications using the service provider's resources. It is the delivery of a computing platform and solution as a service. An example of PaaS is GoogleApps. Running software on the provider's infrastructure and providing licensed applications to users to use services is known as SaaS. An example of SaaS is the Salesforce.com CRM application. This model represents the second layer in the cloud environment architecture. Cloud deployment models include public, private, community, and hybrid clouds. A cloud environment that is accessible for multi-tenants and is available to the public is called a public cloud.

Layer	Cloud Computing Components
Five Characteristics	On-demand self-service Broad network access Resource pooling Rapid elasticity Measured Service
Three Delivery models	IaaS PaaS SaaS
Four Deployment models	Public Private Community Hybrid

Figure 1: Cloud Environment Architecture.

A private cloud is available for a particular group, while a community cloud is modified for a specific group of customers. Hybrid cloud infrastructure is a composition of two or more clouds (private, community, or public cloud). This model represents the third layer in the cloud environment architecture. The infrastructure that is owned and managed by users is in the private cloud. Data that is accessed and controlled by trusted users is in a safe and secure private cloud, whereas the infrastructure that is managed and controlled by the cloud service provider is in a public cloud. In particular, this data is out of the user's control, and is managed and shared with unsafe and untrusted servers.

2.2 Cloud Service Providers Examples

In the commercial world, various computing needs are provided as a service. The service providers take care of the customer's needs by, for example, maintaining software or purchasing expensive hardware. For instance, the service EC2, created by Amazon, provides customers with scalable servers. As another example, under the CLUE program, NSF joined with Google and IBM to offer academic institutions access to a large-scale distributed infrastructure.

There are many features of cloud computing. First, cloud storages, such as Amazon S3, Microsoft SkyDrive, or NirvanixCloudNAS, permit consumers to access online data. Second, it provides computation resources for users such as Amazon EC2. Third, Google Apps or versioning repositories for source code are examples of online collaboration tools.

Cloud service providers should ensure the security of their customers' data and should be responsible if any security risk affects their customers' service infrastructure. A cloud provider offers many services that can benefit its customers, such as fast access to their data from any location, scalability, pay-for-use, data storage, data recovery, protection against hackers, on-demand security controls, and use of the network and infrastructure facilities.

Reliability and availability are other benefits of the public cloud, in addition to low cost. However, there are also concerning issues for public cloud computing, most notably, issues surrounding data integrity and data confidentiality. Any customer will be worried about the security of sensitive information such as medical records or financial information.

3. Security Risks in Cloud Computing

Although cloud service providers can offer benefits to users, security risks play a major role in the cloud computing environment. Users of online data sharing or network facilities are aware of the potential loss of privacy. Protecting private and important information such as credit card details or patients' medical records from attackers or malicious insiders is of critical importance . Moving databases to a large data centre involves many security challenges such as virtualization vulnerability, accessibility vulnerability, privacy and control issues related to data accessed from a third party, integrity, confidentiality, and data loss or theft.

In different cloud service models, the security responsibility between users and providers is different. According to Amazon, their EC2 addresses security control in relation to physical, environmental, and virtualization security, whereas, the users remain responsible for addressing security control of the IT system including the operating systems, applications and data.

For instance, any damage which occurs to the security of the physical infrastructure or any failure in relation to the management of the security of the infrastructure will cause many problems. In the cloud environment, the physical infrastructure that is responsible for data processing and data storage can be affected by a security risk.

As the cloud services have been built over the Internet, any issue that is related to internet security will also affect cloud services. Resources in the cloud are accessed through the Internet; consequently even if the cloud provider focuses on security in the cloud infrastructure, the data is still transmitted to the users through networks which may be insecure. As a result, internet security problems will affect the cloud, with greater risks due to valuable resources stored within the cloud and cloud vulnerability. The technology used in the cloud is similar to the technology used in the Internet. Encryption techniques and secure protocols are not sufficient to protect data transmission in the cloud. Data intrusion of the cloud through the Internet by hackers and cybercriminals needs to be addressed and the cloud environment needs to be secure and private for clients.

We will address three security factors that particularly affect single clouds, namely data integrity, data intrusion, and service availability.

3.1 Data Integrity

One of the most important issues related to cloud security risks is data integrity. The data stored in the cloud may suffer from damage during transition operations from or to the cloud storage provider. Examples of the risk of attacks from both inside and outside the cloud provider, such as the recently attacked Red Hat Linux's distribution servers . Another example of breached data occurred in 2009 in Google Docs, which triggered the Electronic Privacy Information Centre for the Federal Trade Commission to open an investigation into Google's Cloud Computing Services . Another example of a risk to data integrity recently occurred in Amazon S3 where users suffered from data corruption. Further examples giving details of attacks can be read in

Although this protocol solves the problem from a cloud storage perspective, arguing that they remain concerned about the users' view, due to the fact that users trust the cloud as a single reliable domain or as a private cloud without being aware of the protection protocols used in the cloud provider's servers. As a solution, using Byzantine fault -tolerant protocols across multiple clouds from different providers is a beneficial solution.

3.2 Data Intrusion

Another security risk that may occur with a cloud provider, such as the Amazon cloud service, is a hacked password or data intrusion. If someone gains access to an Amazon account password, they will be able to access all of the account's instances and resources. Thus the stolen password allows the hacker to erase all the information inside any virtual machine instance for the stolen user account, modify it, or even disable its services. Furthermore, there is a possibility for the user's email(Amazon user name) to be hacked, and since Amazon allows a lost password to be reset by email, the hacker may still be able to log in to the account after receiving the new reset password.

3.3 Service Availability

Another major concern in cloud services is service availability. Amazon mentions in its licensing agreement that it is possible that the service might be unavailable from time to time. The user's web service may terminate for any reason at any time if any user's files break the cloud storage policy. In addition, if any damage occurs to any Amazon web service and the service fails, in this case there will be no charge to the Amazon Company for this failure. Companies seeking to protect services from such failure need measures such as backups or use of multiple providers. Both Google Mail and Hotmail experienced service down-time recently. If a delay affects payments from users for cloud storage, the users may not be able to access their data. Due to a system administrator error, 45% of stored client data was lost in LinkUp (MediaMax) as a cloud storage provider.

4. Multi-Clouds Computing Security

This section will discuss the migration of cloud computing from single to multi-clouds to ensure the security of the user's data.

4.1 Multi-Clouds: Preliminary

The term "multi-clouds" is similar to the terms "interclouds" or "cloud-of-clouds". These terms suggest that cloud computing should not end with a single cloud. Using their illustration, a cloudy sky incorporates different colors and shapes of clouds which leads to different implementations and administrative domains.

Recent research has focused on the multi-cloud environment which control several clouds and avoids dependency on any one individual cloud.

Two layers are identified in the multi-cloud environment: the bottom layer is the inner-cloud, while the second layer is the inter-cloud. In the inter-cloud, the Byzantine fault tolerance finds its place. We will first summarize the previous Byzantine protocols over the last three decades.

4.2 Introduction of Byzantine Protocols

In cloud computing, any faults in software or hardware are known as Byzantine faults that usually relate to inappropriate behavior and intrusion tolerance. In addition, it also includes arbitrary and crash faults . Much research has been dedicated to Byzantine fault tolerance (BFT) since its first introduction , . Although BFT research has received a great deal of attention, it still suffers from the limitations of practical adoption and remains peripheral in distributed systems .

The relationship between BFT and cloud computing has been investigated, and many argue that in the last few years, it has been considered one of the major roles of the distributed system agenda. Furthermore, many describe BFT as being of only "purely academic interest" for a cloud service . This lack of interest in BFT is quite different to the level of interest shown in the mechanisms for tolerating crash faults that are used in large -scale systems. Reasons that reduce the adoption of BFT are, for example, difficulties in design, implementation, or understanding of BFT protocols. As mentioned earlier, BFT protocols are not suitable for single clouds.

4.3 DepSky System: Multi-Clouds Model

This section will explain the recent work that has been done in the area of multi-clouds. We present a virtual storage cloud system called DepSky which consists of a combination of different clouds to build a cloud-of-clouds. The DepSky system addresses the availability and the confidentiality of data in their storage system by using multi-cloud providers, combining Byzantine quorum system protocols, cryptographic secret sharing and erasure codes.

4.3.1 DepSky Architecture

The DepSky architecture consists of four clouds and each cloud uses its own particular interface. The DepSky algorithm exists in the clients' machines as a software library to communicate with each cloud (Figure 2). These four clouds are storage clouds, so there are no codes to be executed. The DepSky library permits reading and writing operations with the storage clouds.

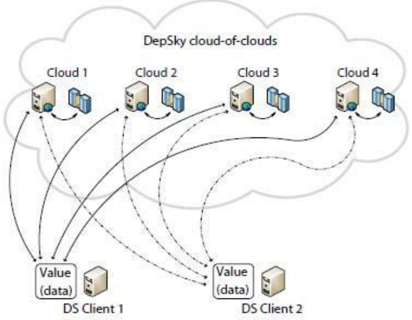

Figure 2: DepSky Architecture [8].

DepSky Data model. As the DepSky system deals with different cloud providers, the DepSky library deals with different cloud interface providers and consequently, the data format is accepted by each cloud. The DepSky data model consists of three abstraction levels: the conceptual data unit, a generic data unit, and the data unit implementation.

DepSKy System model. The DepSky system model contains three parts: readers, writers, and four cloud storage providers, where readers and writers are the client's tasks.It explain the difference between readers and writers for cloud storage. Readers can fail arbitrarily (for example, they can fail by crashing, they can fail from time to time and then display any behavior) whereas, writers only fail by crashing.

Cloud storage providers in the DepSky system model. The Byzantine protocols involve a set of storage clouds (n) where n = 3 f +1, and f is maximum number of clouds which could be faulty. In addition, any subset of (n – f) storage cloud creates byzantine quorum protocols.

4.4 Analysis of Multi-Cloud Research

Moving from single clouds or inner-clouds to multi-clouds is reasonable and important for many reasons. We assumes that the main purpose of moving to interclouds is to improve what was offered in single clouds by distributing reliability, trust, and security among multiple cloud providers. In addition, reliable distributed storage which utilizes a subset of BFT techniques was suggested to be used in multi-clouds. Therefore, the storage load will be spread among several providers as a result of the RACS proxy.

HAIL (High Availability and Integrity Layer) is another example of a protocol that controls multiple clouds. HAIL is a distributed cryptographic system that permits a set of servers to ensure that the client's stored data is retrievable and integral. HAIL provides a software layer to address availability and integrity of the stored data in an intercloud .

We present a design for intercloud storage (ICStore), which is a step closer than RACS and HAIL as a dependable service in multiple clouds. We develop theories and protocols to address the CIRC attributes (confidentiality, integrity, reliability and consistency) of the data stored in clouds.

As mentioned before, presenting a virtual storage cloud system called DepSky consisting of a combination of different clouds to build a cloud-of-clouds. Discussing some limitations of the HAIL protocol and RACS system when compared with DepSky. HAIL does not guarantee data confidentiality, it needs code execution in their servers, and it does not deal with multiple versions of data. None of these limitations are found in DepSky , whereas the RACS system differs from the DepSky system in that it deals with "economic failures" and vendor lock-in and does not address the issue of cloud storage security problems. In addition, it also does not provide any mechanism to ensure data confidentiality or to provide updates of the stored data. Finally, the DepSky system presents an experimental evaluation with several clouds, which is different from other previous work on multi-clouds.

There are a number of studies on gaining constancy from untrusted clouds. For instance, similar to DepSky, Depot improves the flexibility of cloud storage, as believe that cloud storages face many risks . However, Depot provides a solution that is cheaper due to using single clouds, but it does not tolerate losses of data and its service availability depends on cloud availability . Other work which implements services on top of untrusted clouds are studies such as SPORC and Venus. These studies are different from the DepSky system because they consider a single cloud (not a cloud -of-clouds). In addition, they need code execution in their servers. Furthermore, they offer limited support for the unavailability of cloud services in contrast to DepSky.

4.5 Current Solutions of Security Risks

In order to reduce the risk in cloud storage, customers can use cryptographic methods to protect the stored data in the cloud . Using a hash function is a good solution for data integrity by keeping a short hash in local memory. In this way, authentication of the server responses is done by recalculating the hash of the received data which is compared with the local stored data . If the amount of data is large, then a hash tree is the solution . Many storage system prototypes have implemented hash tree functions, such as SiRiUS and TDB . We argue that although the previous methods allow consumers to ensure the integrity of their data which has been returned by servers, they do not guarantee that the server will answer a query without knowing what that query is and whether the data is stored correctly in the server or not. Proofs of Retrievability (PORs) and Proofs of Data Possession (PDP) are protocols introduced to ensure high probability for the retrieval of the user's data. We suggesting using multiple cloud providers to ensure data integrity in cloud storage and running Byzantine-fault -tolerant protocols on them where each cloud maintains a single replica . Computing resources are required in this approach and not only storage in the cloud, such a service provided in Amazon EC2, whereas if only storage service is available. Working with Byzantine Quorum Systems by using Byzantine Disk Paxos and using at least four different clouds in order to ensure users' atomicity operations and to avoid the risk of one cloud failure.

Byzantine fault-tolerant replication to store data on several cloud servers, so if one of the cloud providers is damaged, they are still able to retrieve data correctly. Data encryption is considered the solution to address the problem of the loss of privacy. They argue that to protect the stored data from a malicious insider, users should encrypt data before it is stored in the cloud. As the data will be accessed by distributed applications, the DepSky system stores the cryptographic keys in the cloud by using the secret sharing algorithm to hide the value of the keys from a malicious insider.

In the DepSky system, data is replicated in four commercial storage clouds (Amazon S3,Windows Azure, Nirvanix and Rackspace); it is not relayed on a single cloud, therefore, this avoids the problem of the dominant cloud causing the so-called vendor lock-in issue. In addition, storing half the amount of data in each cloud in the DepSky system is achieved by the use of erasure codes. Consequently, exchanging data between one provider to another will result in a smaller cost. The DepSky system aims to reduce the cost of using four clouds(which is four times the overhead) to twice the cost of using a single cloud, which is a significant advantage .

DepSky uses a set of Byzantine quorum system protocols in order to implement the read and write operations in the system, so it needs only two communication round trips for each operation to deal with several clouds. The use of several clouds needs a variety of locations, administration, design and implementation, which are the requirements of the Byzantine quorum systems protocols. Executing codes in servers is not required in the DepSky system (storage clouds) in contrast to other Byzantine protocols that need some code execution. After using these protocols, the DepSky system aims to deal with data confidentiality by decreasing the stored amount of data in each cloud .

4.6 Limitation of Current Solutions

The problem of the malicious insider in the cloud infrastructure which is the base of cloud computing is considered. IaaS cloud providers provide the users with a set of virtual machines from which the user can benefit by running software on them. The traditional solution to ensure data confidentiality by data encryption is not sufficient due to the fact that the user's data needs to be manipulated in the virtual machines of cloud providers which cannot happen if the data has been encrypted. Administrators manage the infrastructure and as they have remote access to servers, if the administrator isa malicious insider, then he can gain access to the user's data. We present some negative aspects of data encryption in cloud computing. In addition, they assume that if the data is processed from different clients, data encryption cannot ensure privacy in the cloud.

Although cloud providers are aware of the malicious insider danger, they assume that they have critical solutions to alleviate the problem. To determine possible attackers for IaaS cloud providers. For example, propose one solution is to prevent any physical access to the servers. However, arguing that the attackers outlined in their work have remote access and do not need any physical access to the servers. Proposing another solution is to monitor all access to the servers in a cloud where the user's data is stored. We claim that this mechanism is beneficial for monitoring employee's behavior in terms of whether they are following the privacy policy of the company or not, but it is not effective because it detects the problem after it has happened.

We classified four types of attacks that can affect the confidentiality of the user's data in the cloud. These four types of attacks could occur when the malignant insider can determine text passwords in the memory of a VM, cryptographic keys in the memory of VM files, and other confidential data. In addition, they argue that the recent research mechanisms are not good enough to consider the issue of data confidentiality and to protect data from these attacks. This does not mean that these mechanisms are not useful; rather that they do not focus on solving the problems that address in this research. Some of the solutions are mechanisms and are used as part of cloud computing solutions, while different types of solutions focus on solving the whole data confidentiality issue intrinsic to cloud computing. We suggests trusted computing and distributing trust among several cloud providers as a novel solution to solving security problems and challenges in cloud computing. The idea of replicating data among different clouds has been applied in the single system DepSky. We present the limitations of this work which occurs due to the fact that DepSky is only a storage service like Amazon S3, and does not offer the IaaS cloud model. On the other hand, this system provides a secure storage cloud, but does not provide security of data in the IaaS cloud model. This is because it uses data encryption and stores the encrypted key in the clouds by using a secret sharing technique, which is inappropriate for the IaaS cloud model.

5. Future Work

For future work, we aim to provide a framework to supply a secure cloud database that will guarantee to prevent security risks facing the cloud computing community. This framework will apply multi -clouds and the secret sharing algorithm to reduce the risk of data intrusion and the loss of service availability in the cloud and ensure data integrity.

In relation to data intrusion and data integrity, assume we want to distribute the data into three different cloud providers, and we apply the secret sharing algorithm on the stored data in the cloud provider. An intruder needs to retrieve at least three values to be able to find out the real value that we want to hide from the intruder. This depends on Shamir's secret sharing algorithm with a polynomial function technique which claims that even with full knowledge of (k – 1) clouds, the service provider will not have any knowledge of vs (vs is the secret value).

Regarding service availability risk or loss of data, if we replicate the data into different cloud providers, we could argue that the data loss risk will be reduced. If one cloud provider fails, we can still access our data live in other cloud providers. This fact has been discovered from this survey and we will explore dealing with different cloud provider interfaces and the network traffic between cloud providers.

6. Conclusion

It is clear that although the use of cloud computing has rapidly increased, cloud computing security is still considered the major issue in the cloud computing environment. Customers do not want to lose their private information as a result of malicious insiders in the cloud. In addition, the loss of service availability has caused many problems for a large number of customers recently. Furthermore, data intrusion leads to many problems for the users of cloud computing. The purpose of this work is to survey the recent research on single clouds and multi-clouds to address the security risks and solutions. We have found that much research has been done to ensure the security of the single cloud and cloud storage whereas multi-clouds have received less attention in the area of security. We support the migration to multi-clouds due to its ability to decrease security risks that affect the cloud computing user.

References

[1] (NIST), http://www.nist.gov/itl/cloud/.

[2] I. Abraham, G. Chockler, I. Keidar and D. Malkhi, "Byzantine disk paxos: optimal resilience with Byzantine shared memory", Distributed Computing, 18(5), 2006, pp. 387-408.

[3] H. Abu-Libdeh, L. Princehouse and H. Weatherspoon, "RACS: a case for cloud storage diversity", SoCC'10:Proc. 1st ACM symposium on Cloud computing, 2010, pp. 229-240.

[4] D. Agrawal, A. El Abbadi, F. Emekci and A. Metwally, "Database Management as a Service: Challenges and Opportunities", ICDE'09:Proc.25thIntl. Conf. on Data Engineering, 2009, pp. 1709-1716.

[5] M.A. AlZain and E. Pardede, "Using Multi Shares for Ensuring Privacy in Database-as-a-Service", 44th Hawaii Intl. Conf. on System Sciences (HICSS), 2011, pp. 1-9.

[6] Amazon, Amazon Web Services. Web services licensing agreement, October3,2006.

[7] G. Ateniese, R. Burns, R. Curtmola, J. Herring, L. Kissner, Z. Peterson and D. Song, "Provable data possession at untrusted stores", Proc. 14th ACM Conf. on Computer and communications security, 2007, pp. 598-609.

[8] A. Bessani, M. Correia, B. Quaresma, F. André and P. Sousa, "DepSky: dependable and secure storage in a cloud-of-clouds", EuroSys'11:Proc. 6thConf. on Computer systems, 2011, pp. 31-46.

AUTHOR(s)

B. Sujana, received the PG degree in Master of Computer Applications from SV University, 2010 and pursing M.Tech in QCET(2011-2013). She participated in national level conference on Cloud computing at ASCET, Gudur

P. Tejaswini, received the B.Tech degree in Biotechnology from Jippiaar Engueering College Chennai, 2011..and pursing M.Tech in QCET(2011-2013) .she participated in national level conference on Cloud computing at ASCET, Gudur.

G.Srinivasulu, received the PG degree in Master of Computer Applications from SV University..and pursing M.Tech in QCET(2011-2013) .

S. k.Karimulla , received the M.Tech degree in computer science and Engineering from JNTU Ananthapur. At present he is working as asst.professor in quba engineering college. he is dedicated to teaching field from the last 5 years.