Asian Journal of Information Technology 12 (9): 305-311, 2013
ISSN: 1682-3915

A Novel Approach to Enhance Cloud Data Defense

Sudha Devi and Thilagavathy
Coimbatore Institute of Technology, Coimbatore, India

Abstract: Cloud computing is being visualized as a new IT Model for delivering information technology services based on internet. In this model the customer's data are moved to the cloud provider's site where the services and management may not be guaranteed to be completely trustworthy. Cloud computing though suffers against several challenges, data security is the main attribute that act as a barrier to adopt this emerging technology as a new IT procurement model. One of the obstacle that is subjective to cloud data security is data confidentiality and auditability. Cloud users pay for VM to compute data. Data deployed in VM may get leaked in several ways. Either the contents of VM can be hacked by the interference of malicious users or the privileged user accessing VM can leak data intentionally or accidentally. Since, the data owners have no control over their sensitive data in cloud environment it prevents many concerns to adopt cloud computing and avail many of its best services. To solve these issues authentication and access control alone will not be a prominent solution. To ensure reliant security of user's data in the cloud, a better data security technique and data management system should be deployed on this platform. This study addresses data security issues in cloud and proposes possible solutions using crypto techniques. In addition, researchers have discussed about data segmentation and parallelization to manage data efficiently and improve performance. Using proper security solutions, life in the cloud will certainly be successful and advantageous.

Key words: Cloud computing, data defense, data security in cloud, data segmentation, parallelization

INTRODUCTION

The emerging paradigm cloud computing is an Internet-based computing technology. The widely used definition for cloud computing is framed by NIST as cloud computing is a model for enabling ubiquitous, 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. This cloud model is composed of five essential characteristics, three service models and four deployment models. The essential characteristics of cloud model are on-demand self-service, broad network access, resource pooling, rapid elasticity and measured service. The three service models are Software as a Service (SaaS), Platform as a Service (PaaS), Infrastructure as a Service (IaaS) and the four deployment models are public cloud, private cloud, hybrid cloud and community cloud (Mell and Grance, 2011).

The features of cloud computing are incomparable which includes scalability, multi-tenancy, agility, device and location independence, pay for what you use, etc. cloud service providers like Amazon (AWS, 2013) offer these characteristics as services over the internet.

Moving data into the cloud offers cloud users a great convenience since they don't have to take care and handle about the complexities of direct hardware management. Once the requirement for the application is identified, user avails the relevant infrastructure as a service from cloud service providers and moves data into cloud. Infrastructure as a service delivers all form of computing resources as service on demand. The computing resources include hardware, networking, storage services using virtualization technology. Organizations instead of buying and installing the resources in their own data center can rent these resources as needed. Users are totally free away from the technical complexities.

Even though the benefits of cloud computing are tremendous, one unique aspect that hinders the adoption of cloud by many enterprises and individuals is data security issue. Confidentiality and integrity issues are the primary obstacles to the wide acceptance of this emerging trend.

Applications and data are stored on shared servers at some third-party provider's site. As third parties are handling confidential data, the data owner has lack of full control over their data and computation. Cloud users should fully depend on the service providers for availability, privacy and integrity of their data. Even the

Corresponding Author: Sudha Devi, Coimbatore Institute of Technology, Coimbatore, India

provider assures safety measures in these concerns, cloud users in order to fully explore the benefits of cloud should secure their data from their perspective. In this study, researchers propose a novel approach which helps to defend data against security threats from both cloud user and service provider's perspective. This new combined approach includes crypto service, audit service, trapping service and management service which assures secure data access as expected by the cloud users. Using this approach cloud fear regarding data security can be nullified.

CLOUD ENVIRONMENT

Moving data into cloud: The Business processes that are complex to be handled by the organizations internally are allocated to an external service provider which is referred as outsourcing. Once the task is outsourced, it is the responsibility of the service provider to carry out the task successfully. Since, to a third party the data and processes are given, it is essential to analyze the pros and cons of outsourcing.

An organization is in need of concentrating against several threatening. The computational challenge that is threatening many of the organizations is how to collect data and manage such a huge volume of data. The option chosen by many concerns is data outsourcing. Some third party resides on Internet to manage the data. The current trend is to use cloud data storage for data outsourcing. As stated by Hurwitz (2010) that the greatest impact of cloud is its ability to satisfy business requirements efficiently and quickly, enterprises are interested in the journey to cloud computing.

An architecture of the cloud environment is illustrated in Fig. 1. Having enormous storage, the cloud service providers starts their services based on the requirement of the customers on demand. This is how the data is being outsourced in a cloud environment. The entities involved in a cloud environment are:

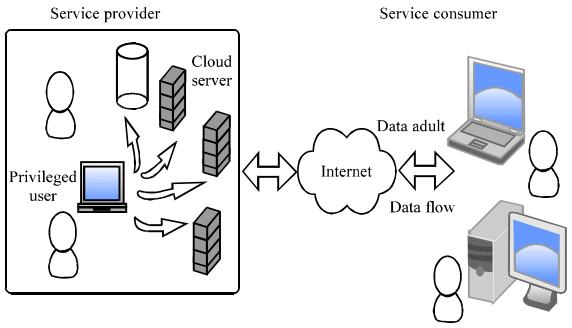

Fig. 1: Cloud environment

- Service provider is the one providing excellence of resources and has responsibility to manage this distributed environment with privileged users
- Service consumer is the one who owns vast amount of data and is in need to rely on cloud for storage and computation

Amazon Web Services, Rackspace, Salesforce, Google, Microsoft, Vmware, Citrix, Verizon, IBM, etc. are the major cloud service providers. Service consumer chooses a service provider who can provide promising services according to their requirement. A cloud user in general compares his need to the available cloud services. Consumers look in for few properties like Flexibility, Interoperability, Security, SLA, Cost, Customer support to choose the right choice of service provider.

Cloud users: Cloud users pay for virtual machines and upload their applications and sensitive data for storage and computation. Cloud users can be categorized into two as shown in Fig. 2. In the first category, users may store data in cloud which can be used for their own use. The user has all privileges to access the data as he is the owner of the data.

In the second category, a user may store data in cloud referred as data owner and the stored data can be accessed by several other authorized users. In this case, the data owner has rights to manipulate and restrict rights to other users and the other authorized users are allowed to do the manipulations according to the privileges assigned to them.

Cloud users in need of security: Even though the cloud service providers has built several measures to protect the

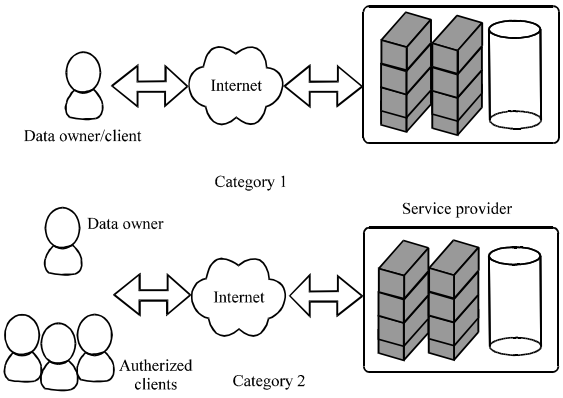

Fig. 2: Cloud users

consumers data and computations against unauthorized and malicious users and assures for security concerns, it is good that users should have some security techniques from their perspective to additionally safeguard the valuable information that is being outsourced and should ensure that the outsourced computation workloads containing sensitive data is secure and the computed results are correct.

The pioneer of Cloud Computing, Amazon Web Services recommends its customers to protect their data using appropriate means and the solution suggested is to run an encrypted file system on top of the virtualized disk device (AWS, 2013). AWS overview of security processes, states that Encryption of sensitive data is generally a good security practice and AWS encourages users to encrypt their sensitive data via an algorithm consistent with their stated security policy (AWS, 2013).

Taking into consideration the importance of cloud data security as discussed by Ren *et al*. (2012), CSA (2009), Dorey and Leite (2011) and Zissis and Lekkas (2012) , researchers propose a security framework which involves crypto services, audit services, intrusion detection service and trapping service.

SECURITY CHALLENGES IN CLOUD

Cloud computing has many unique features compared to the traditional data outsourcing. The characteristics of cloud computing includes scalability, multi-tenancy, agility, device and location independence, cost saving, etc. The scalability on demand feature of cloud is a boon to organizations that are not capable to own and manage huge resources. Cloud has many useful features and although the features of Cloud may be appealing but it suffers with several issues. One unique attribute that hinders the adoption of cloud widely and slows its growth is cloud data security. Confidentiality and privacy requires much attention during outsourcing. Hence, decision making authorities in many organizations do not trust the security of data stored on servers that is shared with other customers in cloud.

The fundamental component of the cloud computing platform is the virtualization technology (YamunaDevi, 2011). In the service model infrastructure as a service, the service provider hosts virtual machines for customers on demand. The privileged user in the provider's site has control over the virtual environment and can monitor that is happening in the VMs. With the dedicated virtual machine the service provider monitor for malware in each of the VM. But anyone with the privileged access being an espionage can use this to manipulate the customer's data.

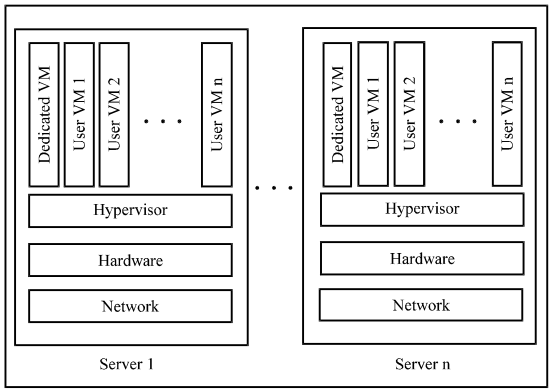

Fig. 3: Server running user's virtual machines at Provider's site

Each VM's virtual disk content are usually stored as a file on the servers. This file can be run by hypervisors on other machines which allows attackers to copy the virtual disk and gain unrestricted access to the contents of VM (Hyde, 2009).

Figure 3 depicts the VMs running in a cloud server at the provider's site. At the provider's site it may happen that a privileged user can give the hypervisor granting permission to a VM that is being dedicated to monitor the other VMs. This dedicated VM gaining permission will know all information such as resource allocation and other details that the other VMs using. Though, it helps in finding malicious attacks, it also paves ways to hack the VM content by an espionage.

The opportunities suggested by many regulations and standards include the requirement for the use of encryption to protect sensitive information. As listed by Armbrust *et al*. (2009), there are several obstacles for adopting cloud computing. Being several issues are in existence, one significant challenge that researchers discussed in this study is data confidentiality and auditability.

The data deployed in cloud servers can be leaked out in several ways. Either the attack may be from outside or from within the provider's site. The privileged user in the cloud service provider site can leak the data either intentionally or accidentally. Hence, it is good to restrict administrative access to the sensitive data by provider's privileged user and also monitoring the access by unauthorized users is also extremely important. To deal with this, researchers categorize the issue into two phases and tried to provide solution as follows:

1 Solution for data confidentiality
2 Solution for data auditability

Phase 1 data confidentiality, includes data security concerns for :

- Data at rest and in transit (a)
- Service provider espionage (b)

Phase 2 data auditability, includes the assurance:

- To perform an audit to prove cloud users that the sensitive data will be handled as they expect (a)

PROPOSED CLOUD DATA SECURITY DESIGN

The goal is to design a security framework for securing cloud users' sensitive data. Considering phase 1, the challenges that are to be solved are:

- Securing data at rest and in transit (a)
- Securing data from service provider espionage (b) and additionally
- Possibility to improve performance after securing data (c)

The proposed cloud data security model confers in enhancing the security to achieve better cloud data defense.

The solution for phase 1a is data owner should change the data in to unintelligent form. That is to secure the data at rest, encrypt the data before moving to cloud helps protect against access to sensitive data by anyone. Since, the user is aware of the critical nature of the data, before moving the sensitive data into cloud it should be encrypted by his own encryption algorithm which should be strong enough to defend data being stored in cloud. The decryption for the first level protection scheme is known only to the user and not to the service provider. Researchers strongly insist on this first level of protection because whatever encryption scheme is being provided by the service provider an insider being an espionage knows how to decrypt and manipulate user's data. Hence, the security schemes provided by the service provider alone are not enough for securing user's data. During uploading, user must use the secure channel for transmission such as SSL. Moving data to cloud through SSL helps protect data during transmission (Dang, 2006). The encryption scheme must be chosen in such a way that the encrypted data are self-defensive. This is the first level of protection given to the outsourcing data. The second level of data protection is given through the security framework.

Compute node is the server in the cloud environment which handles virtual machine creation, deletion and scheduling processes. After acquiring the required

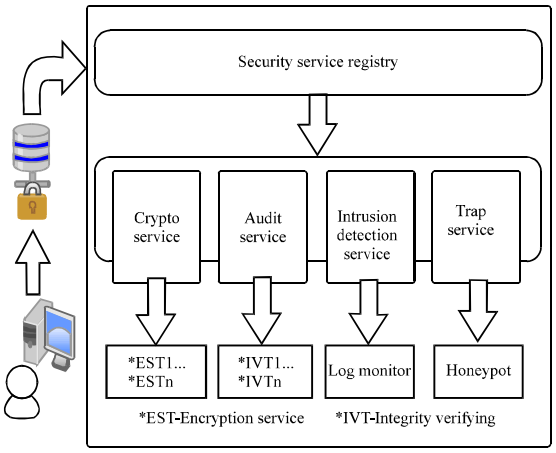

Fig. 4: Cloud data security design

resources, the user enters into the proposed cloud data security framework. The security design comprises of five components as shown in Fig. 4. The user pass in the first component referred as security services registry that is availed in the security framework. The user gets registered with the security service registry. This registry encompasses the list of cloud users using this security framework to enhance their data security. Whenever issue arises for cloud users regarding security, this registry helps in finding the services inherited by the cloud users and the privileges assigned to the cloud users. Backtracking this registry helps in finding where the fault is initiated. Based on the sensitivity of the data the user can determine which encryption service type in the crypto service can be applied. Accordingly the security framework will assign the second level of protection to the user's data.

Researchers suggest two level protection to secure the outsourcing data which will be the solution for the issue stated in phase 1b. Researchers insist to have this two level protection for the following reasons.

Generally, a consumer of service will not fully trust on the security services given by the service provider alone. Since, this security service may protect the data from other malicious outsiders but not completely from the insiders.

Hence, the first level of protection secures data from provider side espionage and the second level of protection through the security framework additionally secures data from both inside and outside attackers. The issue stated in phase 1b can only be overcome through this dual protection. Since, the dual protection includes cloud user security scheme and cloud service provider security scheme, now the data content inside the VM is

safe than before and even if the attacker hacks the VM content in any form, the actual data that is stored and computed is not known and is of no use to the attacker or cloud service provider espionage.

Data owners wish to audit how their data is handled in cloud and ensure that their data is not being altered. Considering phase 2a data auditability, the third component an audit service is included in the security framework. The audit service helps in deploying data integrity verifying techniques. Inheriting an audit service from the security framework can give the assurance to the cloud users that their sensitive data will be handled as they expect.

In the cloud data security framework two more components are designed namely Intrusion Detection service and Trap Service. The action which compromises the security of the information asset is called as the security attack. The security attack that is of type passive is very difficult to detect since, the intention of the attacker is to listen to the data being transmitted and they do not intend to do any alterations in the data. But the security attack of type active involves some modification of the data that is being transmitted. Intrusion is one such security attack on data in which the intension of the intruder is to disrupt the normal working of the system.

The goal of the intrusion detection service deployed in the cloud data security design is to detect the security violations that are not detected and prevented by the available security measures and to improve the quality of security architecture design. That is, it helps in collecting the details about intrusions that takes place and to include proper recovery measures in future to the security framework. If the cloud user inherits this component into the instance, the internal attacker who knows that an intrusion detection service is in use with the instance will be less likely to probe and hack the data. Log monitor is the service that can be inherited into the instance which will keep track of the transactions and check whether any attack is occurring. It maintains a log file which stores the complete transaction that takes place during a session. This will benefit in knowing the security states of the system and sends appropriate alerts to the users whenever it detects an action to be an attack. The monitoring software deployed inside the instance help relates the status automatically whenever a transaction is made with the data to the user.

The trapping service provided in the security framework can be used by the users to notify the presence of an intruder. Honey pot being a trapping system diverts the attackers and make them to stay on the VM processing until proper information is collected regarding the hacking activity of the attacker. As long as the intruder resides on the system the trapping system tries to collect information about the unauthorized access and gives a conclusion based on the network access that the intruder is an insider or an external entity. Users can choose the type of tracking system they need and can import them in their instance. These services will trace for any unauthorised access and let the users know the status.

Assuming the data set moved to cloud is going to be a very large one since then the user opts for cloud, accessing such a large data set will certainly be a tedious process. Along with this trouble, in order to secure the data, researchers apply encryption techniques which results in expanded very large data set. Hence, accessing and retrieving the required content from such a huge volume of data set will be a mind-numbing process. So, along with the encryption technique researchers add some additional measures. This additional measures will be the solution for the issue stated in phase 1c and is also used for dual purpose, one is to enhance the security system and another is to enhance the management of data effectively.

Once the security issues are solved the next question that arises is how to effectively manage the stored data. It is stated by RSA security (2012) that encryption can provide great security and the impacts of encryption can increase data size and decrease performance. To elucidate this fact, researchers bring in the additional measure as stated above. That is, how to efficiently manage the encrypted data stored in cloud that is solution for phase 1c.

The additional measure, is to divide the entire encrypted data set into segments, thereby applying parallelization for accessing the data (Fig. 5).

Applying parallel processing can give prominent solution to manage the encrypted data and to improve its performance. To process the encrypted data in parallel, researchers need to break the large data set into smaller sets and executing it helps to get efficient data

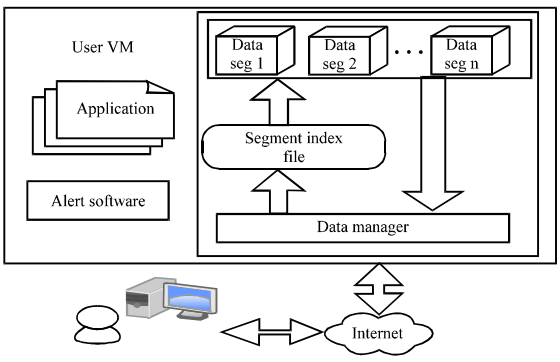

Fig. 5: Improving performance by segmenting data and accessing in parallel

management and faster access and data retrieval. Partitioning efficiently supports very large data by properly segmenting them into smaller and manageable sets. It is revealed that partitioning is useful for many different types of applications, particularly applications that manage large volumes of data. OLTP Systems often benefit from improvements in manageability and availability while data warehousing systems benefit from performance and manageability (Oracle® Database, 2010). Parallel processing gives benefits to improve query performance to optimize resource utilization and minimize execution time. It also enhances the performance and manageability of large data set and ultimately helps to reduce the total cost of ownership for storing large amounts of data (Oracle® Database, 2010).

While considering the data security system, researchers are using the partitioning concept for two reasons: First advantage of parallel execution is to improve the performance by managing the large data set easily. Second advantage is an additional usage that it helps in securing data also. Instead of storing the data set as a whole in the VM, researchers advice to segment the encrypted database into several parts using optimal Hash function. This additional measure will help in protecting the data against unauthorized access attempts. That is, if an unauthorized user gets access into VM fraudulently, the entire data set will be in trouble. Since, the data set is segmented and if an unauthorized user tries to hack the content of VM and if a part of the data is hacked, it is of no use. In the meanwhile the trapping system that researchers inherit from the security framework helps in alerts the data owner regarding unauthorized access to the data set. Segmenting the data in a smart way additionally helps in securing data. The data partitioning is done randomly based on an optimal hash function and the information regarding partitioning is maintained with an index. This index relates information about data residing in segments while accessing the required data. While considering effective data management, the access to the segmented data should be easy, manageable and should take minimum time to retrieve data. By introducing parallel data processing researchers provide a better solution to improve performance.

Finally, if the consumer is satisfied with the security systems and feels that the outsourced data in cloud is secure enough then users relinquish their cloud fear and life in the cloud will be delightful.

CONCLUSION

In this study, one of the issues that hinders the adoption of cloud is taken into consideration and a security framework has been proposed. Researchers aimed to give a solution to overcome the data security threat in cloud computing. Each component in the framework helps to sustain security and privacy in different manner. The user can inherit one or more components into the instance to protect their data and computations. But this dual protection along with other security components may result in a little performance degrade. Instead of compromising security it is good to use this framework and can try to overcome performance degradation. To succeed the performance issue researchers suggested to have data segmentation with parallelism. This will certainly improve the computational performance.

In general many concerns have fear to avail the advantages of cloud due to the loss of control on their sensitive data in cloud environment. In this study, it is proposed to extend security measures from the security service layer into the instance to have a secured data storage and computation. The security measures suggested in the cloud data security framework will certainly alleviate the fear (Chow *et al*., 2009) about secure data storage and computation in cloud computing.

RECOMMENDATIONS

The future research is to design an architecture which can help users to get faster computational services. Working with encrypted data is a monotonous job. So, along with the security features, components that helps to compute the encrypted data comfortably should be deployed. To have better data computational performance in the instance, the mannerism of segmenting the data and working with the segmentation index file should be concentrated and improvisation must be brought in there. The choice of combination of security algorithms has to be analyzed in order to get a better security solution and improved performance.

REFERENCES

AWS., 2013. Amazon web services: Overview of security processes. June, 2013, pp: 1-48. http://media.amazonwebservices.com/pdf/AWS_Security_Whitepaper.pdf.

Armbrust, M., A. Fox, R. Griffith, A.D. Joseph and R.H. Katz *et al*., 2009. Above the clouds: A Berkeley view of cloud computing. Technical Report No. UCB/EECS-2009-28, February 10, 2009. http://www.eecs.berkeley.edu/Pubs/TechRpts/2009/EECS-2009-28.html.

CSA, 2009. Security guidance for critical areas of focus in cloud computing V2.1. Cloud Security Alliance, Security Alliance, USA., December 2009. https://cloudsecurityalliance.org/csaguide.pdf.

Chow, R., P. Golle, M. Jakobsson, E. Shi, J. Staddon, R. Masuoka and J. Molina, 2009. Controlling data in the cloud: Outsourcing computation without outsourcing control. Proceedings of the 2009 ACM Workshop on Cloud Computing Security, Chicago, Illinois, USA., November 13, 2009, ACM, New York, USA., pp: 85-90.

Dang, T.K., 2006. Security protocols for outsourcing database services. Inform. Sec. Int. J., 18: 85-108.

Dorey, P.G. and A. Leite, 2011. Commentary: Cloud computing-A security problem or solution? Inform. Sec. Technical Rep., 16: 89-96.

Hurwitz, J., 2010. The journey to cloud computing: From experimentation to business reality. Hurwitz White Paper. Highland Avenue, Needham, MA.

Hyde, D., 2009. A survey on the security of virtual machines. http://www.cs.wustl.edu/~jain/cse571-09/ftp/vmsec.pdf.

Mell, P. and T. Grance, 2011. The NIST definition of cloud computing. US Department of Commerce, National Institute of Standard and Technology, Special Publication 800-145, Gaithersburg. Maryland, USA., http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf.

Oracle® Database, 2010. Oracle database VLDB and partitioning guide, 11g Release 2 (11.2). E16541-05, August, 2010. http://docs.oracle.com/cd/ E18283_01/server.112/e16541.pdf.

RSA Security, 2012. Securing data at rest: Developing a database encryption strategy. A White Paper for Developers, e-Business Managers and IT.

Ren, K., C. Wang and Q. Wang, 2012. Security challenges for the public cloud. IEEE Internet Comput., 16: 69-73.

YamunaDevi, L., P. Aruna, D.D. Sudha and N. Priya, 2011. Security in virtual machine live migration for KVM. Proceedings of the International Conference on Process Automation, Control and Computing, July 20-22, 2011, Coimbatore, pp: 1-6.

Zissis, D. and D. Lekkas, 2012. Addressing cloud computing security issues. Future Generation Comput. Syst., 28: 583-592.