DATA STORAGE SECURITY CHALLENGES IN CLOUD COMPUTING

Sajjad Hashemi¹

¹Department of Computer Engineering, Science and Research Branch, Islamic Azad University, West Azarbayjan, Iran
laux.hashemi@gmail.com

ABSTRACT

In the digital world using technology and new technologies require safe and reliable environment, and it also requires consideration to all the challenges that technology faces with them and address these challenges. Cloud computing is also one of the new technologies in the IT world in this rule there is no exception. According to studies one of the major challenges of this technology is the security and safety required for providing services and build trust in consumers to transfer their data into the cloud. In this paper we attempt to review and highlight security challenges, particularly the security of data storage in a cloud environment. Also, provides some offers to enhance the security of data storage in the cloud computing systems that by using these opinions can be overcome somewhat on the problems.

KEYWORDS


1. INTRODUCTION

Cloud computing are an important structure with outstanding potential in decreasing the costs by improving and developing functionality and economic outcome which in turn can increase cooperation, pace and scalability acceptance to comprehensible degree [1, 2]. This technology has given many opportunities to large corporations and IT companies in developed countries, however, these opportunities face with challenges like security that is one of the most important concern in the field of cloud computing [2, 3]. If security services use badly the all parts of cloud computing face with problems such as the management of personal information in a public network or stored data users on the servers providing cloud services[2].It can be expressed that safety is a virtual highway to the adoption of the cloud, if the providers of this technology can destroy the obstacle from the path or minimizes it, cloud computing will be an important factor in the field of information technology, so it is easier to companies and public to accept and trust to use it [2]. Today, the main concern in cloud computing is how to make confidence in accepting, sharing applications, hardware, etc., in an environment that we don’t know who is responsible for securing our data [2, 4].So to build trust and develop the cloud computing use, it feels the need to repair the security flaws and minimize the challenges are necessary.

In this paper, security issues in cloud technology, with emphasis on security challenges in the field of saving data examined. In addition to research in the field of cloud security and the issues and weaknesses in their investigation and decided to solutions for a better offer, and we refer to this issue that we build trust with consumers to transfer their data into the cloud and store them in the server side of a provider of cloud services necessary to develop and expand to use green

DOI: 10.5121/ijsptm.2013.2401
technology by users. Also, due to economic problems in the world and reduce the purchasing power of people the act of decrease security challenges and enhance public confidence in the services provided by the technology of cloud computing will follow economy for many people and governments.

2. PREVIOUS WORKS

Since the creation of computer networks and the expansion of Internet security issues of data transfer and storage, it was an important and growing importance of the subject is enhanced because the advancement of technology and the transfer of data from high-volume, high-importance requires channels with a greater safety factor for transferring data is felt. Accordingly, in this section we review presents offers and prior business to improving data security, especially in a cloud environment. Tsai W, et al within [5], framework of four-layer for the development of Web-based was made it interesting, but only one aspect of security in this process is discussed. Sources separation offer’s take place to ensure data security during the process, by separating processor’s cache in the virtual machines and separation of the virtual cache from hypervisor cache [6]. In reference [7], a security framework by different methods provided dynamically, that one of the components of this framework refers to provide data security by storage and access to data based on meta-data, which is similar to storing related data in different areas based on meta-data, and if the destruction of user data takes place, it can be retrieved. Each part of the framework in "security as a service” is provided for practical applications by providers of security as a layer or multiple layers of required applications [7]. This research explains the concept of cloud security and the security system in the real world where security is depend on poses of individuals and organizations. Perhaps this is a good offer, but it should be clear that is security as a service provides with delivers service? In this case, the service provider must be put part of its focus on providing security and this is not good because it maybe decrease the growing of providing application services [7, 8]. M. Ahmed et al. [9], the accuracy of certain security issues related to cloud computing have examined and its aim is to explore and establish a secure channel for communication NO with the CSP, while the reliability and confidentiality of information is maintained. In addition, they have compared the provided protocol by the SSL of the activities associated with the work, along with the trustworthy security way to securing data. In the paper [10], the security problems at different levels of the architecture of cloud computing services have been studied. Security of customer-related data is a substantial need for services which is provided by each model of cloud computing [10]. They have studied matters of on-going security software as a service (SaaS), platform as a service (PaaS) and Infrastructure as a Service (IaaS). This paper focuses on the use of cloud services and security for working cross-domain Internet-connected [10].

3. CLOUD COMPUTING

Cloud computing have different shapes causing different understandings of them. For this reason many consider cloud as web-based applications. Others see it as useful and/or parallel computing because cloud is planned for better and improved efficiency in complicated & large scaled processes [12], [13]. Besides different shapes of cloud, the provided services are also highly different & numerous. Cloud computing have various definitions which some have been brought here. The definition of the national institute of standard and technology of America is as follows [2], [12], [14], [15]: “Cloud computing is 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.” On other common
& acceptable definition is of Mater et al. [4], [16] “A highly scalable tool, with the ability of service technology-enabled, which can be easily used via the internet when needed.”

Following the definition of cloud computing, we should comprehend their important features, developed models, the way of using services and also the way of protecting it, in order to know well and accept it [15]. Here are the five key features of cloud computing [2], [17]:

- **Service demand on self.** Using this feature when needed the customer can easily and automatically access to computing facilities like server, net, storage and soon from any provider.
- **Ubiquitous network access.** It implies that the facilities are accessible on the net and they can be used following standard methods. The methods which support weak and strong clients like laptop and mobile phones.
- **Location-independent resource pooling.** This features pools different customers needed resources in the same place dynamically by the providers. These resources can include the storage, memory, the bandwidth of net and virtual machines.
- **Rapid elasticity.** Using this feature, the facilities can be provided rapidly and with high elasticity and can be expanded or release fast. In other words the services can always be updated and improved and accessible for the users.
- **Measured service.** This feature enables monitoring, control and reporting of the resources, and can apparently control and report the amount and quantity of resource using for both customer and the provider of the infrastructure. In other words all these features cover the coherence and appearance of the clouds [18].

Development models include the aim and identity of cloud and the way they are settled. NIST definition of the development models are of the following four types [1], [2].

- **Public cloud:** The substructure of public cloud is for the public use and accessible to all in which the resources, applications and web-services are provided through internet and public organizations help to provide and supply the substructures [19]. Indeed a cloud service provider organization owns the public cloud.
- **Private cloud:** Private cloud is for the exclusive use and only for an organization, so everyone in the organization can access data, services and applications but others out of organization can’t [19].
- **Community cloud:** Community cloud are provisioned and prepared to make some common facilities and resources available. Its substructure can be shared between one or some organizations but the main point here is that the requested work of them is the same and the demanders follow somehow the same mission, policy, security and soon. In Community cloud a certain group support tasks like security needs. Of course this kind of sharing will have consequences for the organization at work [20].
- **Hybrid cloud:** And latest model is hybrid cloud, which is combination of two or more clouds (including public, private and community). It is in fact an environment which uses some internal and external cloud providers [19].

Various services of cloud are presented into three models which are [1], [2]:

- **SaaS.** The services provided by SaaS include using functional programs on the infrastructure of cloud and access through the web browser [21]. in this section the customer doesn't manage the infra-structure of cloud including the net , servers, operational systems, and saving area, except the functional software to limited degree of adjustment at the level of user [20].
- **PaaS.** In this kind of service, the client has the option of putting the purchased functional programs on the infra-structure of cloud [21]. Here also the client dose not mange or
control infrastructure of the cloud such as the net, servers, storage. He just has control over the functional program installed or Settled by him [20]. In fact the PaaS is similar to SaaS the difference is that PaaS includes exclusive program environment and computing platform, developing and solution strategies [2].

- **IaaS.** This kind of service providing includes process potential, saving space. Nets and other basic computing re-sources and even operational system and application pro-grams [20]. The client dose not manage or control in infra-structure but has control over the operational system, saving area, and the established programs. In this service an artificial server is completely available for the client [21].

4. **SECURITY IN CLOUD ENVIRONMENT**

In the IT industry, the first factor that will ensure the success of a system is information security [8]. Cloud computing is also a component of IT area, and in this rule there is no exception and according to that the users of this technology don’t know where and how their information stored, the role of providing security becomes more important [2]. A number of security problems in cloud computing are provided with cloud models and others are provided by providing services models. So security risks largely depend on service model and deploy of cloud [11] and this reduces the reliable services offers this technology, offered Services by cloud computing because it provides services to the Internet, the protocols for normalization and security measures on the Internet are used to create the encryption and the security fixed part of needs somewhat, but they aren’t text-oriented and, therefore, a powerful set of policies and security contracts for the secure transmission of data in the cloud is needed [2, 7, 8]. The subjects expressed below are refers to some of the risks of security-related data in cloud computing [22, 23].

1) **Data position:** In the cloud environment, the issue of location of an organization or company data, (a place to host them, or even the country that data is located) is very challenging and controversial. One step to securing data is that to agree with provider cloud service to store and process your data in a specific geographical place. You can also use the legal obligations and require them to observe the integrity of your data.

2) **Data separation:** because in the cloud environment, many of data are stored in a shared environment. So customers’ data are together in a cloud. Provider cloud service must assure customers to make the data separated.

3) **Flexibility of servers:** Its full flexibility is one of the advantages of cloud computing and the fact that of this technology [23]. But this matter can cause some problems. Some servers may reconfigure back often without warning to the user. This matter can be challenging for the cloud inside’s data that related to a specific organization. So if these changes continuously be applied the data security threatened.

5. **DATA STORAGE SECURITY IN CLOUD COMPUTING**

Security of data storage refers to data security on storage media, which means non-volatile or fast retrieving after the loss [24]. The security should be to consider by software engineers in the design phase of cloud storage. This not only includes redundancy and dynamic data, but also includes the separation as well. Redundancy is one of the most basic measures to protect the security of data storage, and dynamic means that the user’s information often may change, so effective movements need to ensure the consistency of data. Separation means the time of storing user data in the platform. In order to guarantee the independence of the data, the user can access only to own data and data changing from other users don’t affect current user [24]. In this section we will investigate cloud storage topics and needs and security solutions that around it.
5.1. Cloud Storage

Cloud storage is an online distributed virtual storage that provide by the vendor of cloud computing. [24]. We can access to the service of cloud storage via a web services interface, or a web-based user interface. In cloud data storage system, there is no need to store data locally and users can store their data in the cloud. Therefore, the accuracy and availability of the data files should be guaranteed that they stored on the distributed cloud servers [25]. One of the benefits of cloud storage is its elastic feature that customers can rent the storage space whenever they want to store their data and only pay the amount of their use. [26]. Organization economies dramatically in the storage devices and complexity and its cost by using cloud storage. As cloud computing, cloud storage has also several features like scalability and agility to the cloud storage advantages also there are security problems which like those problems are in the distributed storage system. [26].

5.2. Cloud Storage Security: Requirements & Solutions

In the process of storing data on the cloud, and retrieves data from the cloud, mainly three elements are involved: the client, the server and the connections between them [24]. All three elements must have strong security in order to make necessary of data security. Client is responsible for ensuring that no other unauthorized person can access to the device. When we talk about the security of cloud storage, our purpose is more about the other two elements means server and the connection between server and user and our main concern is in this case. On the server side, the data should be confidential, consistency and available [25]. Privacy and consistency of data can be ensured on the both side of server means server side and client side. Availability of data to the server is guaranteed, and then the server should always ensure that data always are available for retrieving [24]. The last element of importance also is the connection between the server and the client. Communication between client and server must be through a secure channel, i.e. the data should be confidential and consistence during the transfer between server and client. One way to achieve to a secure communication is an encrypted protocol such as SSL [25]. Amazon and Google are the names of two cloud storage provider’s brand that they act as large and well known providers in the market. Drop box is popular also as the provider of cloud storage and file sharing service. In addition, there are many providers of cloud storage that they use different security mechanism like encryption. Some of security solutions that offered or used are mentioned in this research and we refer to ways that they compared. For example, some types of data, such as data on digital libraries. The consistency of data is the main concern but their confidentiality is not so important. In this case, it is vital to have a fast and a simple connection’s mechanism in order to investigate the consistency of data. To achieve this objective, the two methods [25, 27] are recommended.

One is called Proof of Retrievability Schemes (POR), One is called Proof of Retrievability Schemes (POR), a challenge-response protocol which is used by cloud storage provider to demonstrate to the customer that their data is recoverable without corruption or loss. The latter method is Provable Data Possession Schemes (PDP), that also is a challenge-response protocol too, but much weaker than POR, because it does not provide a guarantee for data recovery. These two methods are rationally fast processes, Because Data recovery is confirmed without re-downloading the data [20]. For many other types of users, the confidentiality of their information is important. So many providers of cloud storage business offer secret solutions to their customers.

In comparison with mechanisms used in cloud storage, cryptography access control is client-centred [26]. A user has more control over his local computer. If the data is encrypted locally, they will have more security and confidentiality. Amazon has offered a library called “Amazon
S3 Encryption Client”, which makes local encryption possible, but every user has to implement the all of mechanism using the defined library [26]. This is very complicated and time-consuming for many users which are not familiar with this technology, beside most users prefer to use a ready-made system.

6. EXISTING ALGORITHMS FOR DATA STORAGE SECURITY

6.1. RSA algorithm

Today RSA algorithm is one of the public key cryptography algorithms used for encryption and decryption by many vendors. This is the first generation algorithm that used for providing security to data [28]. It can encrypt a message without the need to exchange a separate secret key. The RSA algorithm can be used for both public key encryption and digital signatures. Its security is based on the difficulty of factoring large integers. Party A1 can send an encrypted message to party B1 without any prior secret keys exchange. A1 uses B1’s public key to encrypt the message and B1 decrypts it using the private key, which only he knows. RSA can also be used to sign a message, so A1 can sign a message using their private key and B1 can verify it using A1’s public key [28]. The RSA algorithm contains three steps, namely key generation, encryption and decryption. Here is the generating of key process which first is choosing two random numbers p and q. Then the number n should be computed [25]: n = pq.

Thereafter a function φ (n) is computed: φ (n) = (p – 1) (q – 1).

Also an integer e is chosen so that 1 < e < φ (n).

Finally, the value of d is calculated: d = e-1 mod φ (n), such that: de mod φ (n) = 1, and e and φ (n) are co-prime. The result (n, d) is the private key and (n, e) is the public key.

Encryption a text m is calculated by: c = me mod n, and decryption a text is calculated by: m = cd mod n [29].

6.1.1. RSA algorithm’s Security

There are a lot of initial attacks on RSA, which are not so powerful, because there are improvements added to RSA, but one of the most famous one of these improvements is on use of common modulus for all users, for example not to choose different n = pq for each user. This problem can occur in a system, where a trusted central authority generates public and private keys for users by using a fixed value for n. In this case user A1 can factor the modulus n, using his own exponents, e and d. Then A1 can use the public key of B1 to recover his private key. The solution is simply not to use a different n for each user. This attack is not applicable in systems, where each user generates the pair of keys on his/her machine. Here the value of n must be different for each user [25]. One of other attacks on RSA is called timing attack. When a user A1 uses RSA algorithm for digital signature or encryption/decryption, a troublemaker can specify the private key by measuring the time it takes to execute signature or decryption. This attack could be applied on the systems that are connected to a network, for example, using a smart card. The intruder cannot read the smart card's content, because it is resistant to unauthorised access, but he can determine the private key by using timing attack. One of the possibilities to Deal with timing attack is to add some delay to the process, because the process always takes a fixed amount of time [25, 29].
6.2. Elliptic Curve Cryptography (ECC) algorithms

Elliptic Curve Cryptography (ECC) is a secure and more efficient encryption algorithm than RSA as it uses smaller key sizes for same level of security as compared to RSA. For example, a 256-bit ECC public key provides equivalent security to a 3072-bit RSA public key [28]. Elliptic Curve Cryptography (ECC) was introduced in 1985 by Victor Miller (IBM) and Neal Koblitz (University of Washington) as an alternative mechanism for implementing public-key cryptography. Public key algorithms provide a mechanism for sharing keys among a large number of participants in a complex information system. Compared to other famous algorithms such as RSA, ECC is based on discrete logarithms that are much more difficult to challenge at similar key lengths [28]. Every participant in the public key cryptography will have two keys, a public key and private key, used for encryption and decryption operations. Public key is distributed to all the participants where as private key is known to a particular participant only.

6.3. Data Encryption Standard (DES)

Data Encryption Standard (DES) is a block cipher with 64 bits of block size. It was developed by IBM in the 1970s, and adopted in the United States of America as a standard encryption technique in 1976. Firstly, it was mostly used in the United States of America, and then it became more and more popular around the world. DES is using substitutions and transpositions one after another in 16 cycles in a very complicated way [25]. For this algorithm, key length is fixed to 56 bits, which seems too weak while it has been proved that the power of computing resources is getting more and more. However it is useful to mention that 3DES, also called triple DES, is a method to make DES more difficult to decode. 3DES uses DES three times on every data block, and in this way the length of the key is increased. In fact it uses a “bunch of keys” containing three DES keys, K1, K2 and K3, which each of them is 56 bits [25].

6.4. Advanced Encryption Standard (AES)

The weakness of the DES has been accepted; In January 1997 NIST (National Institute of Standards and Technology) announced that instead of DES, a new method will be used as the AES (Advanced Encryption Standard). It led to a competition between the open cryptographic community members, and in nine months, NIST received fifteen different algorithms from several countries. In 1999, from the received algorithms NIST choose the algorithm “Rijndael”, which was developed by two Dutch cryptographers, Vincent Rijmen and Joan Daemen [25]. This algorithm officially became the encryption algorithm for AES in 2001. AES is a block cipher with 128 bits of block size. In AES key length is variable (not fixed), then it can be 128, 192, and 256 bits (and probably more). The structure of AES is mainly created from Encryption techniques such as substitutions and transpositions [25]. Same as DES, AES uses repeated cycles, which are 10, 12 or 14 cycles (called rounds in AES). In order to achieve perfect confusion and diffusion, every round contains four steps. These steps are substitutions, transpositions, shifting the bits and applying exclusive OR to the bits [30].

6.4.1. Modes of Operation for AES

AES is a block encryption, so data is divided into various blocks and each block gets encrypted separately. A mode of operation could be defined as the process that is performed during encryption/decryption to each data block [25]. Most of operation modes are using a vector known as initialization vector (IV), which is a block of bits. IV is used to make the encryption process random, such that the corresponding cipher text would be different each time even if the same data is encrypted several times. There are many modes of operations, but NIST has approved six modes for the confidentiality of data, namely ECB (Electronic Codebook), CBC (Cipher Block Chaining), CFB (Cipher Feedback), OFB (Output Feedback), CTR (Counter), and XTS-AES [31].

6.4.2. Security of AES

Regarding the security of AES, there has not been found any flaws in the algorithm. This algorithm has been used two years to analyse before it was approved in USA. It is enough to prove the quality and integrity of AES. Security of key is strong in AES, because the minimum key length is at least twice the length of the key used for the DES [25]. The rounds and Key length are not limited, so if we come close enough to break key by using powerful computing resources, it is possible to increase the key length and the number of rounds. There is a website [32], where researches and activities about AES are stated. The last research paper about Security of AES is presented in 2009 [25, 33]. It is an AES encryption with key lengths of 192 and 256 bits. Here an attack is used named as “Related Key Boomerang”. The paper concludes that this attack is only theoretical threat, and in action it is not possible. The complexity of data and time are so high, which it’s almost impossible to handle with the existing technology [25].

7. CONCLUSIONS

As noted in the system of cloud data storage, users store their data in the cloud, so there is no need to store them locally. Therefore, the security, integrity and availability of data files on storage distributed cloud servers are guaranteed. To accomplish this, the structure and security solutions of involved elements in the process of data storage in the cloud environment should be investigated. About the first element: client; we suggest to use an encryption mechanism from the customer like AES encryption that its high security and resistance has been proven in many testing. AES has been investigated and analysed by the NIST and its security has been approved by this validated Institute, and this encryption is used to encrypt sensitive information in the United States of America. Also we can use encryption algorithm by means of new methods like genetic algorithm or other dynamic algorithm which security can increase dramatically in this way. The next element must gave special consideration to its security is server, because our data store on the server and we possess storage space virtually as a user. Therefore, the accuracy and availability of data and information retrieval is very important and should provide the necessary security to accomplish this on the server side. Therefore, we use a comparison between some security policies by providers known in the field of providing data storage services, we did the comparison can be clearly seen that in order to the confidentiality of information, some providers use the mechanism of encryption control such as symmetric encryption. About the security of our server recommended service providers in this field to expand and to improve security mechanisms on their servers, because the users of cloud technology will go to the side of those providers that their services have enough security, thus server security will be important and providers can success in this technology with high server security and accountability to the users. The third element that its security is important in the storage and transmission of data is the connection channel between cloud service providers and user. In our opinion, the most vulnerable point that can put user’s data and information in the cloud environment at risk are communication channel. Because of the Internet and in most cases of the old mechanisms, therefore we must use new methods in order to avoid of unauthorized influences. In this case we can refer to the established protocols and retrieving or establishing more secure transmission channels that they introduce by using new sciences and methods in the computer science.
REFERENCES


[33] Alex Biryukov and Dmitry Khovratovich, Related-key Cryptanalysis of the Full AES-192 and AES-256, University of Luxembourg, ePrint Archive: Report 2009/317

AUTHORS

Sajjad Hashemi is a M.Sc. student in Computer Engineering Department, Science and Research Branch, Islamic Azad University, West Azarbayjan, Iran.
Email: iau.hashemi@gmail.com