Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 125 (2018) 691–697

Procedia Computer Science

www.elsevier.com/locate/procedia

6th International Conference on Smart Computing and Communications, ICSCC 2017, 7-8 December 2017, Kurukshetra, India

Exploring Data Security Issues and Solutions in Cloud Computing

P. Ravi Kumar*[a], P. Herbert Raj[b], P. Jelciana[c]

[a]School of ICT, IBTE JB Campus, MOE, Kuala Belait, Brunei
[b]School of ICT, IBTE SB Campus, MOE, Seria, Brunei
[c]Laksamana College of Business, Bandar Seri Begawan, Brunei

Abstract

Cloud computing is one of the fastest emerging technologies in computing. There are many advantages as well few security issues in cloud computing. This paper explores the different data security issues in cloud computing in a multi-tenant environment and proposes methods to overcome the security issues. This paper also describes Cloud computing models such as the deployment models and the service delivery models. In any business or Cloud Computing data are exceptionally important, data leaking or corruption can shatter the confidence of the people and can lead to the collapse of that business. Currently cloud computing is used directly or indirectly in many businesses and if any data breaching has happened in cloud computing, that will affect the cloud computing as well as the company's business. This is one of the main reasons for cloud computing companies to give more attention to data security.

Peer-review under responsibility of the scientific committee of the 6th International Conference on Smart Computing and Communications.

Keywords: Cloud Computing; Data Security; Cloud Services; Confidentiality; Integrity; Availability; Authentication and Access Control

1. Introduction

According to National Institute of Standards and Technology (NIST), "*Cloud computing is a model for enabling*

*Corresponding author. Tel.: +673-8171484
E-mail address: ravi2266@gmail.com.

1877-0509
Peer-review under responsibility of the scientific committee of the 6th International Conference on Smart Computing and Communications
10.1016/j.procs.2017.12.089

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 and or service provider interaction" [1]. Cloud computing is one of the fastest emerging technologies in Computing. Everyone is using cloud computing in our day to day life in one form or another without realizing it, like Microsoft Office 365, Gmail and Dropbox etc. There are many advantages of using cloud computing such as anytime-anywhere accessibility, better geographic coverage with the fastest time, less investment on infrastructure, etc., but there are also challenges using cloud computing like data security, lack of resources and expertise etc. Among the challenges, data security stands very tall and this paper explores the challenges of data security in cloud computing and provides methodologies to overcome the data security challenges. This paper is organized as follows. Section 2 introduces the cloud computing model. Section 3 provides different challenges in data security and provides solutions to the challenges. Section 4 concludes this paper with new developments and challenges in cloud computing.

2. Cloud Computing Model

There are five major actors [1] in cloud computing based on their participation as shown in Fig. 1. Cloud consumer or cloud service consumer (CSC) is the one who gets the service from a cloud provider and pays for the service as per the use. Cloud provider or cloud service provider (CSP) is the one who provides the cloud services to the CSC. Cloud auditor is the one who conducts an independent assessment of cloud services, information system operations, performance and security of the cloud implementations. Cloud broker is the one who interacts between CSP and CSC to make the business happen. Cloud carrier is the one who provides the connectivity and cloud services from CSP to CSC.

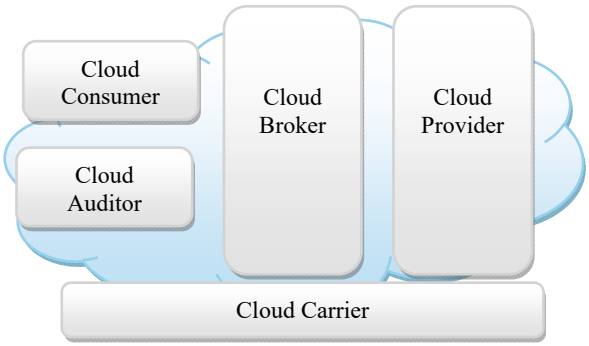

Fig. 1. Cloud Computing Actors

NIST based cloud computing model consists of four cloud deployment models, three service delivery models and five essential characteristics. A Cloud can be deployed as Private, Public, Community and Hybrid clouds. Infrastructure-as-a-Service (IaaS), Platform-as-a-Service (PaaS) and Software-as-a-Service (SaaS) are the three service delivery models have become widely recognized and formalized. Rapid Elasticity, Measured Service, Resource Pooling, Broad Network access and On-Demand Self Service are the five essential characteristics of cloud computing [1, 2 & 3]. Other researchers [4, 5 & 6] says that multi-tenancy is also an important characteristic of cloud computing. The following table 1 shows that a cloud computing delivery model can provide nine services, namely Application, Data, Runtime, Middleware, Operating System, Virtualization, Server, Storage and Networking which are the components in the traditional computing [7, 8, 4].

Table 1 clearly determines the roles of CSP (shown in boldface) and CSC in the delivery models' services. In IaaS, the CSP provides only the infrastructure like server, storage, networks and virtualization. The CSC is responsible for Application, data, runtime, middleware and operating systems. In PaaS, only the Application and Data are CSC's responsibility and the rest of the services are provided by CSP. In SaaS, all the nine services are provided by CSP. The security issues are also subject to the cloud deployment models.

Table 1. Services Provided by Delivery Models and the Responsibility

Services Provided by Delivery Models	IaaS	PaaS	SaaS
Application	CSC	CSC	**CSP**
Data	CSC	CSC	**CSP**
Runtime	CSC	**CSP**	**CSP**
Middleware	CSC	**CSP**	**CSP**
Operating System	CSC	**CSP**	**CSP**
Virtualization	**CSP**	**CSP**	**CSP**
Server	**CSP**	**CSP**	**CSP**
Storage	**CSP**	**CSP**	**CSP**
Networking	**CSP**	**CSP**	**CSP**

Public cloud is more prone to security issues compared to other cloud models because it is open to all and it is using the ubiquitous internet as the connection medium. So, security policies and procedures should be different from one deployment model to another. Generally, CSPs are responsible for the cloud infrastructure and CSCs are responsible for the data and other things they are stored in the cloud. So, security becomes a shared responsibility between CSP and CSC. According to Amazon Web Services (AWS), security responsibility is shared by both CSP and CSC and they called it as Shared Security Responsible Model [9]. There are many security issues or challenges in all the nine services provided by the cloud delivery models. This paper focuses only the data security issues and provides solutions to the issues.

3. Data Security Issues or Challenges

In enterprise computing, data is stored within their organization and it is fully under the control of the enterprise [10]. In cloud computing, the data is stored outside the customer's place (in the CSP's side). So, cloud computing must employ additional security measures apart from the traditional security checks to ensure that data is safe and no data breaches due to security vulnerabilities.

3.1. Data Security Basics

According to [6], there are six stages in the life cycle of data: Create, Store, Use, Share, Archive and Destroy. Once the data is created, it can move freely between any stages. Data should be secured in all the stages of its life cycle from its creation to its destruction. The store and archive stages are also called as data-at-rest, the use stage is called as data-in-use, the sharing stage is called as data-in-transit and the destroy stage can be called as data-after-delete. All these stages are self-explanatory. Generally, encryption is one of the methods in the data-in-transmit stage to protect the data. One of the neglected issues is data-after-delete [11] and this is also called as data remanence. Data remanence is the residual physical representation of the data that has been deleted [12]. After a storage media is deleted, there may be some physical characteristics that allow the data to be reconstructed [12, 13]. Tracing the data path (data lineage) is important for auditing in cloud computing, especially in the public cloud apart from the above stages [11].

Confidentiality, Integrity and Availability are the three important properties of the data and it is popularly called as CIA triad. Authentication, authorization and nonrepudiation are another three important properties associated with people who access the data [14]. Confidentiality refers to data privacy where the data belongs to CSC is not revealed to unauthorized parties on any occasion [15]. The Integrity of data refers to the confidence that the data stored in the cloud is not fiddled by unauthorized parties. It is also applicable when the data is in transit. Availability of data refers to pledge that whenever the CSC needs data, the data should be available to them without any delay or deny. These three basic data security properties are tested a lot in the pubic cloud deployment model. Authentication is the proof for a person to access his or her own data. Authorization is the act of determining whether a person has the right to perform an activity on data like reading or writing. Users must be authenticated before carrying out the

activity they are authorized to do. Nonrepudiation is the assurance that an authenticated user cannot deny after performing a job. The following fig. 2 shows the four major categories of data security issues in cloud computing.

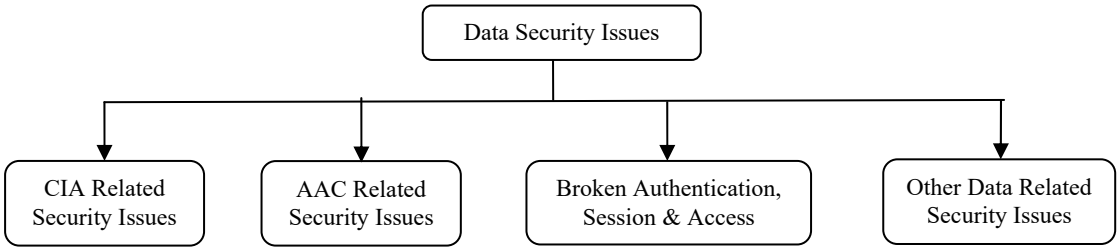

Fig. 2. Types of Data Security Issues

3.2. Security Challenges in the CIA Triad

Confidentiality, Integrity and Availability (CIA) losses can make a big impact in the business of the cloud computing because the data is the core component for any business. Data integrity is the assurance given to the digital information is uncorrupted and only be accessed by those authorized users. Thus, integrity involves maintaining the accuracy, consistency and trustworthiness of data over its entire life cycle [16]. Maintaining CIA is easier in enterprise computing but in cloud computing it is more complicated because of the multi-tenant architecture and the distributed nature of the infrastructure. The following steps can be used to maintain a proper CIA in cloud computing:

- Once the data are created, classify the data, identify the sensitive data, define policies, and create access methods for different types of data. Also, create policies for data archive and data destroy.
- Store data with proper physical and logical security protection, including the backup and recovery plan.
- Identify which type of data can be shared, whom and how it can be shared and define data sharing policies. In cloud computing, many such policies are collectively called as Service Level Agreements (SLA).
- Create a corrective action plan in case data is corrupted or hacked due to network or communication devices, security flaws while data is in transit.

According to Aldossary and Allen, integrity should be checked not only at data, but also at the competition level [4]. Computation integrity refers to only the authorized applications are allowed to access the data and use it for computation. Any abnormality from normal computing should be avoided. An effective Identity and Access Management (IAM) can avoid loss of confidentiality and integrity. Loss of availability can happen through loss of data and data inaccessibility. Cloud computing employs few techniques like scalability and high availability at the architecture level. There are different methods and procedures are followed to improve data security related to the CIA triad at different stages of the data lifecycle. Some of the important methods are listed below:

- Apply data encryption when the data is at rest and also when the data is in transit. Apply strong encryption algorithms like Advanced Encryption Standard (AES) and Rivest Shamir Adleman (RSA) algorithms. Different types of encryption methods are described in [4, 17]. Amazon S3 uses one of the strongest encryption algorithms, 256-bit AES [9].
- Encryption methods are generally to provide confidentiality against attacks from a cloud provider, but it cannot protect data against configuration errors and software bugs [4]. Hash methods can be used to find out accidental and intentional data changes. But they consume more bandwidth and time-consuming.
- Third Party Auditing (TPA) can be employed to check for the data integrity. Many researchers [4, 18, 19] insists to audit data integrity by third-party auditors because they are specialized in that.
- Provable Data Possession (PDP) scheme was initiated by Ateniese et al. to investigate statistically the correctness of the data without retrieving the data outsourced to cloud storage [20]. PDP limitations were overcome by Ateniese et al. themselves at [21, 22], Wang et al. at [23] and Sookhak et al. at [24].
- Do not store the encryption keys along with the encrypted data [25].
- Implement proper Identity and Access Management (IAM) techniques for users to access data.

- Use data duplication, redundancy, backups and resilient systems to address availability issues [25].
- Also, include a failover strategy in case the service fails with the CSP.
- The data dispersion technique can be used to address the availability issue if other methods are not effective. Here the data is stored as fragments in many clouds and the data can be reconstructed when it is required to use fragmentation techniques [6].

3.3. Security Challenges in the Authentication and Access Control (AAC)

Authentication and Access Control (AAC) is the process of verification and confirmation on user's identity to connect, to access and use the cloud resources. In enterprise computing, the credentials are stored in the server in the form of Active Directory (AD) or Lightweight Directory Access Protocol (LDAP). In a private cloud, the authentication is done same as the enterprise computing via a virtual private network. In public cloud, customers use the internet to connect to CSP, applications from different users can co-exist with the same CSP (resource pooling) and CSC can access the applications from anywhere through any devices. So the authentication in public cloud is more subject to vulnerability than private cloud [6]. A Password-based authentication does not provide effective security for the public cloud. Passwords can be cracked using many methods such as a brute force attack, dictionary attack, phishing or social engineering attack. So it is very important that the CSP should include highly secured authentication methods in a public cloud. In cloud computing, customers connect to cloud services through APIs and these API's are designed to accept tokens rather than passwords [6].

In cloud computing, authentication applies to not only users but also to machines. Machines need to authorize certain automated actions like online backup, patching and updating systems and remote monitoring systems [26]. Since the cloud applications are accessed through various devices, there should be a strong authentication method like RSA token, OTP over the phone, smart card / PKI, biometrics, etc., for the original identity confirmation and determine the type of credentials [6]. This will enable identifiers and attributes with a strong level of authentication to be passed on to the cloud application and the risk decisions can be made for access management. There are a number of methods and standards available to avoid security issues related to AAC. The following are some of the important methods to mitigate AAC security challenges:

- Apply single-sign-on policy where ever possible.
- Multi-factor authentication can be employed which enables both identity and access management and it is used Amazon Web Services (AWS) [9].
- Biometric authentication has the potential to be the most secure form of single-sign-on authentication [27].
- RSA cryptosystem can be adopted which can accept different authentication models like two-factor authentication, knowledge-based authentication and adaptive authentication. The above methods are very effective for data protection in cloud computing [28].
- Intrusion Detection System (IDS), firewalls as well as segregation of obligations can be implemented on the different network and cloud layers to enable proper access control in cloud computing for better data protection [28].
- There are many third-party identity management solutions in the market. Employ some third-party solutions like Microsoft Azure Active Directory, Okta identity management, McAfee cloud identity manager, etc. Recently, Identity-Management-as-a-Service (IDaaS) solutions are getting more popular in the corporate infrastructure [29].
- Cloud applications should use open standards where applicable, such as Security Assertion Markup Language (SAML), an XML-based OASIS (Organization for the Advancement of Structured Information Standards) open standard for exchanging authentication and authorization data between security domains and Open Authorization (OAuth), an open standard for authorization, allowing users to share their private resources using tokens rather than credentials [6].

3.4. Security Challenges Due to Broken Authentication, Session and Access Controls

Broken authentication and session control threat occurs due to incorrect implementation of authentication and session management in the application domain. Examples of this type of threats are user credentials can be guessed due to weak account management functions, user credentials are not properly protected, session IDs are exposed in

the URL, etc. Attacker takes advantage of the situation and can compromise passwords, keys, session tokens or to exploit other implementation flaws to assume another user's identity [30]. Attackers generally target the privileged accounts. The Broken access control threat occurs when there is a lack of enforcement of restriction on what authenticated users are allowed to do. Using this loophole, attackers can access another user's account, view sensitive files, modify another user's data, change access rights etc. [30]. The following are some of the methods that can be used to reduce these challenges:

- Implement a single set of strong authentication and session management controls.
- Avoid Cross-site Scripting (XSS) flaws which can be used to steal session IDs.
- Check Access – Each use of a direct reference from an untrusted source must be checked for access control to ensure that the user is authorized for the requested resource.
- Use per user or session indirect object references – In this, the coding pattern prevents attackers from directly targeting unauthorized resource.
- Automated verification – Apply automation to verify proper authentication deployment.

3.5. Other Data Related Security Issues

Other minor data related security issues can occur through Data location, Multi-tenancy and Backup in cloud computing. In cloud computing, data is stored in a diverse geographic location with different legal jurisdictions [6]. If the data location is not safe physically and logically then there is always a threat to the CSC's data. In that scenario, data is vulnerable to malicious insiders and external hackers. Due to the multi-tenant nature of the cloud computing, multiple users can store their data in the same location using physical or virtual storage concept. Because of this scenario, a user can intrude into another user's data location [10].

In cloud computing, all data, especially the sensitive data should be regularly backed up and tested for proper data recovery in case of disasters. A Backup is a form of data-at-rest. Strong encryption schemes can be used to protect backup data, especially if the data is sensitive [4]. There are two types of backup, on-site backup and cloud-based backup. Based on the cost, type of business and data, backup plan can be selected. On-site backup is cheaper, easier to set up and runs faster. The backup and the production environments are the same and if any natural disasters, then all the data including the backup are lost. In the cloud-based backup, CSC's data is stored off-site and if any natural disasters on the CSC's site, then the data is still available with the cloud. The following are some of the methods to overcome other data related security issues:

- CSC should know the logical and the physical location of the data if not at least which state, country and data center due to all the potential regulatory, contractual and other jurisdictional issues [6].
- Establish location and jurisdictional policies to govern the data location [6].
- Adopt intelligent data segregation techniques to segregate the data from different users.
- Use strong encryption techniques for the backup data to avoid data leakage.

4. Conclusion and Future Challenges

The cloud problems are mainly with the security and privacy of the data stored in the cloud. The cloud environments like heterogeneity, resource sharing, multi-tenancy, virtualization, mobile cloud computing and Service Level Agreement (SLA) that makes the cloud security more vulnerable. This paper provides the data security issues and methods to overcome these issues. There are also new developments in cloud computing like Container-as-a-Service (CaaS), Software-defined networking (a concept to design and manage networks that abstracts applications away from the underlying networks), Software-defined-storage (abstracts the logical storage services and capabilities away from the underlying hardware) and Cloud-of-Things (CoT), (a concept combining cloud computing and Internet-of-Things (IoT) for smart city applications). All these new developments bring new challenges in cloud computing and they need to be addressed. When there is a change in technology, always review the security policies and procedures and update accordingly to protect the data and its privacy.

References

1. M. Hogan and A. Sokol. NIST Cloud Computing Standards Roadmap Version 2. NIST Cloud Computing Standards Roadmap Working Group, NIST Special Publications 500-291, NIST, Gaithersburg, MD, 2013, p.1-113.
2. G. Brunette and R. Mogul. Security Guidance for Critical Area of Focus in Cloud Computing V2.1, *Cloud Security Alliance* (CSA), 2009, p.1-76.
3. Z. Xiao and Y. Xiao. Security and Privacy in Cloud Computing, IEEE Communications Surveys & Tutorial, Vol. 15, 2012, No 2, p.843-859.
4. S. Aldossary and W. Allen. Data Security, Privacy, Availability and Integrity in Cloud Computing: Issues and Current Solutions. International Journal of Advanced Computer Science and Applications, Vol. 7, 2016, No. 4.
5. D. Catteddu and G. Hogben. Cloud Computing: Benefits, risks and recommendations for information security, *European Union Agency for Network and Information Security* (ENISA), 2009, p.1-125.
6. A. Reed, C. Rezek, C and P. Simmonds. Security Guidance for Critical Area of Focus in Cloud Computing V3.0, *Cloud Security Alliance* (CSA), 2011, p.1-177.
7. S. Ludwig. Cloud 101: What the heck do IaaS, PaaS and SaaS companies do? VentureBeat, 2011, https://venturebeat.com/2011/11/14/cloud-iaas-paas-saas/ , Accessed on 20th August 2017.
8. M. Sookhak, H. Talebian, E. Ahmed, A. Gani, M. K. Khan. A Review on remote data auditing in single cloud server: Taxonomy and open issues. Journal of Network and Computer Applications, Vol. 43, 2014, p.121-141.
9. Amazon Web Services: Overview of Security Processes. 2016, https://d0.awsstatic.com/whitepapers/aws-security-whitepaper.pdf, Accessed on 20th August 2017.
10. A. Subashini and V. Kavitha. A survey on security issues in service delivery models of cloud computing. Journal of Network and Computer Applications, Elsevier, Vol. 34, Iss. 1, 2011, p.1-11.
11. R. Bhadauria, S. Sanyal. Survey on Security Issues in Cloud Computing and Associated Mitigation Techniques. International Journal of Computer Applications, Vol. 47, 2012, No. 18, p.47-66.
12. F. Sabahi. Secure Virtualization for Cloud Environment Using Hypervisor-based Technology. International Journal of Machine Learning and Computing, Vol. 2, No. 1, 2012, p.39-45.
13. P. R. Gallagher. A Guide to Understanding Data Remanence in Automated Information Systems. The Rainbow Books, Chapter 3 and 4, 1991.
14. L. Pesante. Introduction to Information Security. https://www.us-cert.gov/sites/default/files/publications/infosecuritybasics.pdf, Accessed on 20th August 2017.
15. E. Worlanyo. A survey of cloud computing security: Issues, challenges and solutions. http://www.cse.wustl.edu/~jain//cse570-15/ftp/cld_sec/index.html, Accessed on 20th August 2017.
16. M. Rouse. Data Integrity. http://searchdatacenter.techtarget.com/definition/integrity, Accessed on 21st August 2017.
17. Y. Sun, J. Zhang, Y. Xiong and G. Zhu. Data Security and Privacy in Cloud Computing. International Journal of Distributed Sensor Networks, Vol. 10, Iss. 7, 2014, p.1-9.
18. C. Wang, S. Chow, Q. Wang, K. Ren and W. Lou. Privacy-Preserving Public Auditing for Secure Cloud Storage. IEEE Transactions on Computers, Vol. 62, Iss. 2, 2013, p.362-375.
19. B. Balusamy, P. Venkatakrishna, A. Vaidhyanathan, M. Ravikumar and N. Devi Munisamy. Enhanced Security Framework for Data Integrity using Third-Party Auditing in the Cloud System. In Proceedings of the Artificial Intelligence and Evolutionary Algorithms in Engineering Systems. Advances in Intelligent Systems and Computing, Springer, Vol. 325, 2015, p.25-31.
20. G. Atieniese, R. Burns, R. Curtmola, J. Herring, L. Kissner, Z. Peterson and D. Song. Provable Data Possession at Untrusted Stores. In Proceedings of the 14th ACM conference on Computer and Communication Security, 2007, p.598-609.
21. G. Atieniese, R. Di Puetro, L. V. Mancini and G. Tsudik. Scalable and Efficient Provable Data Possession. In Proceedings of the 4th International Conference on Security and Privacy in Communication Networks, 2008, Art. 9.
22. G. Atieniese, R. Burns, R. Curtmola, J. Herring, O. Khan, L. Kissner, Z. Peterson and D. Song. Remote Data Checking using Provable Data Possession. ACM Transaction of Information and System Security. Vol. 14, no. 1, 2011, p.12-34.
23. Q. Wang, C. Wang, K. Ren, W. Lou and J. Li. Enable Public Auditability and Data Dynamics for Storage Security in Cloud Computing. IEEE Transactions of Parallel and Distributed Systems. Vol. 22, no. 5, 2011, p.847-859.
24. M. Sookhak, A. Gani, M. K. Khan and R. Buyya. Dynamic Remote Data Auditing for Securing Big Data Storage in Cloud Computing. Information Sciences: An International Journal, Vol. 380, Iss. C, 2017, p.101-116.
25. CSCC Security for Cloud Computing Ten Steps to Ensure Success. Cloud Standards Customer Council, 2015, p.1-35.
26. M. Rouse. Authentication, http://searchsecurity.techtarget.com/definition/authentication, Accessed on 21st August 2017.
27. D. L. Shinder. Authentication in the Cloud. http://resources.infosecinstitute.com/authentication-cloud/#gref, 2014, Accessed on 21st August 2017.
28. K. Jakimoski. Security Techniques for Data Protection in Cloud Computing. International Journal of Grid and Distributed Computing, Vol. 9, No. 1, 2016, p.49-56.
29. T. Ferrill. The Best Identity Management Solutions of 2017. https://www.pcmag.com/article2/0,2817,2491437,00.asp, Accessed on 21st August 2017.
30. OWASP Top 10 Application Security Risks - 2017, Open Web Application Security Project (OWASP). https://www.owasp.org/index.php/Category:OWASP_Top_Ten_2017_Project, Accessed on 21st August 2017.