How Data Security Could Be Achieved in The Process of Cloud Data Governance?

An Analysis of Data Security in Cloud Computing From Management and Technique Perspectives

Weijia Liu

Business Information Systems, Australian National University
Australian National University's email: u6810699@alumni.anu.edu.au

Abstract
Cloud computing has been widely deployed and consumed across different application domains. It has been defined as a shared pool of computing resources. In cloud computing, data governance plays a critical role in enhancing overall performance and ensuring data security. This paper analyzes how to achieve data security in the scope of data governance regarding the organization structure and existing techniques by reviewing related research findings. The second part discusses security's role in an organization's data governance framework. The third part analyses the specific security risk and the expected risk mitigation methods reviewed.

Keywords - *Big data, cloud computing, data governance, data security, management model, encryption, data security application*

1. INTRODUCTION

Data governance can be defined as a management framework in company documents, used to allocate rights and obligations regarding decision-making so that data can be processed as company assets [1]. The governance is connected to the optimal utilization of assets and then treats the IT and data as assets in organizations used to cultivate internal data governance development [2]. Data governance procedures connect to the whole data lifecycle, including data organization, storage, backup and recovery, management and maintenance, and finally, data retention and secure destruction. It aims to guide records management and corporate content management. Meanwhile, it provides a framework that integrates IT infrastructure into the organization in the most proper method. So, organizations are expected to have an overall data governance strategy to optimize the value of data information and minimize potential risks [3].

Organizations are increasingly emphasizing their governance in the face of the increased volume and complexity of data and the additional demands for combining, manipulating, storing, and presenting information. However, data security has become an important issue in data governance devices. AI-Ruthe et.al pointed that security is considered to be the biggest challenge in data management. The most critical issue in adopting the cloud computing model is that a data management framework cannot be deployed in the IT infrastructure and service configuration. It requires new design and implementation to satisfy the requirements [4]. This paper discusses how to achieve data security in cloud data governance from management and technical application.

2. DATA SECURITY IN ORGANIZATION MANAGEMENT

2.1 Data security in organization framework

The data governance framework has been investigated regarding its design and implementation. AI-Ruthe et.al has proposed a conceptual framework that the construction of data governance within an organization mainly includes five steps: data governance structure, data governance assets, data governance functions, negotiation, and data governance service-level agreement development. Although data security issues are mainly dealt with in the third step, this is a continuous process based on the execution strategy and the supplier selected in the first step [5]. Regarding asset determination in the process, Haider has analyzed the asset lifecycle data in data governance. Connected his

© The Author(s) 2023
S. A. R. Khan et al. (Eds.): ICMSSE 2022, AHCS 12, pp. 114–120, 2023.
https://doi.org/10.2991/978-94-6463-056-5_16

category to cloud computing, it includes metadata management, document management, data quality management, data structure management, data security management, and database operation management [6]. Considering cloud data governance should interact with the third party, the accountability mechanism has been proposed to strengthen the overall security level. It could be defined as a model which could provide interactive sharing resources to managers and suppliers, such as network, storage, applications, and service [7]. The accountability mechanism connects three parties, the organization, the cloud provider, and the certificate association (CA). The organization could authorize CA to take supervisory behavior to cloud suppliers (as figure 1 shows).

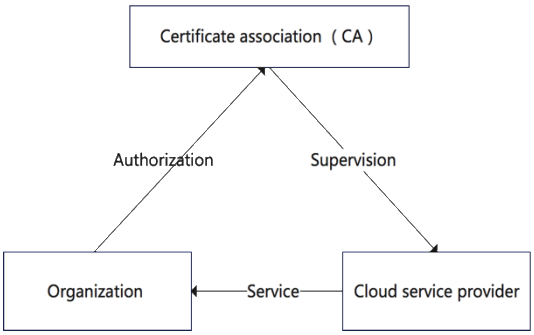

Figure 1: Accountability governance

The reviewing of the data governance in the designing phase presents the conceptual framework in figure 2.

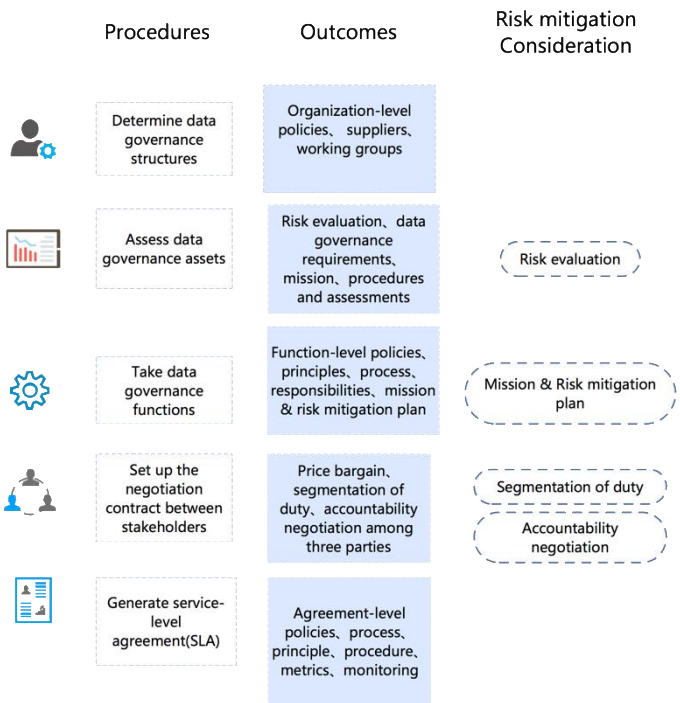

Figure 2: Conceptual framework of data governance design

2.2 the management architecture in data governance

After analyzing the overall data governance framework, this part will study the security issues in the management structure. According to AI-Ruthe et al., It mainly includes three corresponding groups, which can be divided into three-level management models: a group of senior managers, a middle-level management team, and a data governance working group [5]:

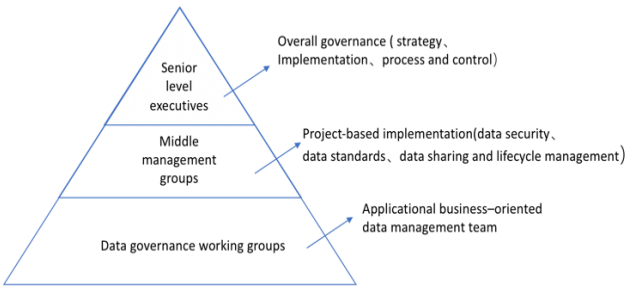

Figure 3: 3-tier management model in data governance

According to Khatri and Brown, Data and IT have been thought of as IT assets in organizations, the same as human assets, financial assets, physical assets, IP assets, and related assets. Security will be controlled in data access, which should be managed by the Chief Information Security and Data Security Officer [8]. Weber et al. also put forward two connotations corresponding to data security. It refers to modern crystal algorithms to protect data security to ensure data integrity and confidentiality. In addition, it is related to the protection of data in modern data storage methods [9]. Felici et al. investigate accountability mechanisms in cloud computing, and the data security issues in organizations and be in charge of the data security program [10]. Wende's research emphasized on data quality management (DQM) approach in data governance. In order to address the organizational problem in terms of IT and management, DQM should establish organization-wide guidelines and standards that are consistent with corporate strategies and data governing laws [11]. In that way, data security is comparatively specific from a management perspective, so it may not reach the strategic level. The overall data security governance model has been concluded below:

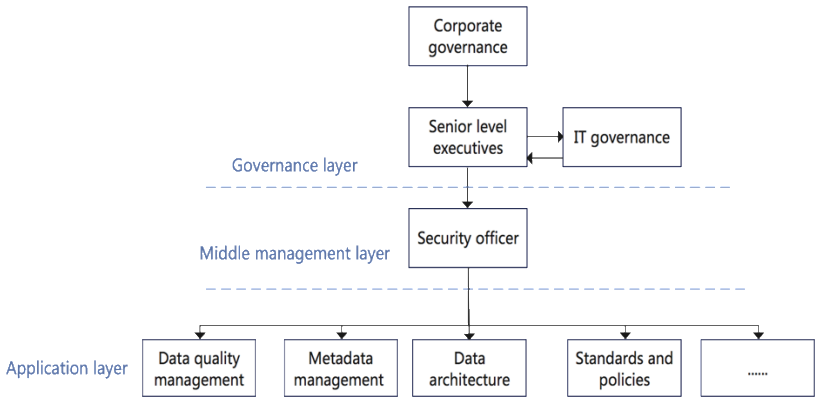

Figure 4: Data security management framework

From the security management framework, it can be seen that the data security in the organization would be accomplished based on the 3-tired architecture. The IT governance will be integrated with the whole process to realize the overall contingency with the business strategies.

3. CLOUD DATA SECURITY ISSUE IN TECHNICAL FACET

3.1 Cloud computing security framework

Specific to technical facets of cloud computing, data security questions could be categorized into several dimensions. Garner has pointed out that cloud users should inquire about seven specific security problems, including privileged user access, regulatory compliance, data location, data segregation, recovery, investigative support, and long-term viability[12]. Cloud data security and cloud computing infrastructure-related policies, controls, and technologies are different from the traditional IT environment [13]. Within this scope, data security and privacy should be well deployed throughout the data life cycle.

Security questions could be categorized into several dimensions. Garner has pointed out that cloud users should inquire about seven specific security problems, including privileged user access, regulatory compliance, data location, data segregation, recovery, investigative support, and long-term viability [12]. Cloud data security and cloud computing infrastructure-related policies, controls, and technologies are different from the traditional IT environment [13]. Within this scope, data security should be well deployed throughout the data life cycle.

Regarding data security risks, Sangroy et.al analyzed the principal risks of data security management. They pointed out seven significant security challenges in the cloud environment, including data location, investigation, data isolation, long-term survivability, compromised servers, regulatory compliance, and recovery [14]. Sood proposed a framework that could offer complete security of data governance in cloud computing. A series of challenges were also concluded, which refer to unauthorized servers, brute force attacks, a threat from a cloud service provider, tampering with data, and loss of user identity or password [15]. Connected to the current finding of the security issue to the cyber risk assessment methods, the overall data security threats can be assessed in figure 5. It could be seen that security mitigation methods should be integrated into every procedure in the data life cycle to realize the confidentiality, integrity, and availability of data.

Data Security in data life cycle				
Phase	**Security focus (Potential threat)**	**Security Requirment**	**Owners**	**Application/Mechanism**
Generation	The overall security plan may not complete enough or cannot be implemented.	Confidentiality	Senior level executives	Need analysis、Risk assessment
		Integrity	Middle management group	
		Availability	Data governance working group	
Transfer	Confidentiality should be ensured in the whole data transfer process, not only between enterprise storage and cloud storage, but between different cloud storage services (cloud providers).	Confidentiality	Middle management group	Access control 、Construct internal data centers
		Integrity	Data governance groups	
Use	Encryption is critical in this procedure for simple storage service, while it could not be utilized in cloud-based applications in the PaaS or SaaS model.	Confidentiality	Middle management group	Encryption、Backup data from the cloud regularly、User authentication
		Integrity	Data governance groups	
		Availability	Cloud providers	
Share	When sharing with third parties, data owners should consider whether the third party could be authorized to use.	Confidentiality	Senior level Executives	User authentication、Segmentation of duty
		Integrity	Middle management group	
		Availability	Data governance working group	
Storage	In this process, the first problem is that professional skills are lacked to manage the keys. Secondly, the integrity is hard to promise. Except for the external attack, the availabilty of cloud computing, the cloud provider authority and the cloud backup service will all threaten the availability of data.	Confidentiality	Middle management group	Encryption、Backup data from the cloud regularly、User authentication、Data backup、Construct internal data centers
		Integrity	Data governance working group	
		Availability		
Archival	Data may take the risk of breach if it is stored in PMP. Also, the storage duration should be consistent with archival requirements.	Confidentiality	Middle management group	User authentication、Segmentation of duty、Negotiation between cloud providers and internal managers
		Availability	Data governance working group	
Destruction	When data is not required, it should be destroyed, which may exist and can be restored. This may lead to the breach of sensitive information.	Confidentiality	Middle management group	User authentication、Security training
			Data governance working group	

Figure 5: Data security threat in the data life cycle

Based on the assessment, there are four technical methods related to the data security risk that should be focused on to promise the confidentiality, integrity, and availability of data: user authentication, data encryption, three-party negotiation, and data backup.

3.2 Data risk mitigation mechanism

3.2.1 Data encryption

Encryption has been thought to be one of the most effective methods to prevent intrusion , which is a means of distinguishing information using algorithms [16]. It is a way to secure the data in an untrusted cloud server [17]. Encryption is the process of disguising information and transferring them into ciphertext, while decryption is restoring them into readable form [16]. In cloud computing, despite symmetrical encryption and asymmetric encryption, attribute-based encryption (ABE) and identity-based encryption (IBE) are efficient systems with fine-grained access control. With the data generated in multiple organizations, access policies could be conducted by multiple authorities [18].

a) The three encryption methods

In symmetrical encryption, the key used to encrypt the message is consistent with the one used to decrypt the message. With asymmetric encryption, the key utilized to encrypt the message varies from the one used to decrypt it [16]. Attribute-based encryption (ABE) and identity-based encryption (IBE) could be realized by authorized management, which means that all the private keys should be managed in the authorization center. To avoid a centralized attack, Hierarchical IBE and Hierarchical ABE be managed in different levels so that the key can be allocated in different levels [18].

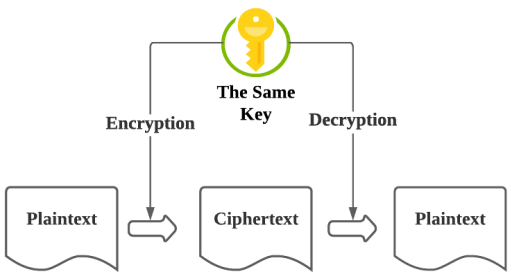

Figure 6: Symmetric encryption

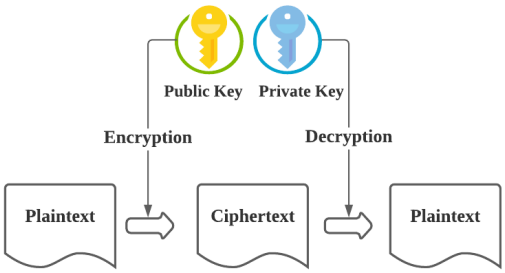

Figure 7: Asymmetric encryption

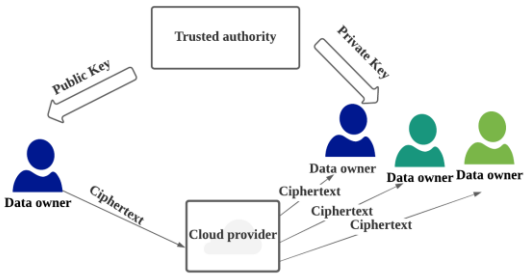

Figure 8: Attribute-based encryption

b) Application of the three encryption methods

In cloud computing, these three encryption methods are used in combination to protect data security. In order to ensure secure cloud storage for personal or enterprise use, data can first be encrypted with AES using a unique key, and then attribute-based encryption can encrypt the unique key. Khamenei and Hanapi proposed a data-sharing method using RSA and AES encryption methods. In this framework, the sender-receiver, and the cloud storage system (CSS). In the first process, the sender transfers his document to the system to CSS. The RSA algorithm will be used to implement encryption. After that, the document should be delivered from the CSS system to the recipient. After the system receives the request, the recipient's public key will also be sent to the CSS using the user's public key. Finally, the AES encryption algorithm will find the required file and send it to the user with the key [17].

In addition, many other encryption applications have been introduced as well. Zarandioon proposed a user-oriented privacy protection protocol called K2C. It permits storing, sharing, and managing their information, which is untrusted anonymously [19].

3.2.2 User authentication

User authentication means that only authorized users can access specific resources inside the network [16]. Critical public infrastructure (PKI) and single sign-on (SSO) technologies are mainly used for cloud computing by reviewing existing authentication methods. Public critical infrastructure (PKI). PKI provides a framework that can effectively deploy mass scale. It can support identity management and online identity authentication within and across the network through Secure Socket Layer (SSL) and Transport Layer Security (TLS). It could support identity management within and across the network and online identity authentication with secure socket layer (SSL) and transport layer security (TLS). The single sign-on technique utilized a single authentication action to permit an authorized user to access independent but related software systems or applications. It decreases the risk of central management for managers and enhances users' working efficiency [20].

3.2.3 Data backup and disaster recovery techniques

ISP provides significant storage space to users in cloud computing, and users can update data to the central cloud. In this way, it creates risks. Once the data stored in the cloud disappears for specific reasons, such as cloud destruction or natural disasters, consumers' data remains in the cloud, and they should continue to rely on cloud providers. In order for addressing this problem, constructing a remote data backup service is an effective method. It is a server that stores the leading cloud's entire data and is located at a remote place, which intends to assist clients to access data from the remote server when network or connection failure comes across [21].

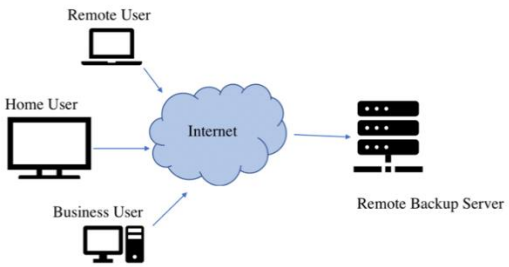

Figure 9: Remote data backup server

4. LIMITATION

One concern about the findings is that the data governance model and security architecture are discussed based on the majority of situations in organizations. It has not taken the business model innovation factors into consideration. With the business innovation dynamic, it can achieve more possibilities and functions in terms of data security.

5. CONCLUSION

This article reviews the literature on cloud data governance and finds out how to achieve data security in the process. First, the overall data governance framework is analyzed, and security risk control methods are deployed throughout the data governance process, from

design to protocol formulation. Then, the management structure of data security is discussed extensively within the organization, and the three-level management architecture is most suitable for data security management. The third part examines the role of the data security framework and assesses the risks in the data life cycle. Then, it explained four security mitigation methods, encryption, user authentication, data backup, and disaster recovery plan technologies to reduce security risks.

This study contributes to figuring out data security's role in the entire data governance process. It provides guidance for deploying data security administration in cloud computing architectures from the strategic level to the specific technique level. In addition, this study emphasizes the significance of data security in managing the model construction of organizations, so that security awareness can be penetrated in every procedure of data governance to avoid and mitigate security risks.

REFERENCES

[1] B. Otto, "Organizing Data Governance: Findings from the Telecommunications Industry and Consequences for Large Service Providers", Communications of the Association for Information Systems, vol. 29, P.47 2011.

[2] N. Horne, "Information as an asset—The board agenda", *Computer Audit Update*, vol. 1995, no. 9, pp. 5-11, 1995.

[3] D. Catteddu, "Cloud Computing: Benefits, Risks and Recommendations for Information Security", *Web Application Security*, pp. 17-17, 2010. Available: 10.1007/978-3-642-16120-9_9

[4] M. Al-Ruithe, E. Benkhelifa and K. Hameed, "A systematic literature review of data governance and cloud data governance", *Personal and Ubiquitous Computing*, vol. 23, no. 5-6, pp. 839-859, 2018. Available: 10.1007/s00779-017-1104-3

[5] M. Al-Ruithe, E. Benkhelifa and K. Hameed, "A Conceptual Framework for Designing Data Governance for Cloud Computing", *Procedia Computer Science*, vol. 94, pp. 160-167, 2016. Available: 10.1016/j.procs.2016.08.025.

[6] A. Haider, "Asset Lifecycle Data Governance Framework", *Lecture Notes in Mechanical Engineering*, pp. 287-296, 2014. Available: 10.1007/978-3-319-06966-1_27

[7] G. Cheng, Y. Liu, Z. Gao and X. Liu, "Cloud Data Governance Maturity Model", *Conference On Software Engineering and Service Science (ICSESS)*, pp. 517-520, 2017.

[8] V. Khatri and C. Brown, "Designing data governance", *Communications of the ACM*, vol. 53, no. 1, pp. 148-152, 2010. Available: 10.1145/1629175.1629210

[9] K. Weber, B. Otto and H. Österle, "One Size Does Not Fit All---A Contingency Approach to Data Governance", *Journal of Data and Information Quality*, vol. 1, no. 1, pp. 1-27, 2009. Available: 10.1145/1515693.1515696.

[10] M. Felici, T. Koulouris and S. Pearson, "Accountability for Data Governance in Cloud Ecosystems", *2013 IEEE 5th International Conference on Cloud Computing Technology and Science*, 2013. Available: 10.1109/cloudcom.2013.157

[11] K. Wende, "A Model for Data Governance - Organising Accountabilities for Data Quality Management", *18th Australasian Conference on Information Systems*, pp. 417-425, 2007.

[12] J. Brodkin, "Gartner: Seven cloud-computing security risks," Jul. 2008.

[13] D. Chen and H. Zhao, "Data Security and Privacy Protection Issues in Cloud Computing", *2012 International Conference on Computer Science and Electronics Engineering*, 2012. Available: 10.1109/iccsee.2012.193.

[14] A. Sangroya, S. Kumar, J. Dhok, and V. Varma, "Towards Analyzing Data Security Risks in Cloud Computing Environments," in *Information Systems, Technology and Management*, Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 255–265. Accessed: Oct. 27, 2022. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-12035-0_25

[15] S. Sood, "A combined approach to ensure data security in cloud computing", *Journal of Network and Computer Applications*, vol. 35, no. 6, pp. 1831-1838, 2012. Available: 10.1016/j.jnca.2012.07.007.

[16] J. Fitzgerald, A. Dennis and A. Durcikova, *Business Data Communications and Networking, Thirteeth Edition*, 13th ed. 2017, pp. 315-330.

[17] N. Khanezaei and Z. Hanapi, "A framework based on RSA and AES encryption algorithms for cloud computing services", *2014 IEEE Conference on Systems, Process and Control (ICSPC 2014)*, 2014. Available: 10.1109/spc.2014.7086230.

[18] Y. Yang, X. Chen, H. Chen and X. Du, "Improving Privacy and Security in Decentralizing Multi-Authority Attribute-Based Encryption in Cloud Computing", *IEEE Access*, vol. 6, pp. 18009-18021, 2018. Available: 10.1109/access.2018.2820182.

[19] S. Zarandioon, D.D. Yao, and V. Ganapathy, "K2C: Cryptographic cloud storage with lazy revocation and anonymous access", Security and Privacy in Communication Networks, Springer Berlin Heidelberg, pp. 59-76, 2012

[20] V. Radha and D. Reddy, "A Survey on Single Sign-On Techniques", *Procedia Technology*, vol. 4, pp. 134-139, 2012. Available: 10.1016/j.protcy.2012.05.019.

[21] K. Sharma and K. Singh, "Online Data Back-up and Disaster Recovery Techniques in Cloud Computing: A Review", *International Journal of Engineering and Innovative Technology (IJEIT)*, vol. 2, no. 5, pp. 249-254, 2012.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (http://creativecommons.org/licenses/by-nc/4.0/), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.