ISSN: 2455 - 1597

ijergs
Open Access

International Journal of Engineering Research and Generic Science (IJERGS)
Available online at: https://www.ijergs.in

Volume - 7, Issue - 3, May - June - 2021, Page No. 49 - 58

Data Security and Privacy in Cloud Computing

[1]Aashish Gupta, Student, Department of Computer Science Engineering, Arya Institute of Engineering Technology and Management, Jaipur, Rajasthan (India)

[2]Ritika Jain, Student, Department of Computer Science Engineering, Arya Institute of Engineering Technology and Management, Jaipur, Rajasthan (India)

[3]Chirayu Mehta, Student, Department of Computer Science Engineering, Arya Institute of Engineering Technology and Management, Jaipur, Rajasthan (India)

[4]Sanjay Tiwari, Associate Professor, Department of Computer Science Engineering, Arya Institute of Engineering Technology and Management, Jaipur, Rajasthan (India)

Abstract

Cloud computing is a method of delivering software, storage, and processing through the internet. It improves the system's capability without modifying the existing infrastructure, training new employees, or purchasing licenses. It enhances and expands existing software capabilities and Information Security resources. Cloud computing has exploded in popularity over the years, boosting the IT industry's business idea. Considering all of the advances in cloud computing, security remains a major concern in the cloud computing environment. These issues also include loss, spillage, and disclosure of confidential information (such as personal and economic information). We reviewed the literature and analyzed several cloud computing models, which revealed that cloud privacy/protection is still in its infancy. Cloud computing is now a global idea that is used by a vastly large number of internet users. Cloud computing is now a global idea that is used by a vastly large number of internet users. Because of the simple and appealing characteristics, it contains, the number of institutions, corporations, and other personal users depending on cloud services and keeping crucial information in the cloud has expanded dramatically over the years. Despite its popularity, there are a number of concerns about the security of data kept in the cloud. Cloud customers are increasingly concerned about the security of their data when it is transported into the cloud. Hackers with advanced skills and motivation are increasingly attempting to intercept or steal large amounts of data, including sensitive information that is being transported or stored on the cloud. Various academics have offered a range of approaches to ensure data security during transmission based on hacker goals.

Keywords: Cloud, Services, information, Global.

Introduction

Cloud computing can be defined as a model for providing on-demand network access to a pool of configurable computing resources such as networks, servers, storage, applications, and services that can be rapidly provisioned and released with minimal management effort or service provider interaction. Cloud computing provides a variety of service and deployment approaches. Infrastructure as a service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS) is the service models, while Public, Private, Hybrid, and Community Cloud are the deployment options. The cloud's primary notion is built on the services it provides, which range from service platform provisioning to grid and service

computing to Software as a Service. Regardless of the architecture, the core premise of this computing model is that users' data, which might come from individuals, organizations, or businesses, are handled remotely by unidentified computers that the user is unaware of. However, the convenience and effectiveness of this method come at the expense of privacy and security. The primary stumbling block in implementing cloud services is data confidentiality. Cloud computing has developed massive data centers, yet data and service deployment remain unreliable. As a result, there is a slew of new security issues to contend with. Vulnerabilities in accessibility, virtualization, and web??? such as SQL injection, cross-site scripting, physical access issues, privacy and control issues arising from third parties with physical control of data, identity and credential issues, data confirmation, changing, and privacy, data loss and theft, issues related to integrity and IP spoofing.

Literature Survey

Various researchers have proposed several alternative security mechanisms. We will present a literature review of previous work on this topic in this part. Jan de Muijnck-Hughes presented a security approach called Predicate Based Encryption in 2011. (PBE). PBE is a type of asymmetric encryption that has its roots in Identity-Based Encryption [1]. This method combines Attribute-Based Access Control (ABAC) and asymmetric encryption, allowing for the creation of a single encryptor/multi decryptor environment with a consistent framework. This Predicate Based Encryption focuses on both Platform as a Service and Software as a Service implementation. This suggested solution also protects cloud resident data from undesired disclosure, leaking, and other breaches of confidentiality.

Venkata Sravan and colleagues published a paper titled Security Techniques for Protecting Data in the Cloud in 2011. The goal of this work is to analyze security concerns in Cloud computing and to develop effective security strategies for mitigating them [2]. A total of 43 security issues and 43 security approaches were discovered in the study. Confidentiality (31%) is the most often assessed trait, followed by Integrity (24%), and Availability (19%) [2].

In 2011, Ali Asghary Karahroudy published a study titled Security Analysis and Framework of Cloud Computing with Partially Distributed File System Based on Parity. This study presented a mechanism known as the Partially Distributed File System with Parity (PDFSP), which is a protocol based on the existing GFS/HDFS [3]. Client Access Machine, User Public Machine, Cloud Management Server, and File Retrieval Server are the four key components of this PDFSP. All of these components work together to guarantee that the data being transferred is not intercepted. Confidentiality, Integrity, and Availability were the three components of security discussed in this study.

In 2013, Nabil Giweli presented the Data-Centric Security methodology, which is a solution-based approach. This technique intends to provide data security by allowing data to self-describe, defend, and protect itself throughout its lifespan in cloud settings. This approach places the full burden of setting and managing data privacy and security safeguards on the data owner. This suggested solution uses symmetric and asymmetric encryption methods and is based on the Chinese Remainder Theorem (CRT). The suggested technique is shown to be highly efficient in this study since it does not require sophisticated key generation methods and does not require the data file to be encrypted more than once [4].

Miao Zhou suggested five strategies for ensuring data security and integrity in cloud computing in 2013. Revolutionary tree-based access control scheme, Privacy enhanced data outsourcing in the cloud, Privacy maintained access control for cloud computing, Privacy improved keyword search in clouds, and Public online integrity check for private data are just a few of the ways available. This work used a Keyword Searching Mechanism, which allows for efficient multi-user keyword searches while concealing personal information in search queries [5]. To enable flexible and fine-grained access control in the cloud, an encryption strategy for a two-tier system was provided. The suggested approach is efficient, according to the experimental results, especially when the data file is big and the integrity check is performed often.

Sudhansu Ranjan Lenka and colleagues published a paper titled "Enhancing Data Security in Cloud Computing Using RSA Encryption and MD5 Algorithm" in 2014. They implemented both the RSA and MD5 algorithms, as the title of the study indicates. The RSA Algorithm is employed in this article for secure communication and file encryption and decryption, while the MD5 Algorithm is employed for digital signature and table protection from unauthorized users [6]. Confidentiality, Integrity, and Availability are the three (3) characteristics of security provided by the two algorithms described.

In 2014, Aastha Mishra suggested a Key Management Scheme for Advanced Secret Sharing. The goal of this study is to provide a more reliable decentralized lightweight key management methodology for cloud systems that will improve data security and key management [7]. The suggested strategy preserves the security and privacy of user data by replicating key shares across many clouds utilizing a secret sharing strategy and a voting method to verify share integrity. The approach described in this research also provides improved security against Byzantine failure, server collusion, and data alteration attacks [7].

Cloud Data Storage Security based on Cryptographic Mechanisms was written by Nesrine Kaaniche in 2014. ID-Based Cryptography (IBC) and CloudSec are two (2) strategies presented by Nesrine in this study to protect data. The study proposes using ID-Based Cryptography to employ each client as a private key generator, generating his own ID-Based Cryptographic Public Elements (IBC-PE). These IBC-PE are used to generate ID-based keys and encrypt data before it is stored and shared in the cloud [8]. There is a public key-based solution for CloudaSec that promotes the decoupling of subscriber access control and nondisclosure asymmetric encryption rules [8]. CloudaSec enables scalable and flexible implementation of the system, as well as solid security assurances for cloud services stored on cloud servers [8]. This study examines and explains why cryptographic activities on the client side are acceptable as compared to upload operations and do not need extensive processing resources. For example, encoding a data set of 8*105 bytes takes only 0.1 seconds, however uploading it takes 10 seconds [8]. As a result, the encryption methods consume 1% of the Openstack upload overhead.

In his work Data Confidentiality and Risk Management in Cloud Computing, Afnan Ullah Khan introduced a system known as Access Control and Data Confidentiality (ACDC). The paper's goal was to create a new methodology for enforcing authentication regulations in cloud computing settings [9]. He utilized a medical/healthcare situation to come up with the following composition: Data Owner (Medical Center), Data Consumers (Patients, Nurses, Doctors, etc.),

Infrastructure Provider, and Trusted Authority. The article uses Infrastructure as a Service as its delivery paradigm, and the presented methodology was utilized to ensure data secrecy and authentication.

Dimitra A. Georgiou published a paper in 2017 outlining security standards for cloud services. The goal of security policies is to safeguard people and information, establish guidelines for anticipated user behavior, reduce risks, and track regulatory compliance. [11]. The focus of the paper was on Software as a Service. The report provided a comprehensive assessment and analysis of previous studies on cloud computing environments. Dimitra concentrated his evaluation of known threats on the ones that aren't relevant to traditional systems [11]. An approach for analysing distinct dangers in the cloud was developed to be able to suggest new rules that should be implemented into the cloud policy. This report examined the security needs of a cloud service provider using a case study of the European E-health system.

Challenges Observed In Literature Survey

The following are some of the obstacles or concerns that were discovered when reading and reviewing the study papers:

- A few of the published studies concentrated on Platform as a Service (PaaS) and Software as a Service (SaaS), leaving Infrastructure as a Service (IaaS) out.
- Other publications related to data confidential information without considering integrity, non-repudiation, or authenticity.
- Only a few of the publications were theoretical, implying that no practical implementation was carried out.
- In other articles, the proposed methodology appears to be dependable, but it appears to be strange, hard, and time-consuming to apply.
- Some proposed approaches, such as Access Control and Data Confidentiality, were also not empirically proven (ACDC).

Data Integrity

One of the most important aspects of any information technology is data integrity. Data integrity, in general, refers to the protection of data against unlawful deletion, alteration, or fabrication. Managing an entity's access and rights to certain corporate resources helps to guarantee that sensitive data and services are not misused, misappropriated, or stolen. In a solitary system with a single database, data integrity is simple to achieve.

In a standalone system, data integrity is ensured by database constraints and transactions, which are normally completed by a database management system (DBMS). To maintain data integrity, transactions should adhere to the ACID (atomicity, consistency, isolation, and durability) criteria. The majority of databases can handle ACID transactions and maintain data integrity.

Authorization is a method of restricting data access. It's the method through which a system determines what level of access a certain authorized user should have to the system's secure resources. In a cloud system, data integrity refers to the preservation of information integrity. Unauthorized users should not be able to access or modify the data. The foundation for providing cloud computing services such as SaaS, PaaS, and IaaS is data integrity. Aside from storing vast

amounts of data, cloud computing environments frequently offer data processing services. Techniques such as RAID-like schemes and digital signatures can be used to ensure data integrity.

Data Confidentiality

When storing private or secret data on the cloud, data confidentiality is critical. To maintain data confidentiality, authentication and access control mechanisms are employed. Increased cloud reliability and trustworthiness might resolve data confidentiality, authentication, and access control challenges in cloud computing [16]. Users do not trust cloud providers, and it is nearly hard for cloud storage service providers to eliminate potential insider threats, therefore storing sensitive data in cloud storage directly is extremely unsafe. Simple encryption has a key management difficulty, and it can't handle complicated requests like queries, simultaneous modification, and fine-grained authorization.

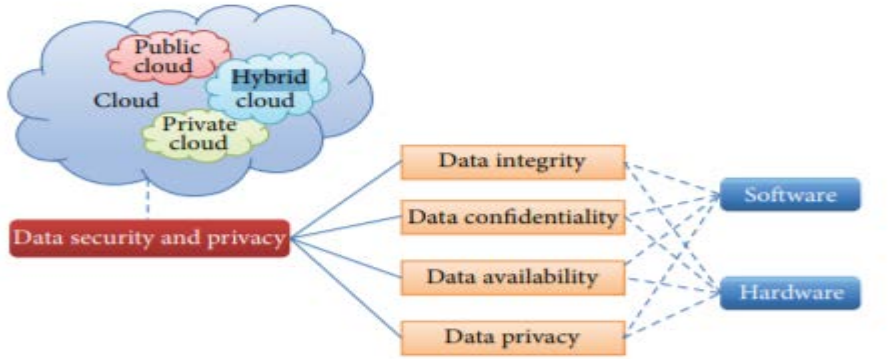

FIGURE 1: Organization of data security and privacy in cloud computing.

Data Availability:

When events like hard disc damage, IDC fires, and network failures occur, availability of data refers to the extent to which users' data can be utilised or restored, as well as how users validate their data using procedures rather than relying solely on the cloud service provider's credit guarantee. Clients are concerned about keeping data on cross-border servers since cloud suppliers are bound by local regulations, and cloud clients should be aware of those rules as well. Furthermore, the cloud service provider must assure data security, including data confidentiality and integrity. All such issues should be shared with the customer, and a relationship of trust should be established in this connection. Clients should be informed about the cloud vendor's data security promises and the jurisdiction of local laws. The paper's major focus is on data difficulties and problems related to data storage site and migration, as well as cost, accessibility, and privacy. Users may boost their confidence in the cloud by locating data. Cloud storage offers consumers a transparent storage solution, which reduces the complexity of the internet but also limits customers' control over their data storage. Benson et al. investigated geographic replication proofs and were successful in identifying information held in the Amazon cloud.

Data Privacy

A person or a group's capacity to seclude themselves or information about themselves and selectively expose them is known as privacy [18]. The following are the components of privacy:

(i) When: a respondent may be more interested in present or future information than information from the past being exposed.

(ii) How: A user may feel at ease if his or her friends may manually request his or her information, but he or she may not like automated and frequent warnings.

(iii) Extent: Rather than a particular spot, a user's information may be given as an unclear zone. Consumer context and privacy must be safeguarded and utilized correctly in trade.

Privacy in the workplace comprises the use of laws, systems, standards, and procedures to control personally identifiable information [19]. When users visit sensitive material in the cloud, privacy implies that cloud services can prevent a prospective adversary from inferring the user's conduct based on the user's visit model (not direct data leakage). Oblivious RAM (ORAM) technology has been the topic of research. ORAM technology visits many copies of data to conceal the true purpose of users' visits. As a promising technique, ORAM has been widely employed in software protection and in preserving privacy in the cloud.A route ORAM method was presented by Stefanov et al. as the state-of-the-art implementation [20]. The privacy concerns vary depending on the cloud environment, and they may be grouped into four groups as follows [44, [20], [21]].

(i) How to provide consumers access to their data while it's stored and processed in the cloud while preventing theft, malicious usage, and illegal sales,

(ii)How to ensure data replications in a jurisdiction and consistent state, where duplicating user data to numerous acceptable places is a common option, while avoiding data loss, leakage, and illegal alteration or fabrication

(iii) Which party is in charge of ensuring that legal obligations for personal data are met?

(iv) How many cloud subcontractors are engaged in processing, and how can they be appropriately recognized, reviewed, and verified?

Data Protection Model

Although public awareness of cloud privacy problems is growing, little progress has been made in this area. Pearson et al. recently presented accountability methods to meet end users' privacy concerns [13] and then developed a simple solution, a privacy manager, based on obfuscation methods [14]. Their main concept is that the user's data is stored in the cloud in encrypted form, and only the encrypted data is handled there.

The features of cloud services are discussed in this part, as well as a policy-driven architecture for data privacy protection. Strategy ranking, policy integration, and policy execution are the three main components of the data protection architecture.

Policy Ranking Models: In comparison to consumers' privacy checks, policy ranking is used to rate the supplier with the most often privacy policies (or policies). Two significant aspects to consider are privacy and efficiency needs. We presented three policy rankings projections based on the relative relevance of the two factors: I consumer rating model; (ii) service-provider-oriented ranking model; and (iii) agent-based ranking model.

User-oriented Ranking Model: Users in this paradigm are in charge of policy comparison and have processing and storage capacity. To begin, users must get privacy policies from service providers who provide the necessary services. The second approach is for users to upload their service requirements to the cloud, after which the relevant service providers will email the consumers their privacy policies. Users can begin ranking and selecting the best appropriate policy after receiving policies from suppliers.

Service-provider-oriented Ranking Model: The strategy comparisons task is used in this model. Users must demonstrate their service requirements as well as their privacy requirements. Service providers that want to attract more customers will use the same ranking software to compare privacy regulations and present the user with similarity ratings. After that, individuals may choose their favorite service providers. Because policy comparison takes place on the service provider's end, the service provider is responsible for policy privacy. The more policies that are similar, the less work is expected to be put into integrating them.

Agent-based Ranking Model: The very last two methods show either that users need some additional ability to avoid disclosing their privacy settings to all service providers, or service providers may need to bring out this problem and resolve it by enlisting the help of certificated third parties who can act as intermediaries among users and service providers. The agent compiles service provider policies and distributes them to customers along with a rating list based on their needs.

4.2 Policy Integration Models :

Following the consumers' successful selection of a service provider, the next stage is to achieve an agreement on both sides' data privacy issues. All privacy needs are fed into the policy integration model, which then helps to produce regulations that should be applied by upcoming superstars**.**

A Policy Integration Approach: The following are some characteristics to consider when designing a good policy integration system.

P1: Each service provider and subcontractor in the cloud has their own set of privacy policies. To resolve the discrepancy and achieve agreement on all requirements, the policy integration approach is used.

P2: The second approach to policy integration is to create or generate actual policies as an output automatically. In real-life situations, this property of the policy integration approach is critical. P3: The policy integration approach should be flexible in order to allow policy updates and reduce maintenance costs, and it should not necessitate re-executing policy changes every time.

Common point approach : Different parties are involved at different levels in cloud computing for a particular service. The issue here is determining which party policies should be merged such that privacy needs are met but overall performance is not harmed. The focus is on the cost and policies for this aim, without sacrificing the simple. These issues may be solved using a binary expression, in which the leaf nodes represent the policies to be integrated and the interior nodes represent the actions to be merged.

Shared approach: Adjacent parties with direct interaction can incorporate policies using this method. These integration costs should be split evenly among the contributors. In comparison to the common point strategy, achieving the common ideal answer is similarly challenging.

Policy implementation Model: After the policies have been written and effectively integrated, the following stage is to implement them. However, before implementing policies, it must meet the requirements and avoid the problems that may arise during implementation. It must meet the following criteria:

P1: Data and policy implementation must ensure data integrity, availability, and confidentiality. On a need-to-know basis, approved service providers should have access to integrity data. Only authorized data can access private information and policies.

P2: The policies can be tailored to the data owner, but they must be individually identifiable.

P3: Authorized parties should have their policies changed so that service providers can change their rules based on the needs of the users.

Tight coupling: Using a tight coupling strategy, the integrity of data and policies will always be guaranteed. Some of the methods for applying sticky policies to data [15] have been assumed, but no convincing explanation has been offered.

Loose coupling: This method is used to dynamically change rules based on data access. The policies are saved using this method.

Conclusion

Cloud computing refers to the supply of computing services such as servers, storage, databases, networking, software, analytics, intelligence, and the internet to enable quicker innovation, more flexible resources, and cost savings. Cloud computing is a new social phenomenon that is being used by people daily. Any key emergent technology has its own set of problems that make it difficult to embrace. Cloud computing is now viewed as a rapidly emerging sector that may provide instantaneous extendable service through the internet with the use of hardware and software virtualization. (12) Cloud Computing is a highly promising phenomenon that promises organizations to enhance productivity while lowering production costs. Although it has been implemented and utilized in a production setting, data protection and security in cloud computing are still crawling on their knees and need additional research. [11] Data security in Cloud Computing is an essential topic that has to be addressed. Today, a large quantity of data is stored on the cloud, making it easier for intruders and eavesdroppers to access it. As a result, it is critical to do extensive research into how to develop and implement solid and functional security mechanisms that would prevent hackers from gaining access to data being transported to and from the cloud. Based on the data given in this study and the knowledge gained from the execution of the recommended methodologies, it is clear that the majority of papers place a high priority on data confidentiality, while only a few articles meet the three (3) aspects of security; Confidentiality, Availability, and Integrity.

References

1. T. Mather, S. Kumaraswamy, and S. Litif, Cloud Security and Privacy: An enterprise perspective on Risks and Compliance (Theory in Practice). O' Reilly, 2009.
2. Venkata S. et.al (2011) *Security Techniques for Protecting Data in Cloud Computing*, 12 Jan, 2019 [Online] Available: https://www.bth.se/com
3. Ali Asghary K. (2011) *Security Analysis and Framework of Cloud Computing with Parity-Based Partially Distributed File System*, 26, Jan, 2019 [Online] Available; https://www.academia.edu/27767213/security_Analysis_and_Framework_of_cloud_computing_with_Parity_Based_Partially_Distributed_File_System
4. Nabil Giweli (2013) *Enhancing Cloud Computing Security and Privacy*, 20, Jan, 2019 [Online]Available:https://www.researchdirect.westernsydney.edu.au/islandora/object/uws%3AI7310/.../view
5. Zhou Miao (2013) *Data Security and Integrity in cloud computing*, Doctor of Philosophy thesis, School of Computer Science and Software Engineering, University of Wollongong. http://www.ro.uow.edu.au/thesis/3990
6. L.M. Kaufman, Data security in the World of Cloud Computing. IEEE Security and Privacy, 2009. 7(4): p. 61-64.
7. Sudhansu R. L. et.al *Enhancing Data Security in Cloud Computing Using RSA Encryption and MD5 Algorithm*, International Journal of Computer Science Trends and Technology (IJCST) – Volume 2, Issue 3, June 2014
8. Nesrine Kaaniche (2014) *Cloud Data Security based on Cryptographic Mechanisms*, 26 Jan, 2019 [Online] Available: https://www.tel.archives-ouvertes.fr/tel-01146029/document
9. Afnan U.K. (2014) Data Confidentiality and Risk Management in Cloud Computing 2 Feb, 2019 [Online] Available: https://www.ethesis.whiterose.ac.uk/13677/1/Thesis_Final_Afnan _27072016_ EngD.pdf
10. Sarojini G. et.al (2016) Trusted and Reputed Services using Enhanced Mutual Trusted and Reputed Access Control Algorithm in Cloud, 2nd International Conference on Intelligent Computing, Communication & Convergence (ICCC-2016). www.sciencedirect.com
11. Dimitra A. G. (2017) Security Policies for Cloud Computing, 26 Jan, 2019 [Online] Available: https://www.dione.lib.unipi.gr/xmlui/bitstream/handle/unipi/11007/Georgiou_Dimitra.pdf?
12. https://en.wikipedia.org/wiki/Cloud_computing_issues retrieved on 23/04/2019 @ 19:15
13. S. Pearson and A. Charlesworth, Accountability as a way forward for privacy protection in the cloud. Hewlett-Packard Development Company, 2009.
14. Siani Pearson, Yun Shen, and M. Mowbray, A privacy manager for cloud computing. In CloudCom, 2009: p. 90-106.
15. M.C. Mont, S. Pearson, and P. Bramhall. Towards accountable management of privacy and identity information. in Proc. of the European Symposium on Research in Computer Security. 2003
16. D. H. Rakesh, R. R. Bhavsar, and A. S. Thorve, “Data security over cloud,” International Journal of Computer Applications, no. 5, pp. 11–14, 2012.

17. K. Benson, R. Dowsley, and H. Shacham, “Do you know where your cloud files are?” in Proceedings of the 3rd ACM workshop on Cloud computing security workshop, pp. 73–82, ACM, October 2011.
18. J. Krumm, “A survey of computational location privacy,” Personal and Ubiquitous Computing, vol. 13, no. 6, pp. 391–399, 2009.
19. S. Pearson and A. Benameur, “Privacy, security and trust issues arising from cloud computing,” in Proceedings of the 2nd IEEE International Conference on Cloud Computing Technology and Science (CloudCom ’10), pp. 693–702, IEEE, December 2010. E. Stefanov, M. van Dijk, E. Shi et al., “Path oram: an extremely simple oblivious ram protocol,” in Proceedings of the ACM SIGSAC Conference on Computer & Communications Security, pp. 299–310, ACM, 2013.
20. S. Paquette, P. T. Jaeger, and S. C. Wilson, “Identifying the security risks associated with governmental use of cloud computing,” Government Information Quarterly, vol. 27, no. 3, pp. 245–253, 2010.
21. S. Subashini and V. Kavitha, “A survey on security issues in service delivery models of cloud,” Journal of Network and Computer Applications, vol. 34, no. 1, pp. 1–11, 2011.