International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Cloud Computing Data Storage Security: A Comprehensive Review

Dr. L. K. Suresh Kumar

Submitted:25/01/2024 **Revised:** 08/03/2024 **Accepted:** 15/03/2024

Abstract: Cloud computing is a transformative paradigm that alters the design and purchase of corporate hardware and software. The convenience of cloud computing is prompting the migration of data and application software to cloud data centers. The Cloud service provider (CSP) must guarantee integrity, availability, privacy, and confidentiality; yet, the CSP is failing to provide dependable data services to customers and to safeguard stored customer data. This paper delineates the challenges associated with cloud data storage, including data breaches, data theft, and the unavailability of cloud data. This document provides an overview of security concerns related to data storage and potential solutions. Numerous firms are unprepared to use cloud computing technology owing to inadequate security control policies and vulnerabilities in protection, resulting in various challenges in cloud computing. An autonomous procedure is necessary to ensure that data is accurately housed on the cloud storage server. This study will examine several strategies used for safe data storage in the cloud. This article presents a method to prevent collusion attacks resulting from illegal server modifications.

Keywords: *Introduction, Cloud Computing And Cloud Storage, Cloud Storage Security, issues, solution, conclusion. Cloud service provider (CSP), cloud data storage, security issues, policies & protocols;*

1. Introduction

Cloud computing is a transformative technique that alters the design and purchase of corporate hardware and software. Cloud computing offers several advantages to customers, including cost-free services, resource flexibility, and convenient internet access. From small to big firms, there is a significant inclination towards cloud computing to enhance their company operations and collaborations with other organizations. Despite the substantial advantages of cloud computing, users are reluctant to save their secret or sensitive information, including personal health records, emails, and government-sensitive documents. Once data is stored in a cloud datacenter, the cloud client relinquishes direct control over their data sources. The Cloud Service Provider (CSP) promises to provide data security for cloud customers' stored data using technologies such as firewalls and virtualization. These systems would not provide comprehensive data security due to their weaknesses across the network, and Cloud Service Providers (CSPs) have full control over cloud applications, hardware, and client data. Encrypting sensitive data before to hosting helps ensure data privacy and confidentiality against Cloud Service Providers. A common issue with encryption schemes is their impracticality due to substantial communication overheads associated with cloud access patterns. Consequently, cloud computing need safe techniques for data storage and administration to maintain confidentiality and privacy. [2][5]. This article primarily examines security risks and concerns about confidentiality and privacy of customer data. Cloud computing amalgamates several pre-existing technologies that have evolved at disparate speeds and within distinct settings. The objective of cloud computing is to enable consumers to use these technologies.

Numerous enterprises are transitioning to the cloud since it enables users to store data remotely and access it at any time from any location. From small to big firms are increasingly using cloud computing to enhance their company operations and partnerships with other organizations. Security and privacy are significant obstacles in cloud computing, namely in maintaining the confidentiality, integrity, and availability of data. Cloud computing has transformed the outsourcing landscape (SaaS, PaaS, and IaaS) by offering more cost-effective and powerful processors inside cloud computing architectures. A computer primarily stores information in available space and retrieves it upon request from an authorized user.

Associate Professor, Department of Computer Science & Engineering , UCE, Osmania University
lksureshkumar@osmania.ac.in

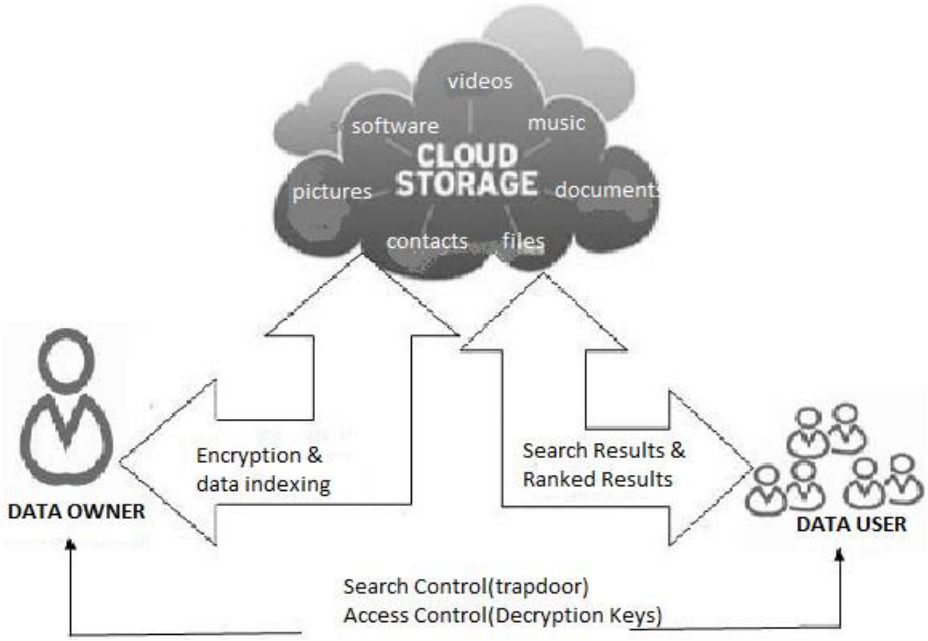

Fig 1: Cloud data storage model.[1]

2. Literature Review

SecCloud, as introduced by Wei et al. [12], offers a storage security protocol for cloud customers' data, ensuring the protection of both stored and computed data. The SecCloud protocol employs encryption for the safe storage of data. The multiplicative groups and cyclic additive pairs are used for key generation for cloud clients, cloud service providers, and other business partners or trustworthy third parties. The encrypted data, accompanied by the verified signature, is sent to the cloud data center together with the session key. The Diffie-Hellman method is used for the production of session keys for both bilinear groups. The cloud receives encrypted data, decrypts it, validates the digital signature, and saves the original data in a designated area inside the cloud. SecCloud ascertains the presence of data at the designated location. The Merkle hash tree is used for computational security in the SecCloud protocol. The verification agency will authenticate the computational findings generated via the use of a Merkle hash tree. The File Assured Deletion (FADE) protocol offers key management with data integrity and privacy in [15].

The File Assured Deletion protocol (FADE) developed in [18] ensures key management, data integrity, and privacy.Due to its simplicity, FADE is a lightweight protocol that employs both asymmetric and symmetric key encryption for data. The Shamir system safeguards both symmetric and asymmetric keys to enhance confidence in key management. A cohort of principal managers is used by the FADE protocol, serving as a trusted intermediary. The key k serves as the encryption key for the client's file F, while another key is used for encrypting the data key (k). The policy file contains the information of which files are accessible. To upload data, the user asks the key pair from the third party by transmitting the policy file p. The key manager transmits public and private keys to the user using the policy file. The uploaded file is encrypted using a randomly generated key, which is further encrypted using a symmetric key.

The encrypted file is decrypted using the public key from the formed key pair, and a MAC is produced for integrity verification. The receiver will do the inverse operation to get the original data.

Liu et al. [15] developed a time-based re-encryption strategy using the ABE algorithm to provide safe data exchange within a group, including access control. This approach guarantees the secure delivery of forwarded data to group users while preserving user revocation. This approach associates a time period with each user, leading to automated revocation by the Cloud Service Provider (CSP) upon expiry. This time-based encryption approach enables users to pre-share keys with the Cloud Service Provider (CSP), which thereafter generates re-encryption keys upon user request. The ABE protocol guarantees access control by evaluating a collection of qualities instead than relying on identification. This approach guarantees the privacy and availability of data among group members but does not focus on data integrity.

Probabilistic sampling is used to mitigate computational redundancy rather than reconstructing the whole tree. The following list comprises essential suggestions from the Computer protection Alliance (CSA) [18] for data protection and efficient key management. The scope of the key should be maintained by either a group or an individual. Standard encryption techniques should be used, whereas weak algorithms should be eliminated. Optimal protocols for key management and encryption software should be used; using genuine software technology is preferable to assure security in storage. The

consumer, companies, and/or trustworthy third parties should maintain efficient key management. If the auditing protocol is improperly built, the encryption process may regulate the data flow to external entities during the audit. However, encryption does not entirely inhibit data transmission to other entities; rather, it may only mitigate it to a small extent. However, it necessitates a substantial variety of key management processes and overheads for key creation during data storage. However, the vulnerability of encryption keys results in data leakage, which remains an issue in cloud environments. This issue is resolved by integrating the homomorphic authenticator with the random masking procedure [19]. The authors introduced Simple Privacy-Preserving Identity Management for Cloud Environments (SPICE) in [20] for identity management systems. SPICE guarantees group signatures to provide anonymous authentication, access management, accountability, unlinkability, and user-centric permission. The SPICE offers the aforementioned qualities with a single registration. Upon user registration with a trustworthy third party, unique credentials are obtained for all services offered by the CSP. The user produces an authentication certificate using the credentials.

Various CSPs want diverse qualities for authentication, necessitating the user to create the requisite kind of authentication certificate using same credentials.

The Role-Based Multi-Tenancy Access Control (RB_MTAC) has been suggested in [21]. The RB_MTAC integrates role-based access control with identity management. User registration with the CSP is necessary to get a unique single credential. The user must choose a password when registration with the CSP site. Utilizing these credentials, the user can access the cloud environment via an identity module that distinctly identifies the user. Subsequently, the user will be redirected to a role assignment module that connects to the RB_MTAC database and allocates roles to the registered user based on their enrolled information.

3. Cloud Data Storage Challenges & Issues

Data stored in cloud data centers is not something that cloud computing gives you control over. The data is completely at the mercy of the cloud service providers, who may alter, delete, or duplicate it at will. With cloud computing, you can manage your virtual computers with ease. Figure 1 shows that compared to the general cloud computing model, this one has more security problems because of the absence of data management.

While the one encryption method isn't foolproof, it does provide some protection over raw data. Cloud computing is characterized by virtualization and multi tenancy, which also presents more attack vectors than the typical cloud architecture. Several problems with figure 2 are detailed below.

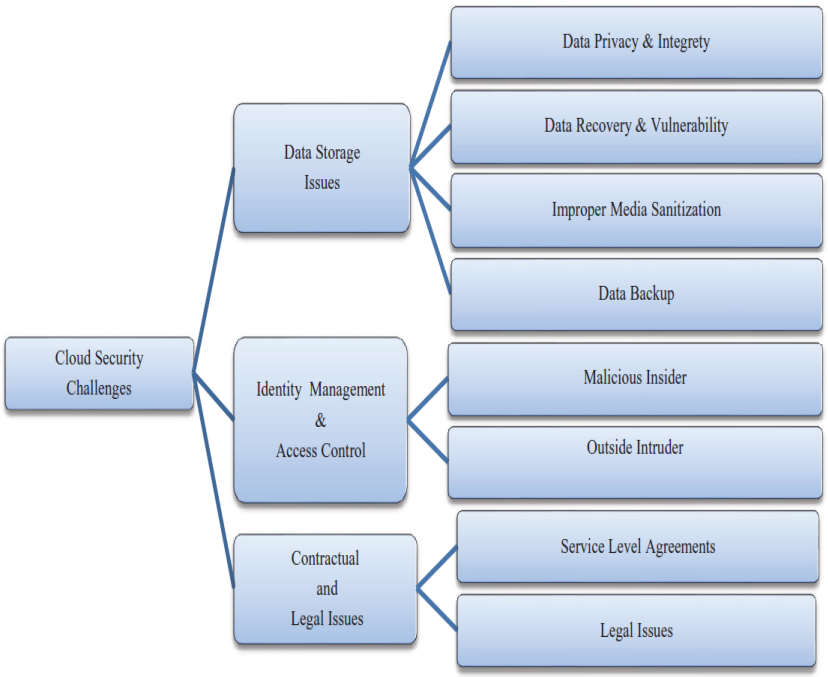

Fig 2. Cloud security Challenges[1]

CLOUD COMPUTING AND CLOUD STORAGE

Grid computing, distributed computing, parallel computing, utility computing, and virtualization are all examples of technologies that combine conventional computing with network technology to form cloud computing.

One method that offers services like data storage and company access is cloud storage. Incorporating the features of cluster applications, grid methods, distributed

file systems, etc., into its application software allows for the assembly of a wide variety of storage devices.

Powerful computing and processing functions are available on clouds, and users may tailor their service orders to their own requirements.

Among the many uses for cloud computing, cloud storage is among the most basic. Cloud storage options include Google Drive, Microsoft OneDrive, and Dropbox. All they do is let you put data in the cloud, and that's about it. On the other hand, cloud service providers like Google, Microsoft, Amazon, Dropbox, and countless more have access to far larger budgets and resources. With redundancy and backups in place, they can provide petabytes of storage for almost nothing. To get access to it, all you have to do is pay.

4. Cloud Computing Security

Cloud Computing Service Model

Numerous security issues must be evaluated before determining whether to transition to cloud computing. Organizations that transfer substantial volumes of sensitive data to an Internet-connected cloud environment increase their vulnerability to cyber assaults. Malware assaults provide a prevalent danger to safe access, and data indicates that as cloud utilization rises, security breaches become more probable for almost 90% of enterprises. Data leakage is becoming alarming for enterprises, with over 60% identifying it as their primary cloud security issue. The progression of technologies like network connection, distributed computing, and application computing significantly facilitates the implementation of cloud computing. Individuals will have swift access to virtual servers with less maintenance and interface costs [24]. The five functional characteristics of cloud computing include on-demand self-service, extensive network access, efficient resource sharing, high elasticity in computation, and metered payment support. The organization's use of cloud technology may be categorized as Software as a Service (SaaS), Platform as a Service (PaaS), and Infrastructure as a Service (IaaS), based on the distinct types of services offered by cloud computing. In the software as a service (SaaS) distribution paradigm, a cloud provider hosts programs and provides online access to customers. In this arrangement, an independent software vendor (ISV) may engage a third-party cloud provider to host the application. The server canister rental service offered by Microsoft and IBM may be regarded as an innovative industrial paradigm, referred to as hardware as a solution (HaaS). Figure 3 depicts this. Advanced cloud providers may independently provide client services or purchase information services from subordinate cloud providers [26] The ongoing elevation of the service mode level correlates with a growing complexity of the required service functions and circumstances.

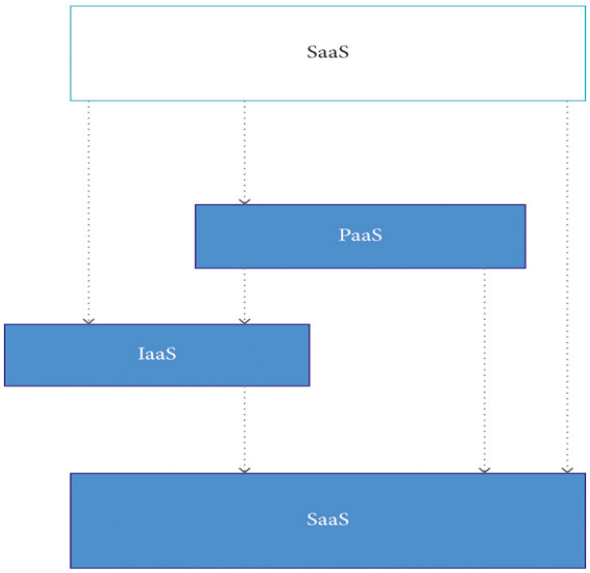

Fig 3: Cloud Computing Service Model[4]

Data Backup Technology

When you use a cloud backup service, your online data and apps are backed up and stored on a distant server. In order to ensure that their files and data are available in the event of a system failure, outage, or natural catastrophe, businesses choose to back up to the cloud. The underlying principle of business cloud storage is to replicate your server's data and store it on another server located in a different physical location. Whatever a business requires, it can back up its server files or just a subset of them. Software and hardware are the two main components of a computer system. Damage to the computer's hardware renders data inaccessible or recoverable, and the loss is permanent. Technology for backups and security audits is built into cloud computing. One of them is backup technology, which involves storing data on computer

hardware in a safe place [29]. In the event that a computer's software or hardware fails, the cloud computing system's recovery equipment will securely restore the system's crucial data. In the event of a data breach, the host cluster will immediately move the system's primary files to the recovery system, where they will be processed by the host.

5. Cloud Computing Storage Backup and Recovery

The data security of cloud computing has distinct features due to the fact that data storage processes are offered as services in cloud computing: The transmission method and the danger of data leakage are both increased since (1) user data is kept on the cloud server and (2) both upload and download must transit over the network. (2) Data is kept by an intermediary that is only partially trustworthy; (3) A dispersed network is the foundation of cloud computing; each user's data is saved on a node in this network. In principle, as mentioned before, an attacker may theoretically access the nodes around it via a certain node using a specific way [20].

Many in business, academia, and government are curious about cloud storage, which is a natural progression from cloud computing and related technologies. An up-and-coming method of storing data is cloud computing. It all comes down to putting data storage and management on a cloud platform, which gives users real-time Internet access. Numerous domestic and international IT behemoths, including Tencent, Baidu, and Ali, as well as global powerhouses like Amazon, Google, and Microsoft, have invested much in cloud storage research and development. Figure 4 shows an architectural design of a cloud-based mobile phone backup system[26].

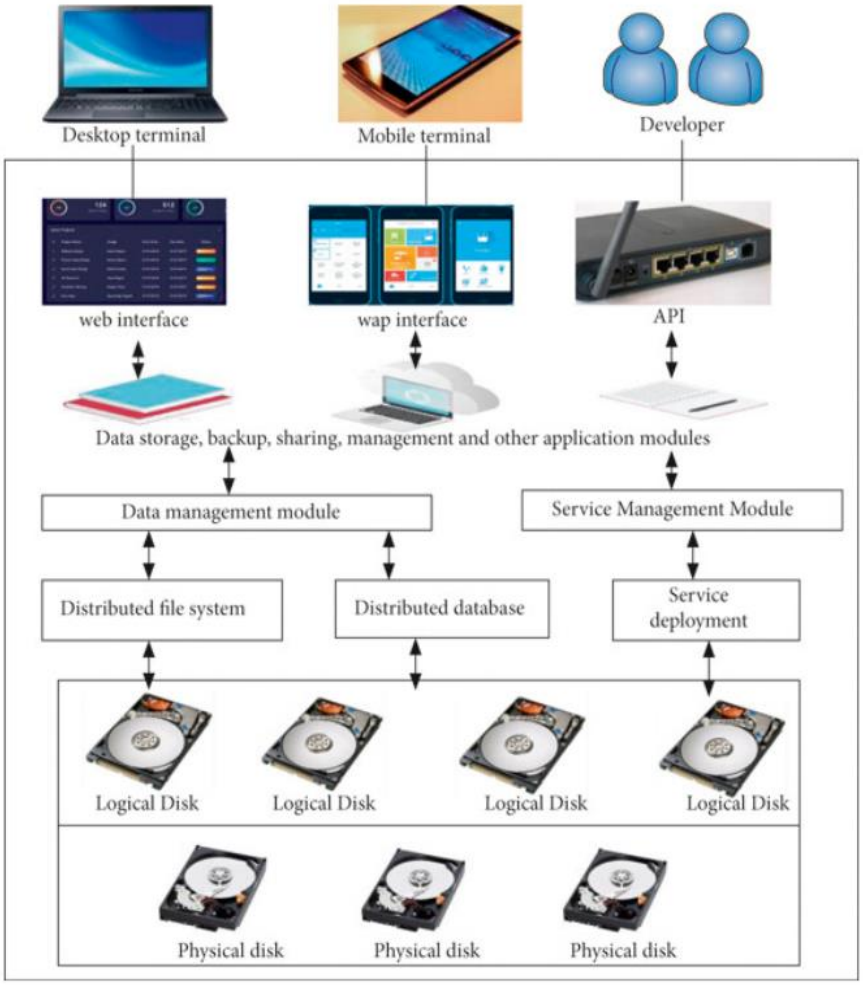

Fig 4: Architecture diagram of mobile phone backup system based on cloud storage (pictures from Baidu picture).[3].

Identity Management and Access Control

Data and service privacy and security are interdependent on identity and access management systems. To prevent unwanted access to the stored data, it is crucial to keep track of user identities. Since the data owner and the stored data are on separate executive platforms, cloud computing introduces complicated identification and access constraints. Organizations utilize a wide range of authentication and authorization strategies in the cloud. Over time, a complex scenario might be created by employing several mechanisms for authentication and authorization. Due to the elastic nature of cloud resources and the pay-as-you-go pricing model, IP addresses are dynamically assigned to new users whenever services are

launched or resumed. With this on-demand access policy in place, users of cloud resources may join and depart as needed. Access control and identity management must be efficient and effective for all these functions. When users join or leave the cloud, the identity management system must be updated and managed swiftly. Weak logging and monitoring capabilities, XML wrapping attacks on web sites, easily-reset credentials, and denial-of-service attacks to temporarily freeze the account are just a few of the numerous problems with identity and access management.

6. Solutions

In their presentation of SecCloud, Wei et al. outline a storage security protocol that protects both the data saved in the cloud and data used in computation.

Secure data storage is made possible by the SecCloud protocol through the use of encryption. When creating keys for cloud clients, CSPs, and other trustworthy third parties, multiplicative groups and cyclic additive pairing are used.

1) Data Discovery and Classification: Identify relevant data stored in databases and organise it into predetermined folders using descriptive labels, tags, or digital signatures.

2) Keep an eye on any configuration changes made to the cloud via change auditing. Thirdly, control and logging of events: make thorough records and report on user and workload audit trails. 4) Keep an eye out for any unwanted access to your sensitive data and take swift action to prevent it.

5) Authentication: To lessen the likelihood of illegal access to your data, apps, and systems, use multi-factor authentication (MFA).

6) Encryption of Data: Protect your data by including this essential extra layer of protection against illegal access. The session key and encrypted data are sent to the cloud data center with the verified signature. The session keys for both bilinear groups are generated using the Diffie-Hellman protocol.Data integrity may also be ensured via the deployment of effective auditing tools.

7. Conclusion

The cloud computing architecture facilitates data and application software storage with low administrative effort, delivering on-demand services to clients over the internet. However, with cloud management, customers lack reliable promises or rules. This will result in several security concerns with data storage, including privacy, confidentiality, integrity, and availability. This research concentrated on data storage security challenges in cloud computing, first presenting service models, deployment models, and various security concerns related to data storage in a cloud context. In the concluding part, we examined potential solutions for data storage challenges that provide privacy and secrecy in a cloud setting. Optimizing the performance of distributed storage systems and ensuring their high dependability in large-scale, big data, virtualization, and highly scalable cloud computing environments is a critical research challenge. The cloud computing architecture facilitates data and application software storage with low administrative effort, delivering on-demand services to clients over the internet. However, with cloud management, customers lack reliable promises or rules.

References

[1] Srikanta Patnaik, A Study on Data Storage Security Issues in Cloud Computing, 2nd International Conference on Intelligent Computing, Communication & Convergence (ICCC-2016), Procedia Computer Science 92 (2016) 128 – 135.

[2] Shweta Sukhram Prasad1, Aarti Kanahiyalal Yadava2, Research on Cloud Data Storage Security, International Journal for Research in Applied Science & Engineering Technology (IJRASET), ISSN: 2321-9653; IC Value: 45.98; Volume 10 Issue VI June 2022- Available at www.ijraset.com.

[3] Dajun Chang, Li Li, Ying Chang, Zhangquan Qiao, Cloud Computing Storage Backup and Recovery Strategy Based on Secure IoT and Spark, 23 November 2021, https://doi.org/10.1155/2021/9505249

[4] Fenglian Cao, Lihong Zhang, Darshana A. Naik, José Luis Arias Gonzáles, Neha Verma, Amit Jain, Rituraj Jain, Ashutosh Sharma, Application of Cloud Computing Technology in Computer Secure Storage, https://doi.org/10.1155/2022/4767725

[5] A. Abbas, K. Bilal, L. Zhang, S.U. Khan, A cloud based health insurance plan recommendation system: a user centered approach, Future Gener. Comput. Syst. (2014)

[6] P. Mell, T. Grance, The NIST definition of cloud computing (draft), NIST Special Publ. 800 (145) (2011) 7.

[7] J. Che, Y. Duan, T. Zhang, J. Fan, Study on the security models and strategies of cloud computing, Proc. Eng. 23 (2011) 586–593.

[8] R. Chandramouli, M. Iorga, S. Chokhani, Cryptographic key management issues and challenges in cloud services, in: Secure Cloud Computing, Springer, New York, 2014, pp. 1–30.

[9] C. Wang, Q. Wang, K. Ren, N. Cao, W. Lou, Toward secure and dependable storage services in cloud computing, IEEE Trans. Services Comput. 5 (2)(2012) 220–232.

[10] M. Balduzzi, J. Zaddach, D. Balzarotti, E. Kirda, S. Loureiro, A security analysis of amazon's elastic compute cloud service, in: Proceedings of the 27th

Annual ACM Symposium on Applied Computing, 2012, pp. 1427–1434.

[11] Duncan, Adrian, Sadie Creese, and Michael Goldsmith. "Insider attacks in cloud computing." Trust, Security and Privacy in Computing and Communications (TrustCom), 2012 IEEE 11th International Conference on. IEEE, 2012.

[12] [8] Khorshed, Md Tanzim, ABM Shawkat Ali, and Saleh A. Wasimi. "A survey on gaps, threat remediation challenges and some thoughts for proactive attack detection in cloud computing." Future Generation computer systems 28.6 (2012): 833-851.

[13] [9] A. Andrieux, K. Czajkowski, A. Dan, K. Keahey, H. Ludwig, T. Nakata, J. Pruyne, J. Rofrano, S. Tuecke, M. Xu, Web services agreement specification.

[14] S. Marston, Z. Li, S. Bandyopadhyay, J. Zhang, A. Ghalsasi, Cloud computing the business perspective, Decis. Support Syst. 51 (1) (2011) 176–189.

[15] B. Hay, K. Nance, M. Bishop, Storm clouds rising: security challenges for IaaS cloud computing, in: 44th Hawaii International Conference on System Sciences (HICSS), IEEE, 2011, pp. 1–7.

[16] L. Wei, H. Zhu, Z. Cao, X. Dong, W. Jia, Y. Chen, A.V. Vasilakos, Security and privacy for storage and computation in cloud computing, Inform. Sci. 258 (2014) 371–386.

[17] O.D. Alowolodu, B.K. Alese, A.O. Adetunmbi, O.S. Adewale, O.S. Ogundele, Elliptic curve cryptography for securing cloud computing applications, Int. J.Comput. Appl. 66 (2013).

[18] M. Aslam, C. Gehrmann, M. Bjorkman, Security and trust preserving VM migrations in public clouds, in: IEEE 11th International Conference on Trust,Security and Privacy in Computing and Communications (TrustCom), 2012, pp. 869–876.

[19] Y. Tang, P.P. Lee, J.C.S. Lui, R. Perlman, Secure overlay cloud storage with access control and assured deletion, IEEE Trans. Dependable Secure Comput.9 (6) (2012) 903–916.

[20] Q. Liu, G. Wang, J. Wu, Time-based proxy re-encryption scheme for secure data sharing in a cloud environment, Inform. Sci. 258 (2014) 355–370.

[21] Z. Tari, Security and privacy in cloud computing, IEEE Cloud Comput. 1 (1) (2014) 54–57.

[22] Cloud security alliance, security guidelines for critical areas of focus in cloud computing v3.0, 2011.

[23] Y. Fu, Z. Lin, Exterior: using a dual-vm based external shell for guest-os introspection, configuration, and recovery, in: Proceedings of the 9th ACMSIGPLAN/SIGOPS International Conference on Virtual Execution Environments, 2013, pp. 97–110.

[24] S.M.S. Chow, Y. He, L.C.K. Hui, S.M. Yiu, Spicesimple privacy-preserving identity-management for cloud environment, in: Applied Cryptography and Network Security, Springer, Berlin, Heidelberg, 2012, pp. 526–543.

[25] S. Yang, P. Lai, J. Lin, Design role-based multi-tenancy access control scheme for cloud services, in: IEEE International Symposium on Biometrics and Security Technologies (ISBAST), 2013, pp. 273–279.

[26] Shahanawaj Ahamad, Mohammed Abdul Bari, Big Data Processing Model for Smart City Design: A Systematic Review “, VOL 2021: ISSUE 08 IS SN : 0011-9342 ;Design Engineering (Toronto) Elsevier SCI Oct : 021;Q4 Journal