Journal of Advances in Science and Technology
Vol. 17, Issue No. 1, March-2020, ISSN 2230-9659

A Study of Security Architecture for Big Data Cloud and Big Data Outsourcing

Surina Jaiswal[1]*, Dr. Tryamvak Hiwarkar[2]

[1] Research Scholar,Sardar Patel University Balaghat MP,School of Computer Science and Technology

[2] Professor, Sardar Patel University Balaghat MP, School of Computer Science and Technology

Abstract - Due to the exponential growth of unstructured data in today's information world, big data infrastructure is essential for a wide range of applications across various industries. Because of the variety of information sources and the volume of data involved, data ingestion and retrieval are important processes during which user data must be protected. Accordingly, authors proposed a big data security architecture using split and merge security methods in big data. Organizations conducting big data initiatives and data security specialists will benefit from this work. And mainly study in which discussed about Cloud Security, Cloud Data Security Issues, Cloud Security Requirements, and Challenges in Storage of Big Data over Cloud, Cloud Infrastructure, and Security in Cloud Infrastructure, Data Security, and Security Issues in Cloud Computing.

Keyword - Data, Cloud

----------X----------

INTRODUCTION

Cloud is a word established as a common paradigm in the current digital world. Cloud computing employs the virtualized platform containing the elastic resources including software, hardware, services, and data sets by enabling dynamic provisioning of the pool of virtual resources to the cloud users. The aim of cloud computing is to transform the conventional desktop computing towards service oriented computing employing huge clusters and data centers. Cloud computing affects its ease and cheap cost to the cloud clients utilizing the notion of virtualization. Cloud computing known as Internet based computing has arisen from the principles of distributed computing, utility computing, grid computing. It incorporates the concepts like virtualization, multitenancy, service oriented architectures for offering infrastructure, platform, and software as the service rather than a product. A cloud model contains the five important elements to benefit the consumers and suppliers. On-demand, self-service: The cloud customers may automatically employ the computing resources such as servers, CPU cycles, and memory and network storage whenever required without intervening human involvement with the cloud service provider. A wide network access, the cloud services are constantly becoming accessible by heterogeneous thin or thick client devices such mobile phones, laptops, desktops, tablets and workstations. (S. Subashini and V. Kavitha) Resource pooling: The cloud service provider's computing resources are brought together to serve several cloud users utilizing a multi-tenant architecture. Rapid elasticity: The computer resources may be provided and released elastically, in certain circumstances frequently, to scale up and scale down according to the consumer's demand. Measured service: The utilization of cloud resources may be monitored, tracked, managed, and reported daily or weekly or monthly by both the service provider and the customer.

Cloud Security

Additionally, cloud computing adds an additional degree of risk since vital services are typically outsourced to a third party. It is more difficult to protect data integrity and privacy when it is externalized, to support data and service availability and to show compliance as a result. As a result, cloud computing shifts much of the control over data and operations away from the client organisation Even the most basic tasks, such as updating software and configuring firewalls, may fall under the purview of the cloud service provider rather than the end user. Customers must build confidence with their service providers and understand risk in terms of how these providers create, deploy, and maintain security on their behalf in order to protect themselves. Because of the dangers of outsourcing services, several firms opt for private or hybrid clouds instead of public ones. Many other areas of cloud computing must be re-examined in terms of risk and security. There are four authors whose names begin with the letters T: When data is kept on the cloud, it's almost impossible to pinpoint where it is physically. Previously evident security procedures have been abstracted away. Many security and

www.ignited.in

compliance issues can arise as a result of this lack of visibility.

Security in cloud computing is fundamentally different from that of more conventional IT settings because of the vast infrastructure sharing involved. Workload balancing, changing Service-Level Agreements (SLA) and other characteristics of today's dynamic IT infrastructures present even more potential for misconfiguration, data breach, and criminal activity.

An automated approach to infrastructure sharing may assist to eliminate operator error and supervision by reducing the possibility of human mistake and oversight. Cloud computing solutions, however, must still place a major focus on isolation, identification, and compliance because of the dangers associated with a widely shared infrastructure.

1. Cloud Data Security Issues

Data security is one of several concerns with cloud computing security. When moving data to the cloud, companies must ensure the safety of both the data while it is in transit and the data while it is in storage, many distinct forms of cloud data exist, including user identification, audit, runtime, and application data. The level of protection required will vary based on the data type. The level of security should be very high if the data is particularly sensitive, such as a user database. (V. Kavitha and S. Subashini) a simple privacy promise is all that's required if the data is only user identifying information. Depending on the kind of data, the user will have a varying level of security concern. Authentication, Confidentiality, Integrity, Scalability of keys, Access control, Policy and Compliance are some of the data security challenges that are addressed in this diagram of a cloud system's architecture.

Cloud Security Requirements

Confidentiality, integrity, and availability are three of the most critical characteristics of security. Authentication, access control, privacy, trusts and audit and compliance needs must also be met by cloud models.

1. Confidentiality

The word "confidentiality" refers to the fact that only those who have the proper authority and permission may access protected data. Data security is jeopardized in the cloud, where an increasing number of devices and services are sharing the data. Virtualization, infrastructure reusability, and multi-tenancy all pose serious privacy and confidentiality risks. (J. Yang, Z. Chen) Data remanence makes it possible for privacy to be breached accidentally. Even after data has been deleted or eliminated, its remnants should be wiped.

2. Integrity

The data, hardware, and software that make up the cloud are all considered resources. Cloud resources may only be edited, destroyed, or invented by those who have been granted access to them. It is essential that cloud providers and cloud users enforce the attributes of the stored data since the user's data is scattered around the globe. SLAs (Service Level Agreements) and technical compliance need cloud providers to assure accurate and trustworthy data modifications. Researchers: (H. Aljahdali, A. Albatli, P. Garraghan and others). Security procedures including auditing, fault recovery, replicating and responding to incidents and remediation should be available to cloud customers.

3. Availability

Users of the cloud should be able to access any computer resource (including infrastructures and applications) at any time, from any location. Denial-of-service attacks, Byzantine failures, resource shortages, and natural calamities may all have long-term or short-term effects on a system's availability. (M. Bahrami and M. Singhal, respectively) even in the event of system failures, security breaches, or distributed denial of service assaults, the cloud system should be able to continue out its normal activities. In order to keep clients happy, cloud service providers need guarantee that all of their data and resources are accessible at all times.

4. Privacy

Privacy is a person's willingness to share certain sensitive information with others. Personal and sensitive data must be handled in accordance with a country's laws. Cloud computing presents a number of legal and privacy issues owing to the fact that data is stored on computers located across the globe. This is due to the fact that national laws and regulations, as well as the level of compliance with them, might differ greatly. The cloud service providers should make sure that their consumers have complete visibility and control over the way their activities are carried out in the cloud. It is essential that cloud users have access to the right security measures to secure sensitive data.

5. Access control

Access control is a policy or practice that allows refuses or limits access to a system. Access control is critical in cloud computing because various users may have access to different resources at the same time or at different points in time. It is possible for cloud service providers to employ one of the pre-existing access control models, such as DAC, Mandatory Access Control, or RBAC, or to design their own.

6. Trust

www.ignited.in

The process of persuading consumers that the system is safe and proper relies on the concept of trust. Cloud customers are completely reliant on the cloud providers for a wide range of services. In order to utilize these services, consumers must provide the supplier access to their private information. For this reason, the creation of a general collection of trust-establishing factors is essential.

7. Audit and compliance

To guarantee the safety of the system, all data and resources will be monitored and recorded. Verifying that different access control rules are being followed is made easier for auditors with the aid of audit and compliance. Customer authorization and monitoring should be possible in the cloud. It's Regulatory and legal concerns should be addressed by a cloud provider, and customers should check the rules and processes of providers. Additionally, consumers want to be able to show that their data is safe, regardless of where it is stored.

Challenges in Storage of Big Data over Cloud

When managing applications, distributed computing focuses on pooling computing resources rather than using servers or individual devices, "Cloud" connotes "The Internet," hence Cloud Computing proposes a kind of computing in which service is delivered over the Internet. The goal of Cloud Computer is to use ever-increasing computing power to carry out a huge number of tasks at once. An vast collection of servers with particular relationships and proper data processing is used in cloud computing. Rather of delivering a pre-installed product suite for each PC, this sophisticated technology must include software that enables users to access Web-based support. In the words of R. A. Sana Belguith and Abderrazak Jemai: In addition, it includes all of the tasks that the client has specified. In a distributed computing system, there is still a significant amount of work to be done. When it comes to executing programmes, PCs no longer have to bear the brunt of the burden. Cloud computing technology is being used to reduce the cost of processing resources while they are not in use. The pile is managed by the Cloud organisation, which consists of a network of PCs. On the client side, the cost of programming and equipment is reduced. Big data systems need the use of traditional security measures, but they are not adequate. Data and application security for Big Data provide some new issues. In the present stage of Big data security, for example, the emphasis is on providing fine-grained protection via a wide inspection of stored data. The problem is that these models incentivize the mistreatment of customer information by apps and service providers. As a result, differential privacy has gained popularity as a means of protecting sensitive user data while also facilitating data analytics.

Cloud Infrastructure

In Cloud infrastructure, software components and hardware components are used such as storage, servers, virtualization and networking. These components are used to support the computing requirements of a cloud computing model. Cloud computing environment include the front end and back end components, cloud infrastructure consists of the back end components. Cloud infrastructure is present in cloud computing modes, i.e., Infrastructure as Service (IaaS), Software as a Service (SaaS) and Platform as a Service (PaaS). IaaS is a foundation; PaaS is the middle layer and SaaS is the top layer of the cloud computing stack. 1) IaaS: IaaS provides the cloud computing infrastructures like servers, storage, network and operating systems as on demand service. Rather than purchasing software's, network equipment, data center space services as on demand of users. Characteristics of IaaS: Allows for dynamic scaling, Resources are distributed as service, IaaS has a variable cost and utility model, and it includes multiple users on a single piece of hardware. Security requirements in IaaS: Cloud protection, communication security, images security. 2) SaaS: SaaS delivered applications to end user over the web. SaaS provides the software licenses to customers and software delivery model on a subscription basis and it's hosted centrally. It is specified to as on demand software. Characteristics of SaaS: SaaS provide the Central location for software delivery, one too many models are used for delivered the software, for doing the communication between different software's Application Programming Interfaces (APIs) is used, Software upgrades and patches are not handling by users. Security requirements in SaaS: Data protection, Access control, service availability, software security. 3) PaaS: PaaS is the set of Services and tools which are used to make coding and to deploy those applications efficient and quick. Services in PaaS are constantly updated with additional features added and with existing features are upgraded. Characteristics of PaaS: PaaS provide the services to test, host, develop, deploy and maintain applications in the same development environment, PaaS provides web-based user interface creation tools, which is used for creating, modify, test different user interface scenarios, PaaS provides the features to software like load balancing, scalability, and failover, PaaS provides communication tools and project planning for software development team. Security requirements in PaaS: Application security, access control.

Security in Cloud Infrastructure

Cloud computing produces services on demand basis, Characteristics of cloud computing are network access, self-service location independent resource pooling, transference of risk. Cloud computing is important for both the academic research word and the industrial world. Cloud computing is important for IT applications. Cloud computing provides different services and resources

to different sites and organizations. Cloud computing share the services and resources in distributed environment via network thus it makes security problem.

Data Security

In cloud computing environment data is distributed in different machines and storage devices like mobile devices and PCs, because of this reason data security is a major issue in cloud computing environment. Two important functions in cloud computing environment: data storage and computing of data. Customers of cloud services can get access to their data and finish their computing task just through the Internet connectivity. Clients have no idea where the data are stored and which machine execute the task of computing. In the research of the cloud computing, the some of the data security and data protection techniques have been proposed.

Data confidentiality: Data confidentiality technique is designed for protecting sensitive information from unauthorized users. Outsourced data is stored in a cloud computing environment, from these environment owners directly control data. Only authorized users can access the data services for, e.g., data computing, data sharing, data search. Data encryption method is used for ensuring confidentiality.

Data Access Controllability: In cloud computing environment all access of the data is controlled by the data owner. When the user accesses the data, owner checks authority of that user. Without authorization of the user, the owner cannot get the permission to access the data. Owner provides different access privileges to different users for controlling the access data.

Data Integrity: Data integrity maintains the accuracy and completeness of data. An owner always regards that data which comes from different sources in a cloud can be stored correctly and reliable. This shows that data should not be illegally modified, improperly tempered, maliciously fabricated or deliberately detected. If unwanted operation deletes or corrupt the data, the data owner should be able to detect the loss or corruption data.

A. Firewall Security

To protect the private network from unauthorized access firewall system is used. A Firewall can be implemented in software and hardware form or a combination of both. Firewall work like a filter between computer and internet. Firewall management program can be configured in two ways,

A default-deny policy: This firewall administrator allowed network services, and everything else is denied.

A default-allow policy: This firewall administrator provides the network services which are not allowed, and everything else is accepted.

A default-deny approach to the firewall is more secure than default-allow approach, but the difficulty occurs in configuring and managing the network.

The Cloud-based firewall provides three basic things: availability, scalability, and extensibility. Availability in cloud-based firewall provides high availability (>99.99%) through the infrastructure with the power redundant, network services as well as backup strategies when site failure. High availability is depending on the manufacturer. Scalability in cloud-based firewall delivered services to multiple users. Scalability comes in an enterprise when bandwidth increases. For increasing the bandwidth in accessing data firewalls are designed in cloud environments. Extensibility in cloud-based firewalls, network manager provides the secure communication path. Cloud-based firewalls extend the new functionality in network security.

Security Issues In Cloud Computing

Despite of its reputation, still, the use of cloud computing has introduced a range of significant privacy and security issues which deters its acceptance in sensitive environments. An organization that is migrating to a cloud computing model faces all these privacy and security challenges. It is because of the concepts of dynamism, heterogeneity, distributed nature of the cloud, and also due to the implementation of virtualization and multitenancy concepts. And also, the constituents of even a single service executed on cloud span across various trusted or untrusted domains. In this regard, data and services in the cloud are stored in the servers that are geographically distributed anywhere in the world so that the cloud users are unknown about the location of their data and services. These issues introduce many vulnerabilities and threats that complicate the problem of authentication and access control. Consequently, ensuring confidentiality and assuring the integrity of the stored data is the most challenging issue in cloud when compared with the traditional models. The security threats faced by the organizations with the adoption of the cloud computing model are as follows.

- Data theft– Stealing of sensitive information by unauthorized individuals,
- Broken authentication and Compromised credentials- Act of stealing users credentials like password, fingerprint, etc.,
- Insecure APIs- Risk may increase when the users access the CSP through insecure APIs.

www.ignited.in

- Exploited vulnerabilities in virtual machines- Attacks in one virtual machine harm all the other virtual machines running under the same hypervisor.
- Account hijacks- Stealing of user account details by the act of phishing.
- Malicious insiders- Data theft or data loss can occur due to the malicious insiders in the organization.
- Data scavenging- Searching in the residual representation of data to gain knowledge of the sensitive data,
- Permanent data loss- Data can be destroyed by Byzantine failures or neglecting the rarely accessed data.
- Inadequate diligence- Risk may increase when the cloud users do not fully understand the environment and its associated risks.
- Abuse of cloud services- Misuse of cloud services for launching attacks like DoS (Denial-Of-Service) attacks, phishing, etc.,
- Denial-Of-Service attacks- Attacks are imposed to make the service or data unavailable

CONCLUSION

Cloud can be envisioned as a flexible computing paradigm for its cost-effectiveness and dynamic provisioning. However, if the security challenges in the cloud are not handled well, it may deteriorate the growth and adoption of cloud. This thesis predominantly addresses the problem of storing and managing sensitive data in the cloud. The main conclusion of the research is the adoption of Confidentiality, Integrity and Authentication model by the cloud users and service providers. In this model, confidentiality of data stored in cloud has been ensured by encrypting the data using 2-Keys symmetric encryption. Data and computation integrity are assured by the Data Owners, by applying the data and computation assurance schemes. The remote user authentication scheme can be applied for allowing only authorized users to access the data and services. This Confidentiality, Integrity and Authentication model also allows the Data Owners to dynamically change the data stored in cloud without retrieving the entire data from cloud

REFERENCE

1. S. Subashini and V. Kavitha, “A survey on security issues in service delivery models of cloud computing,” Journal of Network and Computer Applications, vol. 34, no. 1, pp. 1 – 11, 2011.
2. M. T. Khorshed, A. S. Ali, and S. A. Wasimi, “A survey on gaps, threat remediation challenges and some thoughts for proactive attack detection in cloud computing,” Future Generation Computer Systems, vol. 28, no. 6, pp. 833 – 851, 2012.
3. Z. Zhou and D. Huang, “Efficient and secure data storage operations for mobile cloud computing,” in Proceedings of the 8th International Conference on Network and Service Management. International Federation for Information Processing, 2012, pp. 37–45
4. K. Kumar and Y.-H. Lu, “Cloud computing for mobile users: Can offloading computation save energy?” Computer, no. 4, pp. 51–56, 2010.
5. M. AlZain, E. Pardede, B. Soh, and J. Thom, “Cloud computing security: From single to multi-clouds,” in System Science (HICSS), 2012 45th Hawaii International Conference on, Jan 2012, pp. 5490–5499
6. Talebian, E. Ahmed, A. Gani, and 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, pp. 121–141, 2014.
7. S. Subashini and V. Kavitha, “A survey on security issues in service delivery models of cloud computing,” Journal of network and computer applications, vol. 34, no. 1, pp. 1–11, 2011.
8. J. Yang and Z. Chen, “Cloud computing research and security issues,” in Computational Intelligence and Software Engineering (CiSE), 2010 International Conference on. IEEE, 2010, pp. 1–3.
9. H. Aljahdali, A. Albatli, P. Garraghan, P. Townend, L. Lau, and J. Xu, “Multi-tenancy in cloud computing,” in Service Oriented System Engineering (SOSE), 2014 IEEE 8th International Symposium on, April 2014, pp. 344–35
10. R. A. Sana Belguith, Abderrazak Jemai, “Enhancing data security in cloud computing using a lightweight cryptographic algorithm,” ICAS 2015 : The Eleventh International Conference on Autonomic and Autonomous Systems, 2015.
11. Sookhak, M, Gani, A, Khan, MK & Buyya, R 2017, ‗Dynamic remote data auditing for securing big data storage in cloud computing‘, Information Sciences, vol. 380,

pp. 101-116

12. Jianghong Wei, Wenfen Liu & Xuexian Hu 2017, „Secure Data Sharing in Cloud Computing Using Revocable-Storage Identity-Based Encryption", IEEE Transactions on Cloud Computing, vol. 99, pp. 1-13.

13. Cihan Tunc, Salim Hariri1, Mheni Merzouki, Charif Mahmoudi, Frederic J de Vaulx, Jaafar Chbili, Robert Bohn & Abdella Battou 2017, ‗Cloud Security Automation Framework', IEEE 2nd International Workshops on Foundations and Applications of Self Systems (FASW), pp. 307-312.

14. Alhumrani, SA & Jayaprakash Kar 2016, „Cryptographic Protocols for Secure Cloud Computing", International Journal of Security and Its Applications, vol. 10, Issue 2, pp. 301-310.

15. Jainish Rajesh Jain & Abu Asaduzzaman 2016, ‗A Novel Data Logging Framework to Enhance Security of Cloud Computing', SoutheastCon 2016, 978-1-5090-2246-5/16 IEEE 2016.

Corresponding Author

Surina Jaiswal*

Research Scholar,Sardar Patel University Balaghat MP,School of Computer Science and Technology