Asian Journal of Computer Science and Technology
ISSN: 2249-0701 (P) Vol.9 No.1, 2020, pp.27-39

DOI: https://doi.org/10.51983/ajcst-2020.9.1.2153

An Analysis of Data Security and Privacy for Cloud Computing

Akella Subhadra
Associate Professor
Department of Computer Science & Engineering, BVCITS, Amalapuram, Andra Pradesh, India
E-mail: asubhadra2012@gmail.com
(Received 7 March 2020; Revised 28 March 2020; Accepted 15 April 2020; Available online 24 April 2020)

***Abstract* - Cloud Computing is the important buzzword in the today's world of computer. Cloud is global platform that allows digital information to be stored and distributed at very less cost and very fast to use. In these days since the data is very big in size many users are interested to store their valuable data in the cloud .The application software and databases in cloud computing are moved to the centralized large data centers, where the management of the data and services may not be fully trustworthy. Cloud Computing is scalable, fast, flexible, and cost-effective technology platform for IT enabled services over the internet. There are various advantages of cloud computing but ultimately cloud service users have to put their data over the cloud i.e., third party servers which are not directly controlled by the data owner Data security has consistently been a major issue in information technology In cloud computing in users perspective mainly in government ,industry and business Data security and ivacy protection issues are relevant to both hardware and software in the cloud architecture. Cloud security is becoming a key differentiator and competitive edge between cloud providers. In spite of various benefits that are provided by the cloud computing services, cloud computing service users are very much afraid about the security of their data .So this paper focuses on various issues regarding cloud computing, data security and how cloud provides data integrity, confidentiality, availability over user's data? How data stored over cloud storage servers will be protected from attackers? Risk management of data present on the Cloud is another challenge. There is a requirement to identify the risks an organization would be taking while hosting data and services on the Cloud. In this paper, we present those issues that are preventing people from adopting the cloud and to minimize risks of these issues.**
Keywords: **Cloud Computing, Integrity, Confidentiality, Availability.**

I. INTRODUCTION

There are several different definitions of cloud computing, but all of them agree on the way to provide services to users of the network. Cloud computing is an Internet-based development use of technology. It refers to the utilization of computing resources, hardware, and software, available on demand as a service over the web. It offers a variety of services for users of the network, like applications, storage, and various operations and remote printing, etc. [1]. It typically involves over the web provision of dynamically scalable and sometimes virtualized resources [2]. Businesses are running all types of apps within the cloud. Cloud computing are often considered because the technology that keeps the info, uses in several applications, and is remotely controlled without the necessity to download certain applications on computers.

Several the potential benefits that apply to most sorts of cloud computing includes the following:

1. Cost Savings: Companies can utilize operational expenses and reduce their capital expenses for the sake of increasing their computing capabilities.
2. Flexibility: The pliability of cloud computing allows companies to use additional resources in peak times, to enable them to satisfy consumer demands.
3. Reliability: Services using multi-redundant sites can helps in business continuity and disaster recovery.
4. Reduce Maintenance: Cloud service providers do the system maintenance that does not require application installations onto PCs.
5. Mobile Accessible: Mobile workers have increased productivity to systems accessible in an infrastructure available from anywhere.
6. Transparency: Additional servers to be added to the provisioned service without interrupting the service or requiring reconfiguration of the appliance delivery solution. If the appliance delivery solution is integrated via a management API, then transparency is additionally achieved through the automated provisioning and de-provisioning of resources.

The explanation of "cloud computing" from the National Institute of Standards and Technology (NIST) [2] is that cloud computing enables ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) which will be rapidly provisioned and released with minimal management effort or service provider interaction. According to the reason, cloud computing provides a convenient on-demand network access to a shared pool of configurable computing resources. Resources ask computing applications, network resources, platforms, software services, virtual servers, and computing infrastructure. Cloud computing are often considered as a replacement computing archetype which will provide services on demand at a minimal cost. The three well-known and commonly used service models within the cloud paradigm are software as a service (SaaS), platform as a service (PaaS), and infrastructure as a service (IaaS).

In SaaS, software with the related data is deployed by a cloud service provider, and users can use it through the online browsers.

In PaaS, a service provider facilitates services to the users with a group of software programs which will solve the precise tasks.

In IaaS, the cloud service provider facilitates services to the users with virtual machines and storage to enhance their business capabilities.

Data security becomes more serious within the cloud computing environment because data is scattered in different machines and storage devices including servers, PCs, and various mobile devices like wireless sensor networks and smart phones. Data security within the cloud computing is more complicated than data security within the traditional information systems.

Cloud computing environment provides two basic sorts of functions: computing and data storage. In the cloud computing environment, consumers of cloud services don't need anything and that they can get access to their data and finish their computing tasks just through the web connectivity. During the access to the info and computing, the clients don't even know where the info are stored, and which machines execute the computing tasks. Coming to data storage, protection of data and providing security are the main factors for gaining user's trust and making the cloud technology successfully used. Several knowledge protections and data security techniques are proposed within the research field of cloud computing. However, data protection related techniques got to be further enhanced. Services of cloud computing are provided across the whole computing spectrum. Nowadays, organizations and companies are moving and extending their business by adopting the cloud computing to lower their cost. This can contribute to free more man-powers to focus on creating strategic differentiation and business division of labour is clearer. The cloud is growing increasingly as it provides high performance computational services at lower cost. Famous IT companies such as Amazon, Microsoft azure, google, rock space have provided cloud service on the Internet.

Cloud computing brings several attributes that require special attention when it comes to trusting the system. The trust of the whole system depends on the info protection and prevention techniques utilized in it. Numerous different tools and techniques are tested and introduced by the researchers for data protection and prevention to realize and take away the hurdle of trust but there are still gaps which need attention and are required to be lined up by making these techniques much better and effective. The major issues within the cloud computing include resource security, resource management, and resource monitoring. Currently, there are no standard rules and regulations to deploy applications in the cloud, and there is a lack of standardization control in the cloud. In general, there are three major threats that are identified in cloud computing, namely, security, privacy, and trust. Security plays a critical role within the current era of long dreamed vision of computing as a utility.

The rest of paper is organized as follows:

Section II addresses a general view of cloud computing components, service models, and deployment models. Section III is a description of cloud computing threats, Section IV presents the security risks, Section V describes data security problems of cloud computing,Section VI gives an insight of data security issues ,Section VII representsSecurity Challenges in the CIA,Section VIII is description for how safe cloud data is And finally, the conclusions are summarized in Section IX.

II. CLOUD COMPUTING COMPONENTS & MODELS

Cloud Computing has been considered as the next generation architecture of IT Enterprise. The cloud computing is the type of network in which no central controller is present due to which security is the major issue of the network. This new paradigm makes many new security challenges.

Components in a cloud refers to the platforms, like front end, back end and cloud-based delivery and the network that used. The basic components of cloud computing in a simple topology are divided into 3 parts, namely clients, datacenter, and distributed servers.

Clients on cloud computing architecture are said to be the precise same things that are plain, old, everyday local area networks (LANs). They are mostly desktops. but they could even be laptops, tablet computers, mobile phones, or PDAs - all big drivers for cloud computing due to their mobility. Clients are interacting with to manage their information on the cloud.

Datacenter is collection of servers; It might be an outsized room within the basement of your building filled with servers on the opposite side of the planet that you simply access via the web. A growing trend within the IT world is virtualizing servers. That is, software is often installed allowing multiple instances of virtual servers to be used. In this way, you'll have half a dozen virtual servers running on one physical server.

Distributed Servers may be a server placement during a different location. But the servers don't have to be housed in the same location. Often, servers are in geographically disparate locations.

Cloud Computing services have several components required, namely [3]

A. ***Cloud Clients***, a computer or software specifically designed for the utilization of cloud computing-based services

Example:
Mobile - Windows Mobile, Symbian
Thin Client - Windows Terminal Service, Cherry pal
Thick Client - Internet Explorer, Fire Fox, Chrome

B. ***Cloud Services***, services refer to products; solutions that are utilized are delivered through the media of internet.

Example:
Identity - Open ID, OAuth, etc.
Integration - Amazon Simple Queue Service.
Payments - PayPal, Google Checkout.
Mapping - Google Maps, Yahoo! Maps.

C. ***Cloud Applications,*** applications that use Cloud Computing in software architecture in order that users don't got to install but they will use the appliance employing a computer.

Example:
Peer-to-peer - Bit Torrent, SETI, and others.
Web Application - Facebook.
SaaS - Google Apps, SalesForce.com, and others

D. ***Cloud Platform***, a service within the sort of a computing platform consisting of hardware and infrastructure software. This service may be a service within the sort of a computing platform which contains infrastructure hardware and software.

Example:
Web Application Frameworks - Python Django, Ruby on Rails, .NET
Web HostingProprietary - Force.com

E. Cloud Storage involves the method of storing data as a service.

Example:
Database - Google Big Table, Amazon Simple DB.
Network Attached Storage - Nirvanix CloudNAS, MobileMe iDisk.

F. Cloud Infrastructure, it is a service that delivers infrastructure as a service.

Example:
Grid Computing - Sun Grid.
Full Virtualization - GoGrid, Skytap.
Compute - Amazon Elastic Compute Cloud

III. SERVICE MODELS OF CLOUD COMPUTING

SAAS (Storage-as-a-service) **-** This refers to the disc space we use once we lack a storage platform and thus request it as a service

Database-as-a-service - This component acts as a database directly from a foreign server where its functionality and other features work as if physical DB is present on the local machine.

Information-as-a-service - Information which will be accessed remotely from anywhere called Information-as-a-Service. Highlight the flexibility of accessing information remotely

Process-as-a-*service* - Unlike other components, this component combines various resources like data and services. This is mainly used for business processes where various key services and knowledge are combined to make a process.

Application-as-a-service (AaaS) - because the name suggests, this is often an entire package for accessing and using applications. This is made to attach end users to the web and end users usually use browsers and therefore the internet to access this service. This component is that the main front-end for end users

***Platform-as-a-service (PaaS*) -** during this component, the whole application development process takes place including creating, implementing, storing, and testing the database.

Integration-as-a-service - Mostly associated with application components that are built but must be integrated with other applications. This helps in mediating between remote servers and native machines.

Security-as-a-service - Because security is what most of the people expect within the cloud, this is often one among the foremost needed components.

Management / governance-as-a-service (MaaS and GaaS) - this is often associated with cloud management, like resource utilization, virtualization, and server up and downtime management.

Testing-as-a-service (TaaS)-Using these components, remote-hosted applications are tested against design requirements, database functionality, and security measures among other testing features.

Infrastructure-as-a-service (IaaS) - this is often an entire virtual consideration of networks, servers, software, and hardware on cloud platforms. Users will not be ready to monitor the backend process, but they're going to be presented with a system that's fully configured with all processes found out for direct use.

IV. DEPLOYMENT MODELS OF CLOUD COMPUTING

We may have different types of deployment models users may select a deployment model based on their requirements and availability [4].

***A. Private Cloud*:** a cloud that is used exclusively by one organization. The cloud may be operated by the organization itself or a third party. If the private cloud is properly implemented and operated, it has reduced potential security concerns.

***B. Public Cloud*:** a cloud that can be used (for a fee) by the general public, and involves an organization using a cloud infrastructure which is shared via the Internet with many other organizations and other members of the public; such

as Microsoft, Google, and Amazon [14]. Public cloud has variety of inherent security risks that need to be considered.

***C. Community Cloud*:** is shared by several organizations and is usually setup for their similar security requirements and a need to store or process data of similar sensitivity, such as several agencies of the same government [14].

***D. Hybrid Cloud*:** is a combination of cloud deployment models. Each cloud is independently managed while applications and data would be allowed to move across the hybrid cloud. If more resources are required private cloud can be transferred into public cloud. (14) A specific business and technology requirements are used in designing hybrids, which helps to optimize security and privacy with a minimum IT cost.

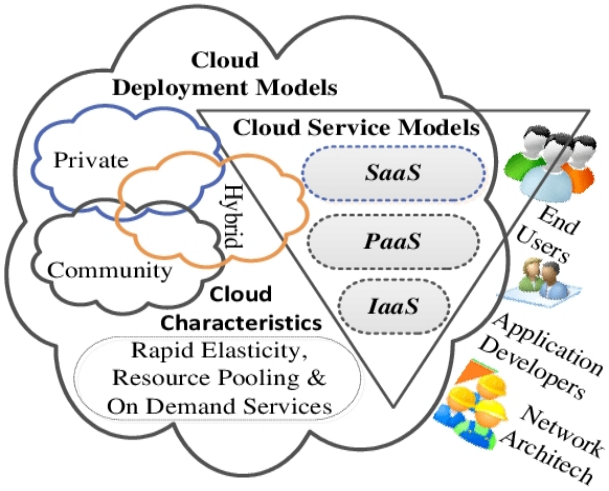

Fig.1 Cloud computing models

V.CLOUD COMPUTING SECURITY THREATS

Cloud computing provides cost savings and operational efficiencies to the users and organizations, it also leads to new security risks and uncertainties. The increased attack surface in a Cloud environment allows for other vulnerabilities to be exploited, thereby increasing the organization's risk .The risk is defined as a given threat that exploits vulnerabilities of an asset or group of assets and thereby cause harm to the organization. The increased attacks in cloud environment; virtual switches and hypervisor that are not present in the traditional data center, allows for other vulnerabilities to be exploited, thereby increasing the organization's risk.

The most important threats are identified as below [5],[6] :

***A. Data Breaches*:** The most important thing is to prevent ny data violation. The challenge addressing the threats of data loss and data leakage is that "the measures you put in place to improve one can worsen the other". Data is encrypted to reduce the impact of a violation, but if the encryption key is lost, then data will be lost. However, if offline backups of knowledge are chosen to scale back data loss, exposure data breaches are increased.

***B. Data loss/leakage*:** There are some ways to compromise data due to insufficient authentication, authorization, and audit (AAA) controls, like deletion or alteration of records without a backup of the original content. Loss of an encoding key may end in effective destruction. Unauthorized parties may gain access to sensitive data. A malicious hacker might delete a target's data.

***C. Account or Service Hijacking*:** Different hijacking methods like phishing, fraud, and exploitation of software vulnerabilities still produces results. If an attacker gains access to user credentials, he can pay attention to user activities and transactions, modifying the data, adding false information, and redirect the user clients to illegal sites.

***D. Insecure Application Programming Interfaces APIs*:** These interfaces must be designed to protect the user against both accidental and malicious attempts. some types of attacks including Anonymous access and reusable tokens or passwords, clear-text authentication or transmission of content, inflexible access controls or improper authorizations, limited monitoring, and logging capabilities; are samples of this sort of threats.

***E. Malicious Insiders*:** A provider might not reveal how it allows employee's access to physical and virtual assets, how it monitors these employees, or how it analyzes. In cloud computing, the organization does not need to know the technical details of how the services are delivered. In situations, the risk is great. Without full knowledge and control, your organization could also be in danger. In situations, the risk is great. Without full knowledge and control, your organization could also be in danger.

***F. Unknown Risk Profile*:** in order to know about organizations security status some important factors that are to be considered are Versions of software, modifications in the code , security policies and applications, vulnerability reports, interference attempts, and security design, and Information about who is sharing your infrastructure

***G. Cloud Abuse*:** Some providers offer free limited trial periods. Currently, spammers, malicious code authors, and other criminals are ready to conduct their activities with relative vulnerabilities, such as password and key cracking. By using cloud servers, a malicious hacker incorporates a Distributed Denial of Service (DDoS) attack, propagate malware, or share illegally copied software. The biggest challenge for cloud providers is how much data is hacked or modified by the hackers.

H. Shared Technology Issues: (IaaS) is predicated on shared infrastructure (e.g. disk partitions, CPU caches, GPUs, etc.), were not designed to offer strong isolation properties for a multitenant architecture. A virtualization hypervisor mediates access between guest operating systems and therefore the physical compute resources. Overlooked flaws have allowed guest operating systems to realize unauthorized levels of control and/or influence on the platform.

Fig.2 Cloud security threats

VI. CLOUD COMPUTING SECURITY RISKS

Virtualized servers are less secure than the physical servers. Even though it's not possible to reduce, all risks by moving operations to a cloud environment. While some risks are reduced, other risks may increase. With the addition of virtual network switches, hypervisors and virtual images, the attack surface increases. A single host with multiple virtual machines may be attacked by one of the guest operating systems, or a guest operating system may be used to attack other guest operating systems.

1. vulnerabilities are particularly risky because other virtual machines residing on the host and the data files stored out sides the owner's trusted domain
2. Movement from one provider to other while unencrypted or logged access. Anyone sniffing the network has an opportunity to extract sensitive data such as passwords or logins.
3. With virtualization, a customer's sensitive data is stored over a shared infrastructure that may be distributed on multiple sharing of servers and data centers.
4. Organizations should consider their risks thanks to anonymous signup, lack of validation, service fraud, and ad-hoc services.

Virtualized platforms Risks are:

1. Management console vulnerabilities
2. Management server vulnerabilities
3. Administrative VM vulnerabilities
4. VM vulnerabilities
5. Hypervisor vulnerabilities
6. Hypervisor escape

VII. DATA SECURITY PROBLEM OF CLOUD COMPUTING

Security problems are identified as follows [7]

A. Security Problem Drive from VM

Whether the IBM's Blue Cloud or the Microsoft's Windows Azure, the virtual machine technology is taken into account as a cloud computing platform of the elemental component, the differences between Blue Cloud and Windows Azure is that virtual machine running on Linux operating system or Microsoft Windows operating system.

Virtual Machine technology bring obvious advantages, it allows the operation of the server which is not any longer hooked into the physical device, but on the virtual servers. In virtual machine, a phase change or migration doesn't affect the services provided by the service provider. if user need more services, without considering the hardware the provider can satisfies the user's needs. However, the virtual server from the logical server group brings tons of security problems. The traditional data centre security measures on the sting of the hardware platform, while cloud computing could also be a server during a number of virtual servers, the virtual server may belong to different logical server group, virtual server, therefore there's the likelihood of attacking one another ,which brings virtual servers tons of security threats. Virtual machine extending the sting of

clouds makes the disappearance of the network boundary, thereby affecting most aspects of security.

B. The Existence of Super-user

For the enterprise providing cloud computing services, they need the proper to hold out the management and maintenance of knowledge, the existence of super-users to greatly simplify the info management function, but it's a significant threat to user privacy. Super-powers is a double-edged sword, it brings convenience to users and at the same time poses a threat to users. In an era of private privacy, personal data should be really protected, and therefore the incontrovertible fact that cloud computing platform to supply personal services within the confidentiality of private privacy on the existence of defects. Individual users and also the organizations have similar potential threats, e.g. corporate users and trade secrets stored within the cloud computing platform could also be stolen. Therefore, the utilization of super user rights must be controlled within the cloud.

C. Consistency of Data

Cloud environment the user's data transmits from the data centre to the user's client. For the system, the user's data is changing all the time. Read and write data concerning the identity of the user authentication and permission issues. In a virtual machine, there could also be different users' data which must be strict managed. The traditional model of access control is made within the fringe of computers, so it's weak to regulate reading and writing among distributed computers. Traditional access control is clearly not suitable for cloud computing environments. In the cloud computing environment, the normal access control mechanism has serious shortcomings.

D. New Technology

The concept of cloud computing is made on new architecture. The new architecture comprised of a spread of latest technologies, like Hadoop, Hbase, which reinforces the performance of cloud systems but brings in risks at an equivalent time. In the cloud environment, users create dynamic virtual organizations, first co-operation usually occurs during a relationship of trust between organizations instead of individual level. So those users supported the expression of restrictions on the idea of proof strategy is usually difficult to follow, which regularly occurs in many of the interactive nodes between the virtual machine, and is dynamic, unpredictable. Cloud computing environment provides a user of the "buy" the complete access to resources which has also increased security risks.

Fig.3 Cloud security problems

VIII. DATA SECURITY ISSUES OF CLOUD COMPUTING

Cloud computing security means providing guard to virtualized IP, data, applications, services by using broad set of policies, technologies, applications, and controls .there are several types of threats associated with cloud data services like network eavesdropping, denial of service attacks, side channel attacks, several vulnerabilities abuse of cloud services.

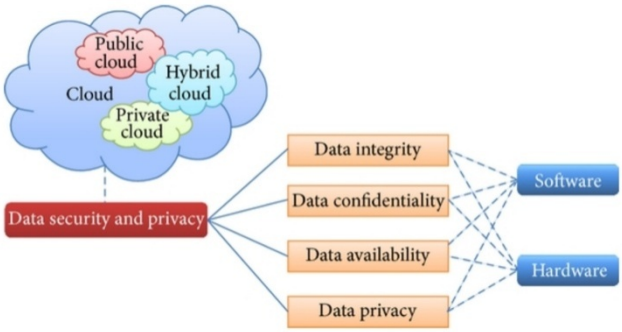

Fig.4 Data storage security issues

IX. DATA INTEGRITY ISSUES

Data integrity demands maintaining and assuring the accuracy and completeness of data. A data owner always expects that her or his data during a cloud are often stored correctly and trustworthily. It means the info should not be illegally tampered, improperly modified, deliberately deleted, or maliciously fabricated. If any undesirable operations corrupt or delete the info, the owner should be ready to detect the corruption or loss. Further, when some of the outsourced data is corrupted or lost, it can still be retrieved by the info users.

Data integrity is one among the foremost critical elements in any data system. Generally, data integrity means protecting data from unauthorized deletion, modification, or fabrication. Data integrity is achieved during a standalone system with one database. Data integrity in the standalone system is maintained via database constraints and transactions, which is usually finished by a database management system. Transactions should follow ACID (atomicity, consistency, isolation, and durability) properties to ensure data integrity. Most databases support ACID transactions and may preserve data integrity. Authorization is employed to regulate the access of knowledge. It is the mechanism by which a system determines what level of access a specific authenticated user should need to secure resources controlled by the system.

Data integrity within the cloud system means preserving information integrity. The data should not be lost or modified by unauthorized users. Data integrity is that the basis to supply cloud computing service like SaaS, PaaS, and IaaS. Besides data storage of large-scaled data, cloud computing environment usually provides processing service. The monitoring mechanisms offer the greater visibility into determining who or what may have altered data or system information, potentially affecting their integrity. Cloud computing providers are trusted to take care of data integrity and accuracy. However, it is necessary to create the third-party supervision mechanism besides users and cloud service providers. Verifying the integrity of knowledge within the cloud remotely is that the perquisite to deploy applications.

X. PROTECTING DATA INTEGRITY

There are two traditional ways of proving the integrity of knowledge outsourced during a remote server. Checking the integrity of knowledge are often by a client or by a 3rd party [10].

The first one is downloading the file then checking the hash value. In this way, a message authentication code algorithm is employed. MAC algorithms take two inputs, which are a secret key and variable length of data, which produce one output, which is a Message authentication code or message digest. In this way this algorithm is run on the client side. After getting a MAC, the info owner outsources those data to the cloud. For checking the integrity after receiving the data, the data owner downloads the outsourced data and then calculates the MAC for it and compares it with the one calculated before outsourcing that data. By using this method accidental and intentional changes are going to be detected. Also, by using the key, the authenticity of knowledge is going to be protected and only the one who has the key can check the info authenticity and integrity. For a large file, downloading and calculating the MAC of the file is an overwhelming process and takes a lot of time. Also, it is not practical since it consumes more bandwidth. Therefore, there is a requirement for employing a lighter technique, which is calculating the hashing value.

The other is to compute that hash value within the cloud by employing a hash tree. In this technique, the hash tree is made from bottom to top where the leaves are the info and fogeys also are hashed together until the basis is reached. The owner of data only stores the root. When the owner must check his data, he asks for just root value and compares it with the one he has. This is also to some extent is not practical because computing the hash value of an enormous number of values consumes more computation. Sometimes, when the provided service is simply storage without computation, the user downloads the file, an equivalent as within the first case, or send it to 3rd party, which will consume more bandwidth. Therefore, there is a requirement to seek out how to see data integrity while saving bandwidth and computation power. Remote data auditing, by which the info integrity or correctness of remotely stored data is investigated, has been given more attention recently [11], [12], [13], [14],[15], [16]

A. Third Party Auditor

Third Party Auditor (TPA) is the person who has the skills and experience to carry out all auditing processes. TPA scheme is employed for checking the info integrity. Since there are many incidents and doubtful actions, users of cloud storage depend upon third party auditors [17]. In [18], This scheme involves the data owner in the auditing process. The owner is conscious of all his resources on the cloud. Therefore, this scheme guarantees the integrity of knowledge for all owner resources on the cloud. First, TPA uses normal auditing processes. Once they discover any modification to the info, the owner is notified about those changes. The owner checks the logs of the auditing process to validate those changes. If the owner of the data suspects that unusual actions have happened, he can check his data by himself or by another auditor assigned by him. Therefore, the owner is usually tracking any modification to his own data. There is an assigned threshold value that a response from the third-party auditor shouldn't exceed. The data owner validates all modifications lesser than or adequate to this threshold. If the time exceeds this threshold, the info owner is meant to try to to surprise auditing.

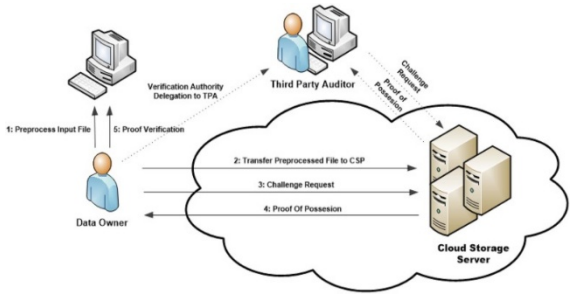

Fig.5 Data integrity using third party auditor

B. Proof of Retrievability

Proof of Retrievability is a cryptographic approach based on a challenge response protocol in which a piece of data is proved to be intact and retrievable without retrieving it from the cloud. The simplest sort of proof of Retrievability is taking the hash of block employing a keyed hash function. Owner of data takes the hash values of the file by using keyed hash function. After getting the hash values, the info owner keeps the key and therefore the hash values. the data owner sends the file to a foreign server. When the info owner must check his data Retrievability, he sends his key and asks the server to send the hash values by using his key so as to compare them with the hash values that data owner has. The advantage of this solution is that it is simple and implementable.

However, there are many disadvantages such the info owner must store many keys to use one whenever. Also, the amount of checking is restricted by the number of keys since the remote server could store all keys and therefore the hash values and use them when it is asked to prove having that file. In addition, it costs more resources on the side of a client and server since the hash values need to be calculated each time when the proof is required. Moreover, some thin client such mobile device and PDA does not have the resources to calculate the hash values of big files. In [21], They used an error correction code and spot checking to prove the possession and Retrievability of the info. Before sending them to the remote server the verified hides some sentinels. When the verifier wants to see Retrievability of the info, it only asks the server for those sentinels. To stay those sentinels indistinguishable for the remote server, the info owner encrypts the file after adding sentinels. In contrast to the straightforward one, it uses one key no matter the dimensions of the file. Also, unlike the straightforward solution that the whole file is processed, it accesses only parts of file. Therefore, the I/O operations is less. This scheme has disadvantages such that the files need to be in encrypted form, so it incurs computation overhead in clients.

C. Proof of Ownership

In this notion, the client proves ownership of the file outsourced by the client to server. The proof of ownership comes after the necessity to save lots of some storage by duplication. The owner of the files needs to prove to the server he owns this file. To prove the ownership are the Collision Resistant Hash functions and Merkle Hash Tree. In [22], The owner of a file creates a Merkle Hash Tree (MHT) and sends the file to the cloud, called verifier. Once it's received by cloud, the file is split into bits using pairwise independent hash then the verifier creates a Merkle Hash Tree for this file. Once the prover asks for the ownership of the file, the verifier sends a challenge, which is the root and the number of leaves. The prover calculates the sibling path and returns it to verifier as proof of ownership of this file. After receiving the sibling path, the verified checks this path against what the merkle tree has and validate the prover.

XI. DATA CONFIDENTIALITY ISSUES

Data confidentiality is important for users to store their private or confidential data in the cloud. Authentication and access control strategies are used to ensure data

confidentiality. The data confidentiality, authentication, and access control issues in cloud computing could be addressed by increasing the cloud reliability and trustworthiness. Because the users do not trust the cloud providers and cloud storage service providers are virtually impossible to eliminate potential insider threat, it is very dangerous for users to store their sensitive data in cloud storage directly. Simple encryption is faced with the key management problem and cannot support complex requirements such as query, parallel modification, and fine-grained authorization.

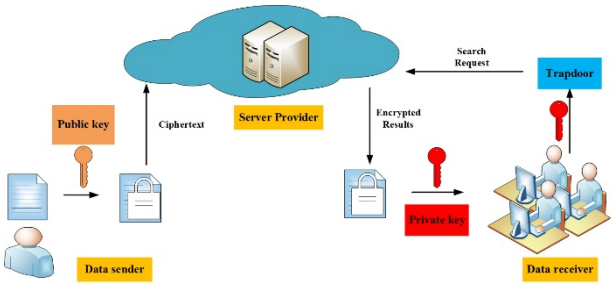

Fig.6 Public key encryption

Data confidentiality is the property that data contents are not made available or disclosed to illegal users. Outsourced data is stored in a cloud and out of the owners' direct control. Only authorized users can access the sensitive data while others, including CSPs, should not gain any information of the data. Meanwhile, data owners expect to fully utilize cloud data services, e.g., data search, data computation, and data sharing, without the leakage of the data contents to CSPs or other adversaries.

Usually the data is encrypted before it is outsourced. The service provider gets encrypted data. Therefore, it is considered not useful or meaningless. However, the client is responsible for handling the access control policy, encrypting the data, decrypting it and managing the cryptographic keys [56]. Even this would cause a burden to the user; sharing it with others exposes it to risks. When the data is shared among many users, there has to be more flexibility in the encryption process to handle users of the group, manage the keys between users, and enforce the access control policy in order to protect the data confidentiality [28].

Sharing the data among a group of users adds more burden on the owner of the outsourced data. In [30], the authors describe a cryptosystem in which the data owner encrypts the data by using his public key and identifiers called a class on the encryption process. Also, the owner has a master key to create others secret keys for one, some classes of data, or all classes of cipher text. Once the user gets his aggregate key, he only decrypts the class of cipher text this key is for. It is an aggregate key where each part of it can decrypt part of the cipher text.

The whole key can decrypt the whole cipher text. Therefore, this cryptosystem helps in sharing data among a group of users with fine grain access control and without giving them a key that can decrypt all that data. This figure8 shows the general view of this system.

A. Access Control

When data is outsourced to the cloud, which is untrusted because it is in a domain where security is not managed by the data owner, data security must be given more attention. When quite one entity wants to share data, there has got to be a mechanism to limit who can access that data. Many techniques are discussed within the literature. Those techniques were proposed to stay data content confidential and keep unauthorized entity from accessing and disclosing the info by using access control while permitting many authorized entities to share those data. The following are some of the techniques that are in the literature.

B. Public Key Encryption

Public key encryption is used to encrypt the data by using he public key. Only the one who has the private key can decrypt this data. There are many issues that make this way hard to apply in the cloud when many people need to access those files.In [60], Sana et el. proposed a light-weight encryption algorithm by utilizing symmetric encryption performance to encrypt files and utilizing asymmetric encryption efficient security to distribute keys. There are many disadvantages of using this method. One of them is vital management issue and therefore the got to get fine-grained access to file, such a part of it. Also, this solution is not flexible and scalable because encryption and decryption is needed when a user leave the group in order to prevent him from accessing the data. Key generation and encryption process.

C. Identity-Based Encryption (IBE)

The owner of data can encrypt his data by specifying the identity of the authorized entity to decrypt it based on that entity's identity, which must match the one specified by the owner.

1. Attribute Based Encryption (ABE) In attribute-based encryption, an identity of a user is identified by a set of attributes. This set of attributes generates the secret key.

Also, it defines the access structure used for access control. This access control is using encryption to encrypt data for confidentiality and share it among group of users. It is a kind of integrating the encryption with the access control. In [33], attribute-based encryption, known as fuzzy identity-based encryption, was proposed a few years after IBE. In this scheme, a group of attributes identify someone's identity. Data owner encrypts his data and only the one who has attributes that overlap with the attributes specified in the cipher text can decrypt it.

2. Key Policy Attribute Based Encryption (KP-ABE): This is more general than ABE because it expresses more conditions than just matching the attributes to enforce more control. In this mechanism, cipher text is linked with a set of attributes. The private key is linked to monotonic access structure. This access structure is predicated on a tree to specify the identity of the user. When the user's private key has the attributes that satisfy the attribute in cipher text, the user decrypts the cipher text. Key generation process is shown in figure 13 and encryption and decryption algorithm are shown in figure 14. A disadvantage of this method is that the descriptor must trust the key generator to get keys for an accurate person with the proper access structure. If the info must be re-encrypted, the new private keys need to be issued so as to stay accessing that data. Therefore, there's a requirement to urge the policy related to the key.

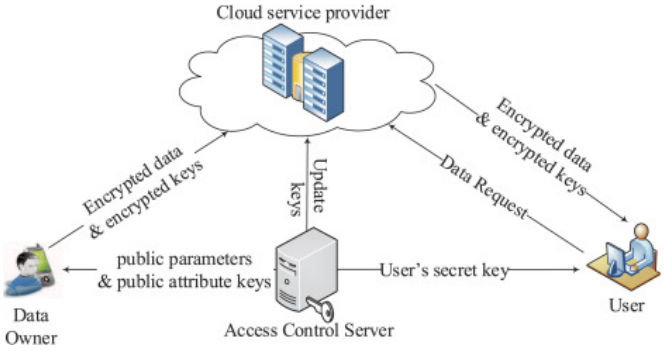

Fig.7 Attribute based encryption (ABE)

XII. DATA AVAILABILITY ISSUES

Data availability means the following: when accidents like hard disc damage, IDC fire, and network failures occur, the extent that user's data are often used or recovered and the way the users verify their data by techniques instead of counting on the credit guarantee by the cloud service provider alone. The issue of storing data over the trans boarder servers may be a serious concern of clients because the cloud vendors are governed by the local laws and, therefore, the cloud clients should be cognizant of those laws. Moreover, the cloud service provider should make sure the data security, particularly data confidentiality and integrity. The cloud provider should share all such concerns with the client and build trust relationship during this connection. The cloud vendor should provide guarantees of knowledge safety and explain jurisdiction of local laws to the clients.

XIII. SECURITY CHALLENGES IN THE CIA

Confidentiality, Integrity and Availability (CIA) losses can make a big impact in the business of the cloud computing because the data is the core component for any business. Data integrity is that the assurance given to the digital information is uncorrupted and only be accessed by those authorized users. Thus, integrity involves maintaining the accuracy, consistency, and trustworthiness of knowledge over its entire life cycle [8]. Maintaining CIA is simpler in enterprise computing but in cloud computing it's more complicated due to the multi-tenant architecture and therefore the distributed nature of the infrastructure. The following steps are often wont to maintain a correct CIA in cloud computing [8]

1. Once the data are created, classify the data, identify the sensitive data, define policies, and create access methods for different types of data. Also, create policies for data archive and data destroy.
2. Store data with proper physical and logical security protection, including the backup and recovery plan.
3. Identify which type of data can be shared, whom and how it can be shared and define data sharing policies. In cloud computing, many such policies are collectively called as Service Level Agreements (SLA).
4. Create a corrective action plan in case data is corrupted or hacked due to network or communication devices, security flaws while data is in transit.
5. Computation integrity refers to only the authorized applications can access the data and use it for computation. Any abnormality from normal computing should be avoided. An effective Identity and Access Management (IAM) can avoid loss of confidentiality and integrity.
6. Loss of availability can happen through loss of knowledge and data inaccessibility. Cloud computing employs few techniques like scalability and high availability at the architecture level. There are different methods and procedures are followed to enhance data security associated with the CIA triad at different stages of the info lifecycle.

Some of the important methods are listed below:

1. Apply data encryption when the data is at rest and when the data is in transit. Apply strong encryption algorithms like Advanced Encryption Standard (AES) and Rivest Shamir Adleman (RSA) algorithms.
2. Encryption methods are generally to provide confidentiality against attacks from a cloud provider, but it cannot protect data against configuration errors and software bugs [4]. Hash methods are often wont to determine accidental and intentional data changes. But they consume more bandwidth and time-consuming.
3. Third Party Auditing (TPA) can be employed to check for the data integrity. Many researchers [4, 18, 19] insists to audit data integrity by third-party auditors because they're specialized therein.
4. Do not store the encryption keys along with the encrypted data.
5. Implement proper Identity and Access Management (IAM) techniques for users to access data.

XIV. HOW SAFE IS CLOUD STORAGE?

For data and security practices Cloud provides end-to-end visibility Vendors frequently offer centralized cloud storage controls for managing users and data. Clouds reduce and sometimes eliminate the necessity for on-premises security architecture which will be configured inconsistently or incorrectly. To reflect changing security threats Cloud storage providers update systems quickly. Clearly, cloud storage adds additional concerns – and sometimes complexity – to a knowledge security strategy that result in lower costs and better overall data protection [9].

Cloud Storage Security Best Practices

It is vital to determine a cloud storage framework and cloud storage security standards. Here are five cloud storage best practices [9].

A. Assess your cloud framework

Secure cloud storage requires a corporation to spot all the devices and apps that hook up with the cloud. It is also vital to know what cloud storage systems exist within an enterprise, who uses them and the way they use them.

An organization can do high security cloud storage by mapping how data flows across systems, devices, applications, APIS and clouds. In general, many cloud storage applications display other apps and services they connect with. This can greatly simplify the task of mapping, and, if necessary, disconnecting from another app service.

B. Determine how cloud storage providers address privacy and security

Terms of service agreements are a good starting point for identifying the general protections a cloud provider offers. But it's not enough to make sure secure file storage. Cloud vendors frequently update terms of service and user agreements. This makes it relatively easy to overlook a seemingly minor change that can have a major impact on privacy and security. In addition, most agreements don't cover the details of how a cloud storage provider implements security, what specific protections it uses, and what happens in the event of a breakdown or breach. As a result, it's important to define policies and procedures closely. This may require further discussion and negotiation.

C. Know What Protections Are in Place?

Cloud security encryption is a fundamental requirement. It is important to know how a cloud storage provider uses encryption, including in transit between data centres, servers and storage devices – along with who controls encryption keys and how they are applied to specific data sets. Likewise, an organization using a cloud provider should know who has access to systems and what other protections it has in place to protect against everything from distributed denial-of-service (DDoS) attacks to application security flaws.

D. Put Data Classification Methods into Motion

All data is not created equal. Treating it an equivalent may be a recipe for security failures inside or outside a cloud environment. What is more, data security is becoming more complex as organizations accumulate larger volumes of unstructured data. Consequently, it's important to understand the value of data, whether it should be stored in the cloud or archived on media such as disk or tape, and how all of this translates into risk tolerance for the enterprise. Another factor is data compliance for government regulations such Sarbanes-Oxley or the General Data Privacy Regulation (GDPR) in the European Union. Some cloud storage services offer built-in tools to provide these processes.

E. Use Multi-Factor Authentication across All Devices and Systems.

The widespread use of multi-factor authentication can reduce the risk of someone gaining unauthorized access to a system or application and using it to unleash malware or gain a backdoor into other data. While the danger could also be greater for administrator accounts, it does not disappear for normal applications and tools. Multi-factor authentication can aid in protecting sensitive data from hackers, disgruntled employees and other insiders that may intentionally or inadvertently put data at risk. In the end, an enterprise can achieve strong cloud data security by focusing on computing and data frameworks across vendors

and learning how to use and manage new tools and techniques. It's also essential to figure closely with cloud providers to make sure that data storage security methods meet their requirements. A best practice approach helps an enterprise achieve the foremost secure cloud storage possible.

XV. CONCLUSION

Cloud computing may be a promising and emerging technology for subsequent generation of IT applications. The barrier and hurdles toward the rapid climb of cloud computing are data security and privacy issues. Reducing data storage and processing cost may be a mandatory requirement of any organization, while analysis of knowledge and knowledge is usually the foremost important tasks altogether the organizations for decision making. So no organizations will transfer their data or information to the cloud until the trust is made between the cloud service providers and consumers. However, there are many security issues coming with this technology as happens when every technology matures. Those issues are related to confidentiality of data from unauthorized people in remote sites, integrity of stored data in remote servers and therefore the availability of the info when it's needed. Also, sharing data in cloud when the cloud service provider is mistrusted is an issue.

Those issues include issues related to the previous issues of the internet, network issues, application issues, and storage issues. Storing data in a remote server leads to some security issues .A number of techniques have been proposed by researchers for data protection and to attain highest level of data security in the cloud. Many studies have been conducted to discover the issues that affect confidentiality, integrity, and availability of data to find a solution for them. Those solutions will lead to more secure cloud storage, which will also lead to more acceptance from the people and the trust on the cloud will increase. More work is required in cloud computing to make it acceptable by the cloud service consumers. To provide a secure environment and to protect sensitive static and dynamic data on cloud computing, firstly, components and service modals and deployment models are explained, and then different threats, vulnerabilities and risks are explained.

With the increase in the growth of cloud computing, security needs to be analyzed frequently. The Users should be aware of the different security problems present in the current cloud computing environment before being a part of the environment. In This paper we have Identified the most data security and Data Integrity, confidentiality, and availability Issues, which may be considered in cloud computing both by users and providers. This paper covered different challenges of CIA and focusing on how safe data storage in the cloud computing environments is to build trust between cloud service providers and consumers.

REFERENCES

[1] Data integrity in cloud computing security, *Article in Journal of Theoretical and Applied science*

[2] Data Integrity in Cloud Computing Security, *Journal of Theoretical and Applied Information Technology*, 31st December 2013. Vol. 58 No.3, Int 2 Brian O. and others, Cloud Computing, authors: 2012-11-06, page 6, publish Swiss.

[3] All info about cloud news article Components of Cloud Computing, By Cloud Storage February 14, 2019

[4] Data integrity in cloud computing security, Article in research gate, *Journal of Theoretical and Applied Information Technology* Vol. 58, No.3, pp.570, December 2013

[5] Cloud Security: Apponix Technologies,scribd,Cloud Security: Apponix Technologies | *Cloud Computing* | Public Key

[6] IT Blog: Top 10 Could Computing Threats

[7] *International Journal of Research in Science and Technology (IJRST*) 2013, Vol. 2, No. 5, Apr-June, Data Security modal for cloud computing

[8] Exploring Data Security Issues and Solutions in Cloud Computing, Article in research gate, *Procedia Computer Science*, January 2018.

[9] Enterprise storage, Cloud Storage Security Standards and Best Practices, By Samuel Greengard, Posted December 5, 2018

[10] Data Security, Privacy, Availability and Integrity in Cloud Computing: Issues and Current Solutions *(IJACSA) International Journal of Advanced Computer Science and Applications*, Vol. 7, No. 4, 2016

[11] C. Erway, A. Kupc¸ u, C. Papamanthou, and R. Tamassia, "*Dynamic provable data possession," in Proceedings of the 16th ACM conference on Computer and communications security*. Acm, 2009, pp. 213–222.

[12] G. Ateniese, R. Burns, R. Curtmola, J. Herring, L. Kissner, Z. Peterson, and D. Song, "*Provable data possession at untrusted stores," in Proceedings of the 14th ACM conference on Computer and communications security*. Acm, 2007, pp. 598–609.

[13] G. Ateniese, R. Di Pietro, L. V. Mancini, and G. Tsudik, "*Scalable and efficient provable data possession," in Proceedings of the 4th international conference on Security and privacy in communication netowrks*. ACM, 2008, p. 9.

[14] K. Yang and X. Jia, "An efficient and secure dynamic auditing protocol for data storage in cloud computing," *Parallel and Distributed Systems*, IEEE Transactions on, Vol. 24, No. 9, pp. 1717–1726, 2013.

[15] Q. Wang, C. Wang, K. Ren, W. Lou, and J. Li, "Enabling public auditability and data dynamics for storage security in cloud computing," *Parallel and Distributed Systems*, IEEE Transactions on, Vol. 22, No. 5, pp. 847–859, 2011.

[16] C. Wang, Q. Wang, K. Ren, N. Cao, and W. Lou, "Toward secure and dependable storage services in cloud computing," *Services Computing*, IEEE Transactions on, Vol. 5, No. 2, pp. 220–232, 2012.

[17] C. Wang, S. Chow, Q. Wang, K. Ren, and W. Lou, "Privacy-preserving public auditing for secure cloud storage," *Computers*, IEEE Transactions on, Vol. 62, No. 2, pp. 362–375, Feb 2013.

[18] B. Balusamy, P. Venkatakrishna, A. Vaidhyanathan, M. Ravikumar, and N. Devi Munisamy, "Enhanced security framework for data integrity using third-party auditing in the cloud system," in Artificial Intelligence and Evolutionary Algorithms in Engineering Systems, ser. *Advances in Intelligent Systems and Computing*. Springer India, 2015, Vol. 325, pp. 25–31.

[19] G. Ateniese, R. Burns, R. Curtmola, J. Herring, O. Khan, L. Kissner,Z. Peterson, and D. Song, "Remote data checking using provable data possession," ACM Trans. Inf. Syst. Secur., Vol. 14, No. 1, pp. 12:1–12:34,Jun. 2011.

[20] M. Sookhak, A. Gani, M. K. Khan, and R. Buyya, "Dynamic remote data auditing for securing big data storage in cloud computing," *Information Sciences*, 2015.

[21] A. Juels and B. S. Kaliski Jr, "Pors: Proofs of retrievability for large files," in Proceedings of the 14th ACM conference on Computer and communications security. Acm, 2007, pp. 584–597.

[22] S. Halevi, D. Harnik, B. Pinkas, and A. Shulman-Peleg, "Proofs of ownership in remote storage systems," in Proceedings of the 18th ACM conference on Computer and communications security. *ACM*, 2011, pp. 491–500.

[23] N. Kaaniche and M. Laurent, "A secure client-side deduplication scheme in cloud storage environments," in *New Technologies, Mobility and Security* (NTMS), *2014 6th International Conference on*, March 2014, pp. 1–7.

[24] F. S. Al-Anzi, A. A. Salman, N. K. Jacob, and J. Soni, "Towards robust, scalable and secure network storage in cloud computing," in *Digital Information and Communication Technology and it's Applications* (*DICTAP*), 2014 Fourth International Conference on. IEEE, 2014, pp. 51–55.

[25] K. D. Bowers, A. Juels, and A. Oprea, "Hail: a high-availability and integrity layer for cloud storage," in Proceedings of the 16th ACM conference on Computer and communications security. ACM, 2009, pp. 187–198.

[26] A. Bessani, M. Correia, B. Quaresma, F. Andre, and P. Sousa, "Depsky: ´dependable and secure storage in a cloud-of-clouds," *ACM Transactions on Storage (TOS)*, Vol. 9, No. 4, p. 12, 2013.

[27] D. Chen, X. Li, L. Wang, S. Khan, J. Wang, K. Zeng, and C. Cai, "Fast and scalable multi-way analysis of massive neural data," IEEE Trans. Comput, Vol. 63, 2014.

[28] A. N. Khan, M. M. Kiah, S. A. Madani, M. Ali, S. Shamshirband et al., "Incremental proxy re-encryption scheme for mobile cloud computing environment," *The Journal of Supercomputing*, Vol. 68, No. 2, pp. 624–651,2014.

[29] P. S. Kumari, P. Venkateswarlu, and M. Afzal, "A key aggregate framework with adaptable offering of information in cloud," *International Journal of Research*, Vol. 2, No. 3, pp. 5–10, 2015.

[30] C.-K. Chu, S. S. Chow, W.-G. Tzeng, J. Zhou, and R. H. Deng, "Key-aggregate cryptosystem for scalable data sharing in cloud storage," *Parallel and Distributed Systems*, IEEE Transactions on, Vol. 25, No. 2, pp. 468–477, 2014.

[31] R. A. Sana Belguith, AbderrazakJemai, "Enhancing data security in cloud computing using a lightweight cryptographic algorithm," *ICAS 2015*.

[32] A. Shamir, "Identity-based cryptosystems and signature schemes," in Advances in cryptology. Springer, 1985, pp. 47–53.

[33] A. Sahai and B. Waters, "Fuzzy identity-based encryption," in Advances in Cryptology–EUROCRYPT 2005. Springer, 2005, pp. 457–473.