iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 7 **Issue:** III **Month of publication:** March 2019

DOI: http://doi.org/10.22214/ijraset.2019.3495

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 7 Issue III, Mar 2019- Available at www.ijraset.com

A Research Paper on Data Storage and Security in Cloud Computing

Rekha Sahani[1], Sourabh kaushik[2], Rajiv Dahiya[3], Ruchika Doda[4]
[1, 2]*Student, Department of Electronics and Communication Engineering, MVSIT, Sonipat*
[3]*Head of Electronics and Communication Engineering and Electrical Engineering Department, MVSIT, Sonipat*
[4]*Project Guide, Department of Electronics and Communication Engineering , MVSIT, Sonipat*

Abstract: cloud computing has been envisioned as the next generation architecture of IT enterprise. Cloud computing moves the application software and data bases to the large data centers, where the management of the data and services may not be fully trustworthy. This poses many new security challenges which have not been fully implemented. In this paper, we mainly focus on aspects for providing security for data storage in cloud, also architecture for data storage that are implemented by other service providers vendors in cloud, key points for proving security for data storage.
Keywords: cloud computing, cloud storage techniques, security techniques, architecture, S3, API, TPA

I. INTRODUCTION

Several trends are opening up the era of Cloud Computing , which is an Internet-based development and use of computer technology. The ever cheaper and more powerful processors, together with the software as a service (SaaS) computing architecture, are transforming data centers into pools of computing service on a huge scale. The increasing network bandwidth and reliable yet flexible network connections make it even possible that users can now subscribe high quality services from data and software that reside solely on remote data centers. Moving data into the cloud offers great convenience to users since they don't have to care about the complexities of direct hardware management. The pioneer of Cloud Computing vendors, Amazon Simple Storage Service (S3) and Amazon Elastic Compute Cloud (EC2) are both well known examples. While these internet-based online services do provide huge amounts of storage space and customizable computing resources, this computing platform shift, however, is eliminating the responsibility of local machines for data maintenance at the same time. As a result, users are at the mercy of their cloud service providers for the availability and integrity of their data. Recent downtime of Amazon's S3 is such an example. Benefits of Cloud storage: No need to invest any capital on storage devices, No need for technical expert to maintain the storage, backup, replication and importantly disaster management, Allowing others to access your data will result with collaborative working style instead of individual work.

II. SERVICES IN CLOUD COMPUTING

A. SaaS

Software as a Service (SaaS) are probably the most popular form of cloud computing and are easy to use. SaaS uses the Web to deliver applications that are managed by a third-party vendor and whose interface is accessed on the clients' side. Most SaaS applications can be run directly from a Web browser, without any downloads or installations required. SaaS eliminates the need to install and run applications on individual computers. With SaaS, it's easy for enterprises to streamline their maintenance and support, because everything can be managed by vendors: applications, runtime, data, middleware, O/S, virtualization, servers, storage, and networking. Examples for SaaS are Gmail , Google Apps, Microsoft Office 365 , Google+ , face book , yahoo. Bdeveloper

B. PaaS

Platform as a Service (PaaS) deliver computational resources through a platform. What developers gain with PaaS is a framework they can build upon to develop or customize applications. PaaS makes the development, testing, and deployment of applications quick, simple, and cost-effective, eliminating the need to buy the underlying layers of hardware and software. One comparison between SaaS vs. PaaS has to do with what aspects must be managed by users, rather than providers: With PaaS, vendors still manage runtime, middleware, O/S, virtualization, servers, storage, and networking, but users manage applications and data. Examples for PaaS are AWS Elastic Beanstalk, Windows Azure, Heroku, Force.com and Google App Engine. C. *IaaS*

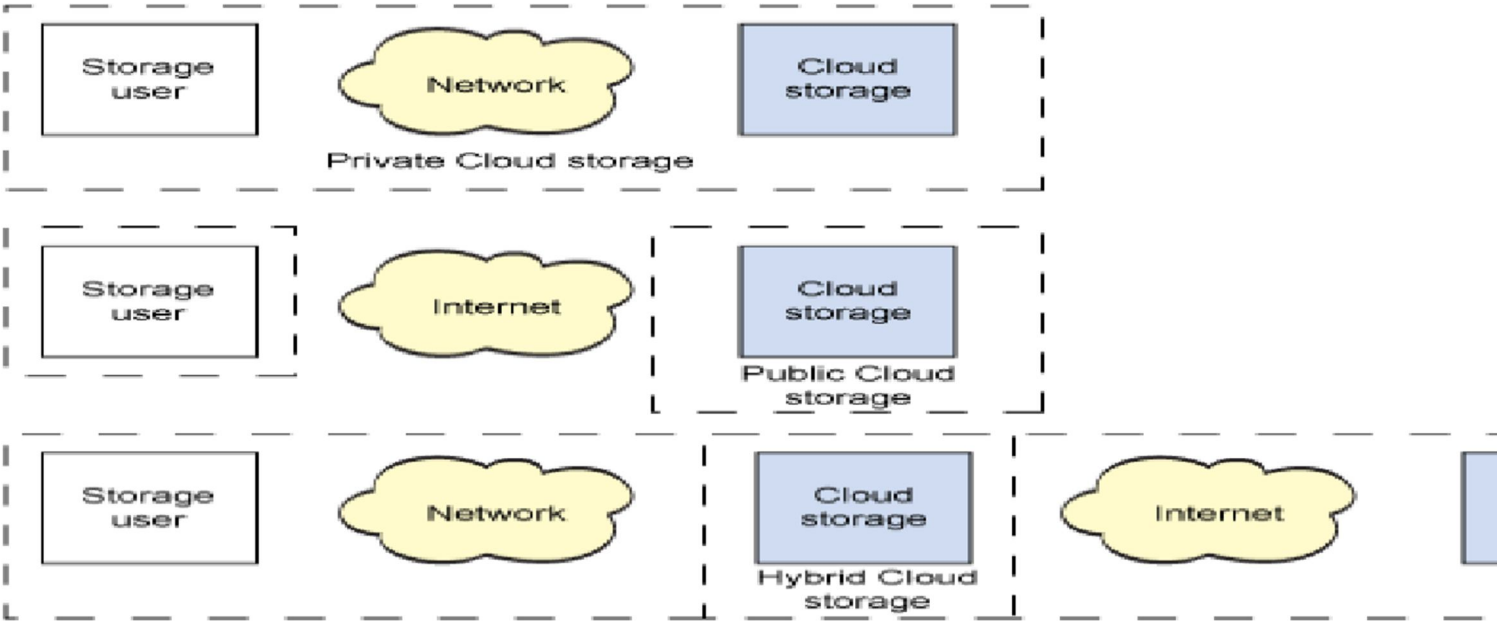
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 7 Issue III, Mar 2019- Available at www.ijraset.com

Infrastructure as a Service (IaaS) delivers computer infrastructure (such as a platform virtualization environment), storage, and networking. Instead of having to purchase software, servers, or network equipment, users can buy these as a fully outsourced service that is usually billed according to the amount of resources consumed. Basically, in exchange for a rental fee, a third party allows you to install a virtual server on their IT infrastructure. Compared to SaaS , PaaS and IaaS users are responsible for managing more: applications, data, runtime, middleware, and O/S. Vendors still manage virtualization, servers, hard drives, storage, and networking. What users gain with IaaS is infrastructure on top of which they can install any required platforms. Users are responsible for updating these if new versions are released. Examples for IaaS are Amazon EC2, Windows Azure, Rack space, Google Compute Engine

C. StaaS (Storage as a Service)

you take plunge and sign up with a cloud storage service provider, there are some things you need to know. Is cloud storage secure? How much will it cost? What services are best for SMBs? In our cloud storage services guide for beginners, we've collected our top tips and expert advice in one place so you can get answers to your most important questions. Learn about cloud backup, cloud archiving, cloud disaster recovery, and using the cloud for primary storage.

III. CLOUD STORAGE MODELS



There are models for cloud storage that allow users to maintain control over their data. Cloud storage has evolved into three categories, one of which permits the merging of two categories for a cost efficient and secure option. Public cloud storage providers, which present storage infrastructure as a leasable commodity (both in terms of long-term or short-term storage and the networking bandwidth used within the infrastructure). Private clouds use the concepts of public cloud storage but in a form that can be securely embedded within a user's firewall. Finally, hybrid cloud storage permits the two models to merge, allowing policies to define which data must be maintained privately and which can be secured within public clouds.

scale computing easier for developers. Amazon S3 provides a simple web services interface that can be used to store and retrieve

Commonly known as Storage as a Service (StaaS) ,it facilitates cloud applications to scale beyond their limited servers. StaaS allows users to store their data at remote disks and access them anytime from any place. Cloud storage systems are expected to meet several rigorous requirements for maintaining users' data and information, including high availability, reliability, performance, replication and data consistency; but because of the conflicting nature of these requirements, no one system implements all of them together.

E. Amazon S3

Amazon S3 is storage for the Internet. It is designed to make web-

any amount of data, at any time, from anywhere on the web. It gives any developer access to the same highly scalable, reliable, secure, fast, inexpensive infrastructure that Amazon uses to run its own global network of websites. The service aims to maximize benefits of scale and to pass those benefits on to developers. According to the Spring 2010 Storage magazine/Search Storage Purchasing Intentions survey, 14% of respondents said they're using cloud storage now, with the largest numbers using cloud storage for disaster recovery (6%). But 4% are using it to hold primary data from their data centers, and an equal number are using it for near line data storage. But before Fig 1: Cloud storage models

The cloud models are shown graphically in Figure 1. Examples of public cloud storage providers include Amazon

(which offers storage as a service). Examples of private cloud storage providers include IBM, Para scale, and Clever safe (which build software and/or hardware for internal clouds).

Finally, hybrid cloud providers include Egnyte, among others.

IV. STORAGE SECURITY TECHNIQUES IN CLOUD COMPUTING

Various existing techniques have been discussed in this paper. Cloud storage is regarded as a system of disseminated data centers that generally Utilizes virtualization technology and supplies interface for data storage.

A. Implicit Storage Security to Data in Online

Providing implicit storage security to data in online is more beneficial in a cloud computing. The use of a data partitioning scheme for implementing such security involving the roots of a polynomial in finite field. In this scheme data is partitioned in such way that each portion is implicitly secure and does not to be encrypted. These portions are stored on different servers on the network which are known only to the user. Reconstruction of the data requires access to each server and the knowledge as to which servers the data portions are stored. Several versions of this scheme are described, which include the implicit storage of encryption keys rather than the data and where a subset of the partition may be brought together to recreate the data.

B. Identify –Based Authentication

An identify based encryption (IBE) and decryption and identity based signature IBS schemes for IBHMCC.Resources and services are distributed across numerous consumer. So there is a chance of various security risks. Therefore authentication of users as well as services is an important requirement for cloud security. When SSH Authentication protocol (SAP) was employed to cloud, it becomes very complex. As an alternative to SAP, proposed a new authentication protocol based on identity which is based on hierarchical model with corresponding signature and encryption scheme. Identify based authentication protocol constrains sequence of steps. In step (1) the client C sends the servers a client Hello message. The message contains a fresh random number Cn.session identifier ID and c specification. In step (2) the server S responds with a server Hello message which contains new fresh random number Sn.

C. Public Auditing with Complete Data Dynamic Support

Verification of data integrity at unreliable servers is the major concern in cloud storage with public audit ability trusted entity with expertise and capabilities data owners do not posses can be delegated as an external audit party to access the risk of outsourced data when needed. It also provides a transparent yet cost effective method for data owners to gain trust in the cloud.

To accomplish, dynamic data support, the existent proof read of PDF (or) POR scheme is improved by spoofing the basic Markel Hash tree (MHT).

D. Efficient Third Party Auditing (TPA)

Cloud consumers save data in cloud server so that security as well as data storage correctness is primary concern. The data owners having huge amount of outsourced data and auditing the data correctness in a cloud environment can be difficult and expensive for data owners. To support third party auditing where user safely delegate in integrity checking tasks to third party auditors(TPA) this scheme can almost guarantee the simultaneous localization of data error(i.e. the identification of misbehaving servers). A novel and homogeneous structure is introduced to provide security to different cloud types. To achieve data storage security, BLS (Bonch-Lynn-Sachems) algorithm is used to signing the data blocks before outsourcing data into cloud. Reed Solomon technique is used for error correction and to ensure data storage correction.

E. Way of Dynamically Store Data in Cloud

Data storage is cloud may not be completely trustable because the clients did not have local copy of data stored in cloud. To address these issues proposed a new protocol system using the data reading protocol algorithm to check the data integrity services providers help the clients to check the data security by the proposed effective automatic data reading algorithm. A flexible distributed storage integrity auditing mechanism (FDSIAM), these mechanisms utilizes the homomorphism tokens, blocking erasure and unblocking factors and distributed erasure coded data.

F. Effective and Secure Storage Protocol

Current trend is users outsourcing data into service provider who have enough area for storage with lower storage cost. A secure and efficient storage protocol is proposed that guarantees the data storage confidentiality and integrity. This protocol is invented by using the construction of elliptic curve cryptography and sober sequence is used to confirm the data integrity]. Data and software process protocol step executed by cloud customers to add the privacy enforcement structure to the software and data before transferring them to the cloud. Challenge response protocol is protocol is credential so that it will not exposes the contents of the data to outsiders. Data dynamic operations are also used keep the same security assurance and also provide relief to users from the difficult of data leakage and corruptions problems.

G. Storage Security of Data

The data is secured in server based on user's choice of security method so that data is given high secure priority resources are being shared across server trouble to data security in cloud. Transmitting data over internet is dangerous due to the intruder attacks data encryption plays an important role in cloud environment. Introduced a consistent and novel structure for providing security to cloud types and implemented a secure cross platform. The proposed effective and flexible distribution scheme two-way handshakes based on token management by utilizing the homomorpic token with distributed verification of erasure coded data, our scheme achieves the integration of storage correctness insurance and data error location (i.e.) the identification of misbehaving server.

H. Secure and Dependable Storage Service

Storage service of permits consumers to the data in cloud as well as allowed to utilize the available well qualified application with no worry data storage maintained. Although cloud providers benefits, such a service gives up the self control of user's data that introduced fresh valuability hazards to cloud data correctness. The proposed a flexible distributed storage integrity auditing mechanism, utilizing the homomorphism token and distributed coded-data. The proposed design further support secure and efficient dynamic operation on outsource data including block modification, deletion and append.

I. Optimal Cloud Storage Systems

Cloud data storage which requires no effort is acquiring more popularity for individual, enteprise and institutions data backup and synchrozation. The proposed system describes, at a high level, a possible architecture for a cryptographic storage service. At its core, the architecture consists of these components- a data processor (DP) that processes data before it is sent to the cloud a data verifier (DV) that checks whether the data in the cloud has been tampers with, and a token generator (TG) that generator token which enables the cloud storage providers to retrieve segments of consumer data.

J. Process of access and store small files with storage

To support services extensively, Hadoop distributed file system server reasons are examined for small file trouble of native Hadoop distributed file system. Burden on Nane Node of HADOOP distributed file system is enforced by large amount of small files, for data placement correction are not considered prefetching mechanism is not also presented. In order to overcome these small size problems, proposed an approach that these small size problem, proposed an approach. That improves the small file efficiency on Hadoop distributed file system, in a large cluster, thousands of servers both host directly attached storage and execute user application task. By distributing storage and computation across many servers the resource a grow with demand while remaining economical at every size.

K. File storage security management

To assure the security of stored data in cloud, presented a system which utilizes distributed scheme. Proposed system consists of a master server and a set of slave server. These are not direct commutation link between clients and slave servers in the proposed model. Master server is responsible to process the client's request and at slave server chunking operation in order to provide data backup for file recovery in future. Clients file is stored in the form of tokens on main server and files were chunked on slave server for file recovery.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 7 Issue III, Mar 2019- Available at www.ijraset.com

V. CLOUD STORAGE ARCHITECTURE

Cloud storage architectures are primarily about delivery of storage on demand in a highly scalable and multitenant way. Generically (see Figure 1), cloud storage architectures consist of a front end that exports an API to access the storage. In traditional storage systems, this API is the SCSI protocol; but in the cloud, these protocols are evolving. There, you can find Web service front ends, file-based front ends, and even more traditional front ends (such as Internet SCSI, or iSCSI). Behind the front end is a layer of middleware that I call the *storage logic*. This layer implements a variety of features, such as replication and data reduction, over the traditional data-placement algorithms (with consideration for geographic placement). Finally, the back end implements the physical storage for data. This may be an internal protocol that implements specific features or a traditional back end to the physical disks.

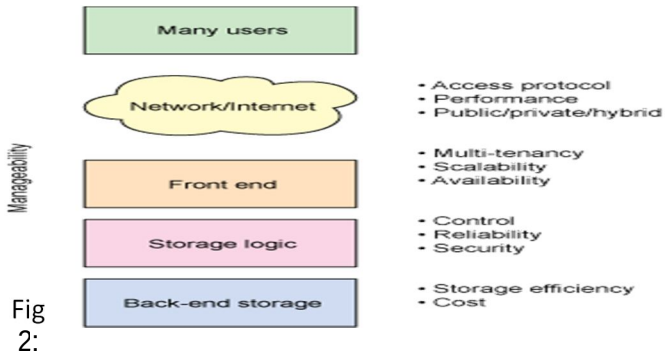


Fig 2:

Cloud storage architecture

From Figure 2, you can see some of the characteristics for current Cost Measure of the cost of the storage (commonly in dollars per gigabyte

Characteristic	**Description**
Manageability	The ability to manage a system with minimal resources
Access method	Protocol through which cloud storage is exposed
Multi-tenancy	Support for multiple users (or tenants)
Scalability	Ability to scale to meet higher demands or load in a graceful manner
Data availability	Measure of a system's uptime
Control	Ability to control a system—in particular, to configure for cost, performance, or other characteristics
Storage efficiency	Measure of how efficiently the raw storage is used
Cost	Measure of the cost of the storage (commonly in dollars per gigabyte

VI. CLOUD STORAGE API (APPLICATION PROGRAMMING INTERFACE)

A Cloud Storage Application Programming Interface (API) is a method for access to and utilization of a cloud storage system. The most common of these kinds are REST (Representational State Transfer) although there are others, which are based on SOAP (Simple Object Access Protocol). All these APIs are associated with establishing requests for service via the Internet. REST is a concept widely recognized as an approach to "quality" scalable API design. One of the most important features of REST is that it is a "stateless" architecture. This means that everything needed to complete the request to the storage cloud is contained in the request, so that a session between the requestor and the storage cloud is not required. It is very important because the Internet is highly latent (it has an unpredictable response time and it is generally not fast when compared to a local area network).REST is an approach that has very high affinity to the way the Internet works. Traditional file storage access methods that use NFS (network files system) or CIFS (Common Internet File System) do not work over the Internet, because of latency. Cloud Storage is for files, which, some refer to as objects, and others call unstructured data. Think about the files stored on your PC, like pictures, spreadsheets and documents. These have an extraordinary variability, thus unstructured. The other kind of data is block or structured data. Think data base data, data that feeds transactional system that require a certain guaranteed or low-latency performance. Cloud Storage is not for this use case. Industrial Design Centre (IDC) estimates that approximately 70% of the machine stored data in the world is unstructured, and this is also the fastest growing data type. So, Cloud Storage is storage for files that is easily accessed via the Internet. This does not mean you cannot access Cloud Storage on a private network or LAN, which may also provide access to a storage cloud by other approaches, like NFS or CIFS. It does mean that the primary and preferred access is by a REST API. REST APIs are language neutral and therefore can be leveraged very easily by developers using any development language they choose. Resources within the system may be acted on through a URL. So, an API is not a "programming language", but it is the way a programming language is used to access a storage cloud. REST APIs are also about changing the state of resource through representations of those resources. They are not about calling web service methods in a functional sense. The key differences between different Cloud

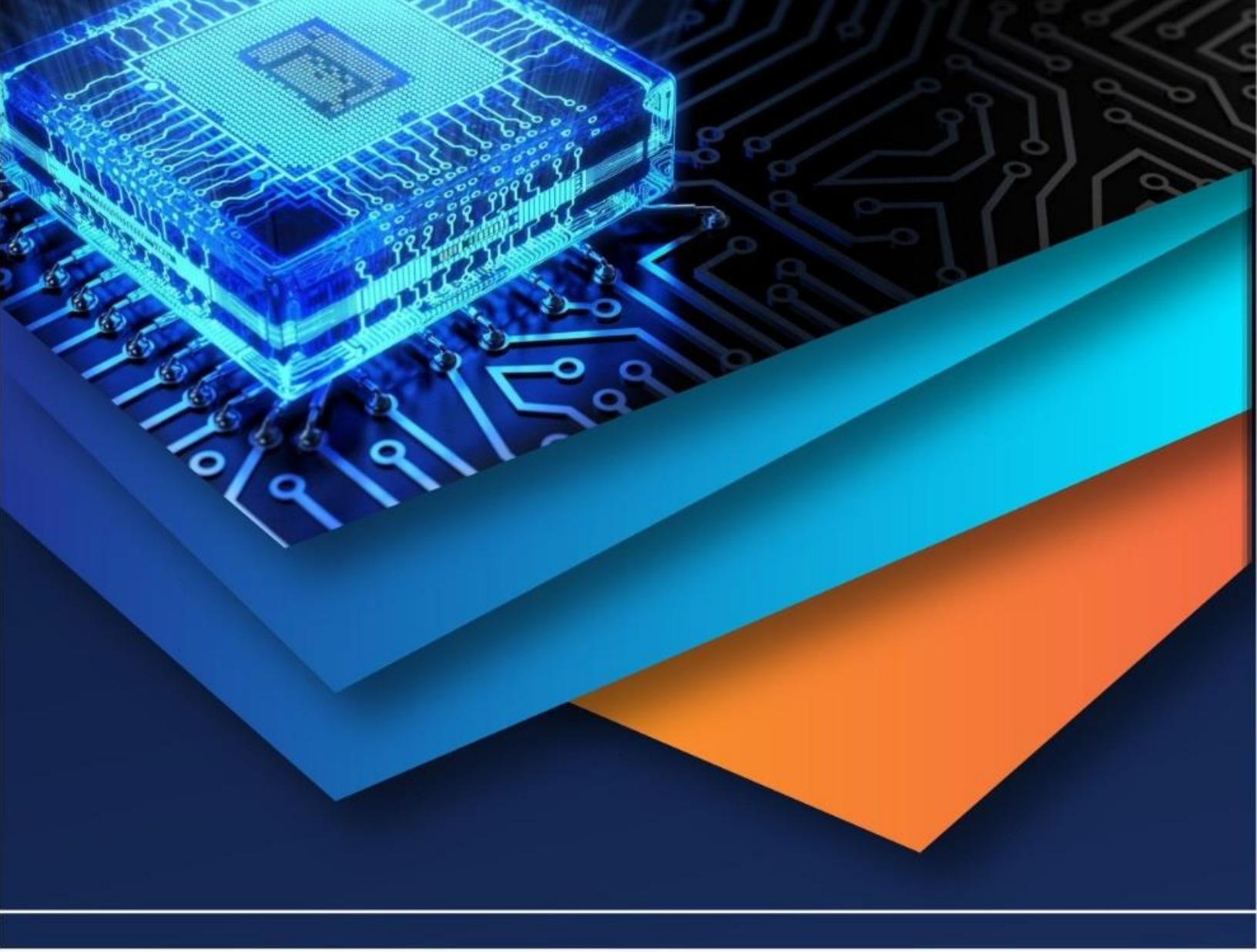
Storage APIs are the URLs defining the resources and the format of the representations. Amazon S3 APIs, Eucalyptus APIs, Rack space Cloud Files APIs, Mezeo APIs, Nivanix APIs, Simple Cloud API, along with the standards proposed by the Storage Networking Industry Association (SNIA) Cloud Storage Technical Work Group, and more.

VII. CONCLUSION

Data storage in cloud is more advantageous than traditional storage because of its availability, scalability, performance, portability and its functional requirements. We mainly focused on data storage aspects that cloud service providers are following to store the data and security aspects to be provided for that data stored in cloud. We took a look on Amazon s3 and third party auditing (TPA) mechanisms which are used for data storage and security for data in cloud.

REFERENCES

[1] http://www.ibm.com/developerworks/cloud/library/clcloudstorage/cl-cloudstorage-pdf.pdf
[2] T. Sivashakthi1, Dr. N Prabakaran A Survey on Storage Techniques in Cloud Computing" Volume3Issue12/IJETAE.
[3] R. Arokia Paul Rajan, S. Shanmugapriyaa "Evolution of Cloud Storage as Cloud Computing Infrastructure Service" IOSR Journal of Computer Engineering (IOSRJCE) ISSN: 2278-0661 Volume 1, Issue 1 (May-June 2012), PP 38-45
[4] Amazon.com, "Amazon Web Services (AWS)," Online at http://aws.amazon.com, 2008.
[5] http://www.business.att.com/enterprise/Service/hostingservices/cloud/storage/
[6] "Cloud Computing-Storage as Service" Gurudatt Kulkarni, Ramesh Sutar, Jayant Gambhir / International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622 www.ijera.com Vol. 2, Issue 1, JanFeb 2012, pp.945-950
[7] http://searchsmbstorage.techtarget.com/feature/Understandi ng-cloud-storage-services-A-guidefor-beginners
[8] E.Gorelik, "Cloud Computing Models", Massachusetts Institute of Technology Cambridge, MA,2013. Available: http://web.mit.edu/smadnick/www/wp/2013-01.pdf
[9] Gurudatt Kulkarni, Rani Waghmar, Rajnikant Palwe, Vidya Waykule, HemantBankar, KudilikKoli."Cloud



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY