ISSN: 2366-1313

Auditing Framework for Data Division, Replication to Check Data Integrity in the Cloud

Mattapalli Anil Kumar, Dr.Prasadu Peddi ,Dr.P.M. Yohan

Research scholar, Shri Jagdishprasad Jhabarmal Tibrewala University, Jhunjhunu, Rajasthan.

Assistant Professor, Dept of CSE, Shri Jagdishprasad Jhabarmal Tibrewala University, Jhunjhunu, Rajasthan.

Professor & Principal of CSI wesley Institute of Technology and Sciences, Secunderabad, Telangana.

anilkumar11282@gmail.com

ABSTRACT

Recently, Cloud Computing is treated as an effective method to provide various services at the request of many users in one physical network, using the virtualization method. Different players of Cloud Computing (CC) Environment are Cloud Service Provider (CSP) cloud consumer, cloud broker, and cloud auditor. The aim of Cloud Service Providers is to maximize the use of resources to increase the value and to also meet the requirements of the Service Level Agreement (SLA). Cloud users' goal is to meet their computing needs with the least amount of expense. Cloud broker's goal is to assist in the choice process for cloud users who use different cloud service providers. Cloud auditors' goal is to gather and analyze the evidence to determine if the cloud service provider is secure with respect to assets, ensure integrity, confidentiality, and confidentiality of the stored data, or does not. The biggest challenge for auditors was to be able to understand CC. Auditors responsible for cloud security must be well-versed in CC terminology and have an understanding of the structure and delivery mechanisms of cloud systems. This knowledge allows auditors pay greater attention to cloud security factors such as transparency and encryption, collocation scale, scope, complexity, and collocation.

KEYWORDS: Cloud Computing, Data Integrity, SLA, Auditing and CSP.

I. INTRODUCTION

Data integrity refers to the assurance that digital information is not corrupted and can still be accessed by authorized users. Data integrity is the preservation of data's stability, accuracy, and reliability. Digital forensics and data assurance are greatly aided by the integrity verification scheme. Because cloud computing is location-independent, the data integrity verification process can be a significant task.

Many types of research have been done to address the shortcomings of public auditing. This was possible by using data integrity verification methods. Data integrity is a protocol that runs between CSSs and users. Private auditing is a data integrity protocol that runs between CSSs and users. Private auditing is where the user must calculate authentication data internally. This scheme requires that the user communicates with CSS to store the data. This can lead to a more computational burden.

To communicate with users, a third-party auditor can be launched. Public auditing is where the auditor can either be a TPA or an authentic user. TPAs are used to increase audit effectiveness and decrease computation costs. It is therefore essential to develop a data integrity verification model that can be used for public auditing. Public audibility of outsourced data is possible using two types of data integrity schemes: Evidence of retrievability and Provable data possessions. Ateniese and his colleagues created the PDP scheme. The PDP scheme allows clients to send pre-processed data to untrusted servers. It also retains a small amount of meta-data. Later, the client requests that the server show that the data stored has not been modified or deleted. If data is corrupted, deleted or altered by malicious users, the PDP scheme can't be properly retrieved. PDP's behaviour is susceptible to financial scratch, data loss and trust loss. PoP is another data integrity scheme that is similar to PDP. It offers an additional advantage over PDP. The PoP will retrieve corrupted or lost data by using error-correcting codes. Data integrity can be broken down into probabilistic and deterministic groups. The deterministic scheme is required to access the entire file. This scheme is not recommended to large files because it takes longer for integrity verification. The probabilistic scheme dynamically retrieves data blocks to verify integrity. This scheme is more suitable for large files, which require less computation time.

Auditing

An auditor's job is to provide an objective opinion about the facts and evidence that a company controls to meet a particular objective, criteria, requirement or other requirements. An auditor may also provide an opinion on whether the control was effective for a specific period. This is also true for cloud compliance audits. Auditors will ask for evidence that controls have been enabled when cloud compliance is required. Auditors will request evidence of controls being

ISSN: 2366-1313

enabled. The cloud auditor will be able to give an opinion as to whether controls are in place and, if so, for how long.

Auditors use variety methods to gather evidence. These include observation, confirmation, analysis and confirmation. You can combine these procedures to get evidence that will enable auditors form an opinion on the service being audited. These are some examples of tests that were done for each IT area. This is by no means an exhaustive list.

Challenges in cloud computing

An IT security audit is an examination of IT organizations' checks and balances. Auditors evaluate, test and evaluate the organization's systems and practices to determine if they can protect information assets, maintain data integrity, and achieve their business goals. To achieve these goals, IT security auditors need data from both internal and external sources.

Cloud computing can also pose security risks. Cloud infrastructure is the outcome of three-way negotiations between service provider, end users and cloud service providers (CSPs). This allows for productivity to be maintained, while still maintaining some security. CSPs should ensure data security and allow clients access to any Internet service. CSPs must ensure that cloud computing companies meet clients' business objectives, goals, and future requirements.

II. Proposed Architecture

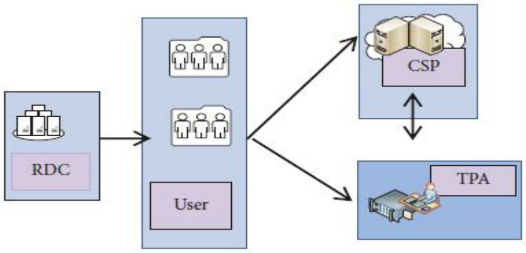


Figure 1: Data integrity Cloud sharing model.

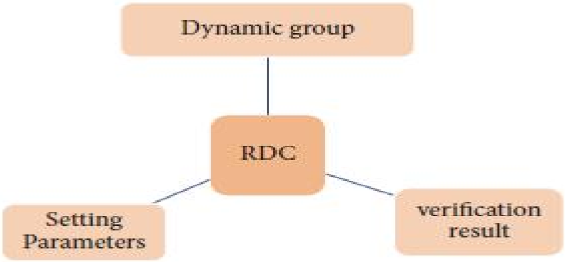


Figure 2: Right Distribution Center for Cloud Auditing.

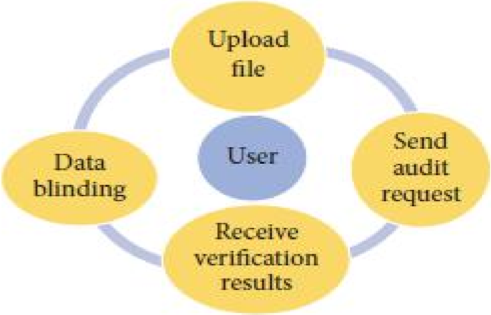


Figure 3: Security in terms of accessing the data using the encrypted keys for Data integrity.

III. Proposed Contributions

Two main entities make up a cloud environment: the CSP is the service provider and the CUs are the service users. These two entities interact using a variety of technologies, including databases, networks and virtualization. The CUs not only use the CSPs' services, but often also outsource sensitive data to the cloud servers. Security issues arise from the different stored in-house. Research shows that encryption and data integrity are top concerns for most CUs. Researchers have suggested the use of a trusted TPA to address the issues of confidentiality and data integrity. The TPA is able to perform many resource-intensive tasks such as managing encryption keys and checking out data integrity. For CUs the TPA helps reduce the computational burden for CUs but it is still possible that malicious insiders are involved at the TPA. There is a need to find a way to detect malicious activity could result in the sharing/transmission encryption keys between CSPs and CU. It is not realistic to assume that all CSPs can be trusted. CSPs can conceal data loss-leakage incidents from CUs in order to keep their good reputation. A security incident can be caused by Byzantine failures or server conspiring attacks, as well as malicious data alteration. There are two types of malicious insiders. The first is responsible for debasing the CUs data files stored on individual servers. A malicious insider can modify and read the data on the server.

A CU could become a malicious entity by intentionally violating the terms of the SLA. A CU can sublet services to third-party individuals or organizations once it has received services from a CSP. This can pose serious security risks and raise management questions. Subletting cloud services to third parties can also slow down service delivery from the CSP website. These malicious activities can have a serious impact on the reputation of a CSP and could result in substantial business loss. To detect potential violations of SLA, it is imperative that CUs are periodically audited. A trusted TPA is the best choice for an impartial and fair audit of CUs.

To authenticate all stakeholders, we develop a triangular data PMM. This model aims at ensuring the integrity of the CU’s cloud-based data, which can be retrieved anytime. Recent research has mainly focused on the reliability of CSP in terms of security and privacy measures, as well as compliance with SLA. The reliability of the CU or the TPA has not been evaluated. Our model assesses the integrity of the CU in terms of its compliance with the rules set forth by the CSP in their SLA. The TPA also audits the services rendered to the CU, and as such ensures that the integrity of the TPA (i.e. the TPA does not disclose the contents of the CU's from the information obtained during the auditing).

IV. Understanding Security Audit Frameworks for Different Cloud Service Providers

Cloud storage services enable users to store data online and avoid the need for local storage or maintenance. To ensure data in cloud storage is secure, there are many data integrity auditing techniques. To perform data integrity auditing, most, if not all, of the available methods require the user to use his private key. A hardware token, such as a smart card or USB token, is required to activate the private key. A hardware token (e.g., smart card or USB token) is required to activate his private key. It stores the user's private key and password. Many of the current data integrity auditing systems will fail if the hardware token is lost or forgotten. This problem can be solved by a new paradigm in data integrity auditing, which does not require private key storage.

Opportunities and challenges of Cloud Auditors

These evaluations use only current Cloud services and configuration settings as set by service providers. The SOA principle makes cloud provisioning transparent for service users.

Second, it is difficult for researchers to determine impact of Cloud infrastructure resource management on service performance. This is because there is not enough control over the infrastructure configuration. Many service users wish to evaluate the performance of new services to assist them in making their decisions.

Existing works that are focused on predictability and general performance evaluation offer some promising solutions. Research Clouds can be used by researchers to verify validity of performance assessment based on measurement. This allows them to access information about service implementations and control the configuration of service deployments.

ISSN: 2366-1313

V. Algorithms and Method of Implementation

This scheme can be used for data integrity auditing. We use a linear sketch to confirm identity. It includes error correction and coding. A new signature scheme was also designed that supports block-less verification and is compatible with the linear sketch. According to security analysis and performance analysis, our proposed scheme is efficient and secure.

Implementation

After the data sample has been copied, it can be accessed via all CSP cloud services. The cloud file can then be accessed and credentials provided to verify data integrity. You can either use Python programming or R programming to get performance results.

Preliminaries and notation

1. F and F[i] are a file and the ith block of data for F.

2. H (*): is a hash function.

3. ph (*): Euler's totient function.

4. m(*): a Pseudo-Random Function (PRF) which maps: m: {0, 1k x {0, 1}|5. m (*): A pseudo-random function (PRF), which maps: m, 0 x 1k x 10.|5. m (*): is a pseudo-random function (PRF), that maps: m: 0, 1k, x m: 0, 1, 1.}

5. s (*) : a Pseudo-Random Permutation (PRP) which maps: s : {0, 1k x {0, 1 ... n} -- {0, 1 ... n}.|6. s (*). A pseudo-random permutation (PRP), which maps: S : 0 x 1k x 1... N -- 0... n.|6. s}

6. p and Q: Two different odd prime numbers of equal length.

7. J : Multiplication two prime numbers p or q.

8. r1,r2 : Random numbers taken from the Galois field

9. T and Ti: All block tags and ith tags of T.

10. R is the number of blocks needed for each challenge.

11. kt: Encryption key for tag.

12. kd: Decryption key for tag.

13. Z and Z*: Group on integer numbers.

14. c: Number of blocks that are available for challenge operation.

15. r: Number of deleted blocks in total file blocks.

VI. Auditing Framework for Data Division and Reduplication

Auditing's Role

Auditing is essential for organizing, planning, and delivering support. It evaluates, monitors, and assesses third-party providers' and service provider performance (IT Governance 2012). It should use auditing systems to verify confidentiality, data integrity and availability, authentication, reliability, and security. It should be more responsible and contribute to key strategic areas such as customer relations, cost reductions and revenue maximization. Business management includes auditing. Auditing should be about adding value by supporting strategic initiatives and providing valuable insight into the company. Auditing should be actively involved in monitoring, evaluation and improvement of regulatory compliance.

Cloud Auditing

Cloud storage and processing of data does not require a lot of resources and local systems. Cloud storage is an efficient way for users to store their data. Users also prefer cloud storage because they can store as many data as they want without restriction. The cloud transfers the application software and databases into large, centralized data centers. However, this can make it less trustworthy because of security concerns like old IP addresses, etc.

Auditing is the process of evaluating and collecting evidence to determine if a computer system's safety, efficiency, security, and data integrity. To verify and maintain data integrity, an auditor must audit the data stored on the cloud. There are many ways to ensure data security.

PPPA for Secure Cloud Storage

Fog empowers to own right to use up to several services. Costumer's loo get entry to sensational supplies on-demand along with getting pleasure from purposes along with providers. Web mustiness sees journal integrity in addition to the certificate. This can be a shocking key outcome. As mentioned above, the effect will be handled with semipublic accountancy tubercle for which concerns mediator examiners (TPA). This technique promotes record kinetics in addition to real-world scrutinize skills. Management consulting is employed to trace track record differences, false positives, furthermore insertions. The auditing procedure is supported by the system, which includes data dynamics, public audibility, and multiple TPA. HARS is used to create ring signatures. Merkle Hash Trees are used to improve block-level authentication. Through the Batch auditing process, the TPA can perform audits simultaneously for multiple users.

ISSN: 2366-1313

Dynamic data operations

The important dynamic data operations are divided into

Data Insertion (DI),

Data Deletion (DD) and

Data Modification (DM)

Data Insertion (DI)

Data insertion operation inserts a new block after the specified position of a file F. Mostly it does not alter logic structure of customers data. Suppose the client wants to insert a new block, this protocol supports insertion of a new block b_{new} after a given block b_i into a file F. As a first step, user initially generates new data information (VER_{new}, TS_{new}) for new block b_{new} contains new version and timestamp of a new file block to be inserted. Then client sends the update request to $update_{ins}$ (F, DI, i, VER_{new}, TS_{new}) to Trusted Third Party Auditor (TTPA)

Algorithm Steps:

```
begin
    Update record in ITS
    begin
        user generates new data information (VERnew, TSnew) for the new block bnew.
        sends update request updateins(F, DI, i, VERnew, TSnew) TTPA
    end
    TTPA performs the following after receiving updateins request
    begin
        TTPA finds the last record in ITS and inserts the new one after it
        update both VERnew, TSnew, increment pointer by 1
    end
    update stored data in cloud
    begin
        user generates new signature Tnew for the new block bnew
        sends update request updateins(F, DI, i, bnew (VERnew, TSnew), Tnew->CSP
        CSP generates new version of file Fnew and tag set Tnew = Fnew||n||sigSsk(Fnew||n)
    end
end
```

Data Deletion (DD)

Data deletion operation refers to deletion of a specified block and requires moving all the blocks after deletion. Suppose the client wants to delete a block, this protocol supports deletion of a particular block bi from the file F. As a first step, user initially generates request for deletion as update del (F, DD, i) to Trusted Third Party Auditor (TTPA). The request consists of a file; DD refers to data deletion and a block number to be deleted. TTPA carries out the update request by deleting a specified block upon receiving the request from clients.

Algorithm Data Deletion (DD)

```
begin
  update record in ITS
  begin
    user sends update request update del (F, DD, i) -> TTPA
    TTPA deletes ith record in ITS
    decrement pointer by 1
  end
  update stored data in cloud
  begin
    user sends update del (F, DD, i) o CSP
    CSP gets new file version Fnew and tag set 7new = Fnew||n||sigSsk(Fnew||n)
  end
```

Data Modification (DM)

In cloud data storage data modification is the most important one among the three operations. This protocol supports the replacing of a specified block with a new block. This is the most frequently used operation in cloud storage. Assume that the client wants to modify the i^th file block b_i in to b_{newc}. Data modification operation replaces the specified block of bi to b_{new}^c.

As a first step, client generates the new information (VER_{new}^c, TS_{new}^c) for the particular new block b_{new}^c. Then it sends the update request $update_{mod}(F, DM, i, VER_{new}^c, TS_{new}^c)$ to the auditor (TTPA).

The server performs the update operation on their side upon receiving the request from user.

i) replaces the old file with the new file version F_{new}^c

ii) Generates tag set F_{new}^{c} for the new block b_{new}^{c} as F_{new}^{c} ||n||sig$_{Ssk}$(F_{new}^{c}|| n). Then server sends the update request modification proof to the client.

Algorithm: Data Modification

```
begin
    user generates new ($VER_{new}^{c}$, $TS_{new}^{c}$) information for the block $b_{new}^{c}$
    update record in ITS
    begin
        Sends update request updatemod(F, DM, i, $VER_{new}^{c}$, $TS_{new}^{c}$) for $b_{new}$cĺ TTPA
    end
    TTPA performs the following after receiving update$_{mod}$ request
    begin
        TTPA finds the appropriate record in ITS
        replace (VER, TS) with $VER_{new}^{c}$, $TS_{new}^{c}$
    end
    update stored data in cloud
    begin
        user generates new signature $V_{new}^{c}$ for the new block $b_{new}^{c}$
        sends update modification request updatemod(F, DM, i,$b_{new}$($VER_{new}$, $TS_{new}$), $b_{new}$ ($VER_{new}^{c}$, $TS_{new}^{c}$), $V_{new}^{c}$ o CSP
        CSP generates new version of file $F_{new}^{c}$ and tag set $TS_{new}^{c}$ = $F_{new}^{c}$ ||n||sig$_{Ssk}$($F_{new}^{c}$|| n) after replacing the old block with a new one
    end
end
```

Proposed a secure and effective cloud data integrity verification scheme with all the implementation is similar to the protocols of RSA DAP, ECC-DAP and the Secret Sharing and Public Verifiable Dynamic Protocol. The usage of the multiple TPA is the only difference between them where one acts as master and other as slave TPAs. The master TPA uses the entire slave TPAs to detect the data corruption or the misbehaving servers effectively. If the master TPA crashes any one of the Slave TPA act as the master

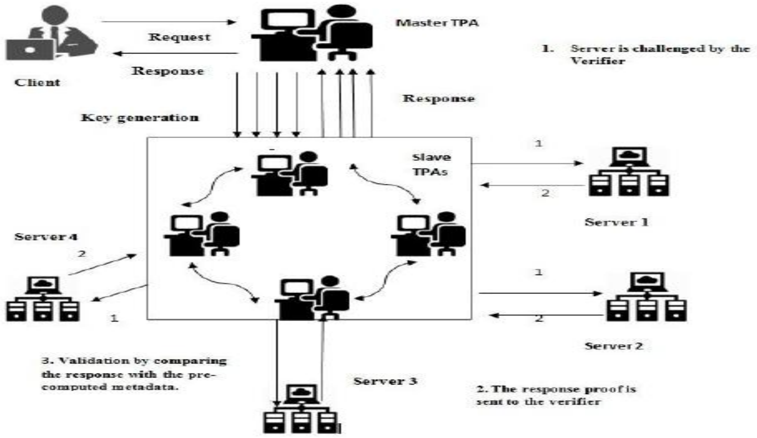

Client
Request
Response
Master TPA
1. Server is challenged by the Verifier
Response
Key generation
Slave TPAs
1
2
Server 1
Server 4
2
1
1
2
Server 2
3. Validation by comparing the response with the pre-computed metadata.
Server 3
2. The response proof is sent to the verifier

Figure 4: Architecture of the proficient distribution verification protocol

All TPA are atomic in nature and ceased from communicating within. All operations are consistent where all perform their assigns operations and take over operation when peer TPA fails.

VII. RESULTS AND ANALYSIS

Google App Engine Study Guide

A. Round-trip Time (RTT).

b. Throughput of the network

Round-trip time is the time it takes for data to travel from source and destination to the destination and back to the user. This is an important metric for cloud computing as it allows for easier comparisons between latencies of Google App Engine compared to traditional web servers. RTT can easily be measured in seconds.

Network throughput is the data transfer rate through a network connection during a given time period. It also measures the system's bandwidth. It can show the bandwidth difference between Google App Engine and traditional web servers in this example. You can measure network throughput in kB/sec.

Two parameters will govern the experiment: data size and number requested by Planet Lab nodes. There are three sizes of data available: small (12kB), medium (355kB), and large (1MB). Planet Lab nodes can accept between 1 to 100 requests.

Table 1: Network-based Measurements on Cloud Computing Services

Image Size	12kB			350Kb			1MB		
# of req/Planetlab	1	10	100	1	10	100	1	10	100

node									
RTT for GAE(sec)	1	5	47	1	10	40	1	15	43
RTT for TWS(sec)	1	13	62	1	50	510	1	120	1380

RTT results for GAE and traditional web host, data are collected from Network-based Measurements on Cloud Computing Services

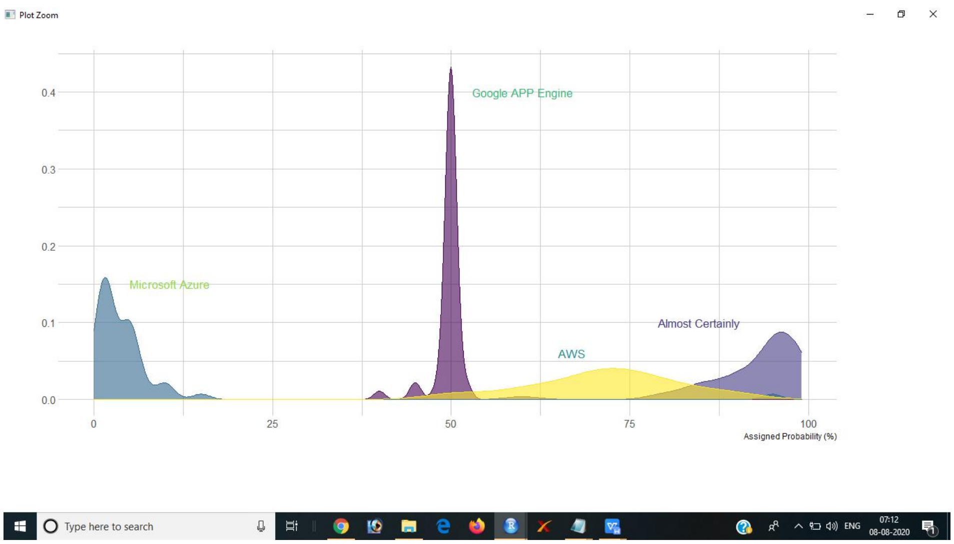


Figure 5: Comparison of the different cloud services used in the models.

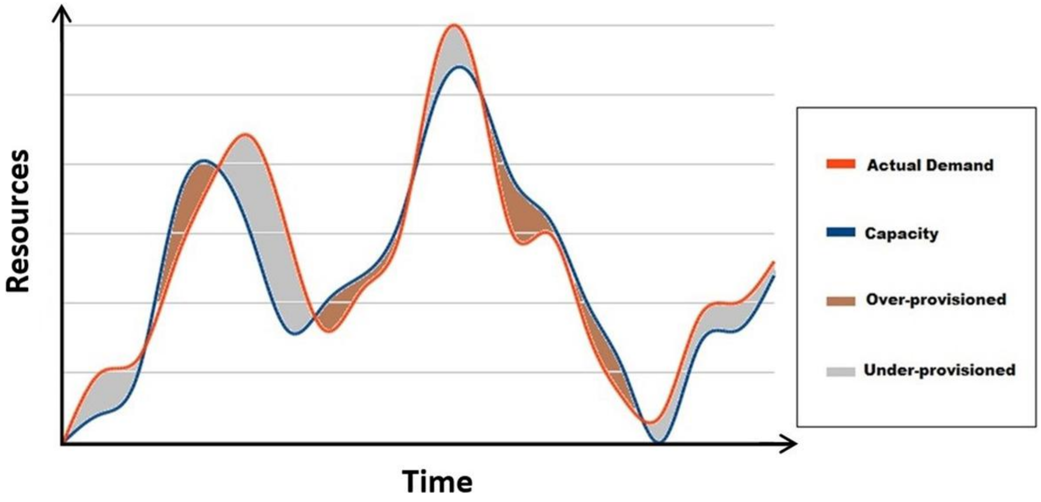
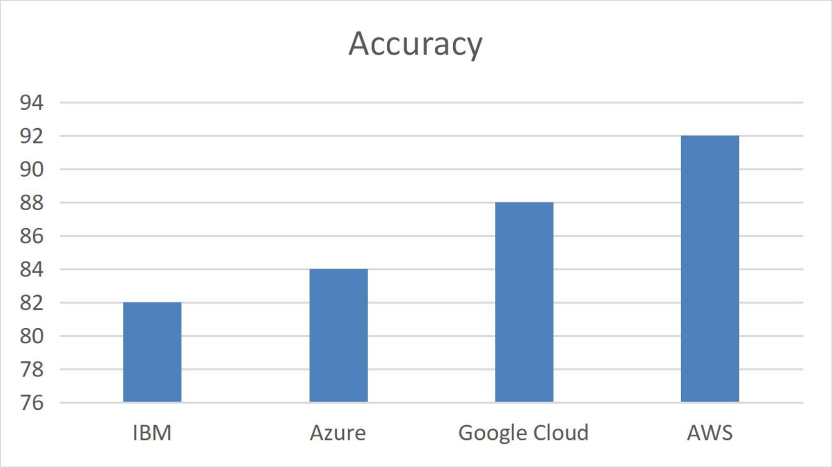


Figure 6: Scalability analysis comparisons of cloud-based software services.



CONCLUSION

In this paper we study and compare different Third-Party Auditors. There are two types of cloud computing auditing public and private. Cloud computing has two types of auditing: public and private. Third-Party Auditors dominate public auditing. To ensure data integrity and security, cloud users need a TPA. There are many organizations that can serve as TPAs. However, it is difficult to identify which TPA will offer better service and be more reliable. This algorithm helps to select legitimate Third Party auditors (TPA). This study also focused on the integrity of the Third Party Auditor. Cloud computing does not guarantee data integrity, but it is important that you know if the TPA is checking this. This paper explains the process of checking integrity of Third-Party Auditor.

REFERENCES

1. Avvari Sirisha & G. Geetha Kumari, (2010), “API access control in cloud using the Role Based Access Control Model”, ISSN: 2325-5919, PP: 135-137.
2. Amit Hendre & Karuna Pande Joshi, (2015), “A Semantic Approach to Cloud Security and Compliance”, ISSN: 2159-6182,PP: 1081-1084.
3. Bo Tang and Ravi Sandhu (2014). “Extending openstack access control with domain trust”.
4. Chow S M. (2016) “A Framework of Multi-Authority Attribute-Based Encryption with Outsourcing and Revocation”. ACM, 2016: pp:215–226.
5. Fan Y, Liu Z. (2017) “Verifiable Attribute-Based Multi-keyword Search over Encrypted Cloud Data in Multi-owner Setting” Second International Conference on Data Science in Cyberspace, IEEE, 2017: pp: 441–449.

6. Guo Z, Zhang H, Sun C, et al. (2018) “Secure multi-keyword ranked search over encrypted cloud data for multiple data owners”. Journal of Systems & Software, 2018, 137:380–395.
7. Kim-Kwang Raymond Choo, 2014, “A Cloud Security Risk-Management Strategy”, ISSN: 2325-6095, Volume: 1 , Issue: 2 , PP: 52-56.
8. Li, M., Yu, S., Zheng, Y., Ren, K., & Lou, W. (2013). Scalable and secure sharing of attributes health files in cloud computing using attribute-based encryption. IEEE Trans. Parallel Cloud Systems, 24(1), 131-143.
9. SushmitaRuj, (2014), “Attribute based access control in clouds: A survey”, ISSN: 2165-0608, 2014 International Conference on Signal Processing and Communications (SPCOM), PP: 1-6.
10. Prasadu Peddi (2016), Comparative study on cloud optimized resource and prediction using machine learning algorithm, ISSN: 2455-6300, volume 1, issue 3, pp: 88-94.