A comparative study on efficient cloud security, services, simulators, load balancing, resource scheduling and storage mechanisms

Balajee R M1 [0000-0003-2928-9509], Hitesh Mohapatra2 [0000-0001-8100-4860] and Venkatesh K3

1 Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram-522502, Guntur, Andhra Pradesh, India, balajee.rm@gmail.com
2 Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram-522502, Guntur, Andhra Pradesh, India
3 Department of Computer Science and Engineering, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Avadi, Chennai, Tamil Nadu, India

Abstract. There are applications in software that can process the user or customer records in hospitals, education sector, government sector, social media and mail communication, etc., which require accessing a single or set of records in a database consisting of millions of records. A basic requirement of those applications is that, the need to access a few data required in a secured manner from the data sets which are very large but simple in structure. Cloud computing provides the computing requirements for these kinds of the new generation of applications involving very large data sets that cannot possibly be handled efficiently without understanding the available cloud computing infrastructures. The objective is to find the cloud security, services, simulators, load balancing, resource scheduling and storage mechanisms which can best suite the need of the cloud environment chosen in an efficient way. This survey would act as a catalyst for the design of cloud environment over the very large data set in the future.

Keywords: Cloud Security, Cloud Services, Load Distribution, Resource Scheduling, Storage Mechanism, Cloud Simulators, Cloud Computing Environment.

1 Introduction

Systematic data plays a major role in minimizing the human effects for calculation, analysis, ranking, sharing emotions, future prediction, purchasing, billing, etc. In a broad classification, there are two challenges in handling these data, (i) The data size and type variation and (ii) The accessibility of data. Different data sizes have happened in network and storage areas because data creation starts from simple tweets to massive video storage clouds like YouTube. Usage of different office packages, video, and audio capture for day-to-day life leads to data type variation besides data generated by smart devices used by the people today. The customers, users, viewers, and administrative persons need to access the required data anywhere and anytime, when it is required. This is the reason for the initialization of great technological advancement in the name of cloud computing. This technical advancement plays a major role in commercial marketing, logical analysis, and individual user communications like mail, Facebook, etc. In the year of 2015, big data starts to have its great exponential growth because of social media, using smart sensor-based electronic devices, medical science, financial sectors, and telecommunication industries [29]. In the year 2015 alone 5,000,000,000,000,000,000 bytes are accumulated in storage and this is the massive growth while comparing previous data storage consideration. There is a sure problem based on increasing data size exponentially and it needs to be addressed with a clear
solution. The researchers had their view on the aspects of (i) Compressing data size, (ii) Efficient data retrieving algorithm and (iii) Selection of files to delete from storage.

2 Literature Survey

Every popular cloud had its service engine to do the job [8]. The cloud service provides the space and software support that is required to store and handle the client's data in the cloud. Amazon had Amazon Web Services (AWS) which allows any company to save the information and to run its software on Amazon's server. The computation or process is the major task of the system and because of its complex nature, it takes more energy to do the job [9, 41]. This computation can also be done on the server to reduce the burden of the client system. Different cloud simulators will offer the simulation environment to predict the behavior of the virtual machine[30], client request, and data center [10] and major simulator comparison is shown in below table.

Table 1: Cloud Simulators Comparison

| Simulator              | Programming Language | Networking | Availability |
|------------------------|----------------------|------------|--------------|
| CloudSim               | Java                 | Limited    | Open Source  |
| GCD (Green Cloud)      | C++ / OTCL           | Full       | Open Source  |
| Cloud Analyst          | Java                 | Limited    | Open Source  |
| Network Cloud          | Java                 | Full       | Open Source  |
| MDC                    | C++ / Java           | Limited    | Commercial   |

The cloud services also providing a major role in the education sector [11]. In general education institution wants to manage a large amount of data in terms of faculty database and student database. T-broker is the mechanism proposed in the manuscript [23]. It is a trust-aware service brokering system for efficient matching multiple cloud services to satisfy various user requests. Various cloud service providers are offering different services and support. [12] Three major cloud service providers are takers here for a survey.

Table 2: Cloud Service and Supportability Comparison

| S. No. | Features   | Google Drive                                                                 | One Drive                                                                 | DropBox                                                                 |
|--------|------------|------------------------------------------------------------------------------|---------------------------------------------------------------------------|-------------------------------------------------------------------------|
| 1.     | Storage Space | The user can have up to 15GB of free data storage and if required, the user can purchase further storage space. | The user can have up to 7GB of free data storage and if required, the user can purchase further storage space. | As per the user requirement, the user can purchase the storage plan. |
| 2.     | Platform Supported | ios, Android, Mac, Windows.                                                  | ios, Android, Windows, Mac, Windows Phone.                               | ios, Android, Blackberry, Windows, Mac, Linux.                         |
| 3.     | File Size Limit       | It supports file size up to 10 GB.                                           | It gives a file size limit up to 2 GB.                                   | It provides file size Unlimited.                                       |
| 4.     | Supported File Format | It supports Word, Excel, PowerPoint, PDF,                                   | It supports Microsoft Word, Excel, PowerPoint, OneNote’s, photos, and    | It supports photos and Videos.                                         |
While speaking about the cloud, cryptography is one of the very important aspects of it. The cryptography is all about encryption and decryption. If the image is been encrypted, it should be in good quality, so that it can be decrypted in an easier way to get the original. [13] The research proved that the best quality of the image can be obtained with lower Mean Square Error (MSE) and it results in higher Peak Signal to Noise Ratio (PSNR). Here sample images are taken for testing to produce the following output in MATLAB.

Table 3: Image Quality for Various Encryptions

| Size   | MSE     | PSNR   | SC      | MD   | NAE   |
|--------|---------|--------|---------|------|-------|
| 177x177| 350.0429| 22.6896| 1.0828  | 133  | 0.1094|
| 280x210| 432.9834| 21.7661| 0.0412  | 154  | 0.0850|
| 533x400| 740.1536| 19.4376| 0.9925  | 183  | 0.0523|
| 800x600| 284.2120| 23.5944| 1.0439  | 139  | 0.0798|
| 900x720| 405.6102| 22.0497| 0.9897  | 117  | 0.1088|
| 1024x768| 730.4694| 19.4948| 0.9908  | 198  | 0.0626|

In cloud large amount of data's are stored and so security to protect those data's are very much important. Considering the security of cloud storage [19], few come in front, SSL (Secure Socket Layer), Encryption, Intrusion Detection System, Multi-Tenancy based Access Control, etc. The security in the cloud can also be provided without the encryption process [14]. The security is achieved by the internal partitioning of data in cloud servers. When speaking about the encryption process for improving the cloud security, new encryption techniques using ASCII values are specified [15]. The evaluation of big data security through the Hadoop incremental security model is the work done by the author [22] over there. Hadoop, Map Reduce, and No SQL are the major big data technology in recent times. Google’s Fusion Table is also been used to confirm the effectiveness of the research. [21] This manuscript has presented a comparative study on different models in various scenarios to plot a comparison table as below:
Table 4: Comparison of Different Models on Data Security in Cloud

| Scenario        | Trust and Risk | Multi-Domain | MTB-AC       |
|-----------------|----------------|--------------|--------------|
| Confidentiality | Very Good      | High         | High         |
| Integrity       | High           | High         | Very Good    |
| Trust           | High           | High         | Very Good    |
| Risk            | Yes            | No           | No           |

MTBAC – A Mutual Trust-Based Access Control

In the development of the cloud system, a large number of vendors can visit their users in the same platform directing their focus on the software rather than the underlying framework [21, 35]. The Metadata management normally solves the problem of renaming directory [16]. The author proposes the new mechanism for Metadata management for redirecting the directory path.

Figure 1: Block Diagram of Equally Spread Current Execution Process

[20] The manuscript had given the main focus to small networks of a particular system with thousands of nodes. This brings out the requirement for the distribution, analysis of storage over the data in the cloud virtualized access and bringing the web services which are scalable on the wider application of cloud. Load balancing plays a major role in cloud computing to distribute the resources and such an area required effective mechanism [17, 34]. When the number of searches increases then by default the server performance decreases due to overload [18]. This scenario can be handled by a two-layered index structure for query processing. [24] Each algorithm had its advantage and disadvantage. Here, two of the existing algorithms are committed which are throttled and equally spread to form the new hybrid algorithm. It had been proven that this hybrid algorithm is most effective than existing in terms of load balancing.
Online health care system is already been supported by various application in mobile [25] such as Practo – It is used to book an appointment with doctors in the local area, Thyrocare – This is the useful application for thyroid patients and E-Wireless – It is the application used to save details of a particular hospital only. Reducing the processing and storing time in the cloud using Hadoop is the concept here [26]. Hadoop receives its input from scoop and flume. Scoop loads the input from the database to Hadoop and flume loads the input from server files to Hadoop. Two problems are discussed here, storage and processing speed. The storage problem is sorted out by Hadoop and performance problem is sorted out by map reduction algorithm. [27] Finding out emotions and matching it is a huge process.

In cloud, scheduling means the order in which the set of tasks to be executed [28]. In the Cloud computing environment, there is a large number of resources are available but the main challenge is to allocate the tasks in such a way so that resources can be maximum utilized. Here there is one comparison table (Table 2.5) is given below that describes about few resource allocation schemes.

![Figure 2: Master / Slave Architecture](image)

![Figure 3: Various Cloud Security Challenges](image)
Table 5: Comparison table of Resource Allocation in Cloud Environment

| Approaches                                           | Advantages                                                                 | Disadvantages                                      | Parameter |
|------------------------------------------------------|-----------------------------------------------------------------------------|----------------------------------------------------|-----------|
| In the Cloud Computing Environment, an Approach of Genetic Algorithm | can be used in complex objective function, Bigger search space can be handled, and being trapped into local optimal solution can be avoided | Still not well efficient technique.                | Makespan |
| In Load Balancing Ant Colony Optimization, an Approach of Cloud Task Scheduling | The size of the tasks doesn't matter.                                       | Should not contain any precedence between the task. | Makespan |
| In the Cloud Computing Environment, an Approach of Improved GEP Algorithm | Reduce the optimization time and good convergence.                         | Not consider the dynamic multi-objective.          | Completion Time. |
| Enhanced Genetic Algorithm based Task Scheduling in Cloud Computing. | Better task optimization and give high performance.                        | Analysis can be made only by using a limited number of jobs & resources. | Execution time |
| In Environment of Genetic Algorithm and Ant Colony Optimization in Cloud Computing, an Approach of Task Scheduling Algorithm | Improve the searching efficiency of algorithm and reducing the slow convergence. | Become a quite complex technique.                  | Execution time. |

The waste space management in the cloud got introduced strongly in this manuscript [31]. The motive of the researchers here is to do two things,

(i) To perform secure deduplication of encrypted data.
(ii) To perform public integrity auditing of data.

From here on, this process will be called as Split-Way-Approach on Efficient Cloud Storage (SWA-ECS). The SWA-ECS performs with the core of challenge-response protocols based on BLS signature with a homomorphic linear authenticator [32]. The manuscript utilizes an auditor from a third party for performing public audit with the motive of supporting low-powered clients. The SWA-ECS consists of the entities, (i) Client / User, (ii) Cloud Storage Service (CSS) and (iii) Third-Party Auditor (TPA)

Figure 4: SWA-ECS Architecture Diagram
(i) **Client / User**: The client encrypts the data and sends it to the cloud storage. In the way of doing this, the client needs to verify the integrity of the data going to be uploaded. To do this, the client delegates integrity auditing to the TPA. The data integrity needs to be checked because of the accessibility of multiple clients over the cloud storage. There is a possibility of uploading the same file in the cloud from the same user or different user.

(ii) **Cloud Storage Server (CSS)**: It offers storage space for client files. When the file is being uploaded by the client and then the Cloud Storage Server provides the mechanism to do the deduplication process. The CSS uses the PoW protocol to ensure the file is already owned by the client or not. It should also respond with a proof for the request from TPA.

(iii) **Third-Party Auditor (TPA)**: To support data integrity, two schemes PDP and POR had been introduced already. Ateniese et al. [33] introduced PDP for ensuring that the cloud storage providers possess the files without retrieving or downloading the entire data. It is a challenge-response protocol between the verifier (a client or TPA) and the proof provider (it is a cloud). Compared to PDP, POR not only ensures that the cloud servers possess the target files, but also guarantees their full recovery [1]. Since then, several POR schemes [2], [5], [6], and PDP schemes [3], [4], [7] have been proposed. Here TPA will perform integrity auditing on behalf of the client and reduces the burden of the client. It will send the challenge to the CSS to perform an integrity audit (in a periodic manner). The TPA is also assumed to be a model of a semi-trusted one.

### 3 Conclusion

The various mechanisms to handle storage, resource scheduling, load balancing in the cloud with different service providers and simulation tools is been surveyed with the consideration of cloud security, but all those mechanisms may not be suitable on any cloud environment. The proper cloud mechanisms need to be chosen according to the cloud environment made. When thinking about the future provided with the thought of increasing cloud storage over a regular period. Such an increase will cost the customer to pay more, so this study will provide insight in choosing the proper cloud service providers. While considering the cloud storage mechanisms available, there are also few researches are emerging with a technique for deleting the old files from the cloud to free up the memory space but the historical data (valuable data in terms of quality and quantity) are also lost from the future. The SWA-ECS technique is only focusing on removing duplicate or repeating files. It is not focusing on the removing technique of older files after separating it using historic and unwanted data. To overcome this drawback in storage handling mechanism, there should be a mechanism to free up the space filled by older / non-accessed data and also to prevent the historical values of data.

### References

[1]. Juels A, Kaliski Jr BS. PORs: Proofs of retrievability for large files. In Proceedings of the 14th ACM conference on Computer and communications security 2007 Oct 28, pp. 584-597.

[2]. Dodis Y, Vadhan S, Wichs D. Proofs of retrievability via hardness amplification. theory of Cryptography Conference 2009 Mar 15 (pp. 109-127). Springer, Berlin, Heidelberg.

[3]. Ateniese G, Di Pietro R, Mancini LV, Tsudik G. Scalable and efficient provable data possession. In Proceedings of the 4th international conference on Security and privacy in communication netowrks 2008 Sep 22, pp. 1-10.

[4]. Erway CC, Küpçü A, Papamanthou C, Tamassia R. Dynamic provable data possession. ACM Transactions on Information and System Security (TISSEC), Vol. 17, Issue. 4, 2015 Apr 24, pp.1-29.

[5]. Shacham H, Waters B. Compact proofs of retrievability. In International conference on the theory and application of cryptology and information security 2008 Dec 7, pp. 90-107. Springer, Berlin, Heidelberg.

[6]. Yuan J, Yu S. Proofs of retrievability with public verifiability and constant communication cost in the cloud. In Proceedings of the 2013 international workshop on Security in cloud computing 2013 May 8, pp. 19-26.
[7]. Wang Q, Wang C, Ren K, Lou W, Li J. Enabling public auditability and data dynamics for storage security in cloud computing. *IEEE transactions on parallel and distributed systems*, Vol. 22, Issue. 5, 2010 Oct 28, pp. 847-59.

[8]. Rao GS, Krishna CV, Rao KR. Rational unified process for service oriented application in extreme programming. In2013 *Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT)* 2013 Jul 4 (pp. 1-6). IEEE.

[9]. Lakshmi J, Reddy VK, Naresh T. Resource optimization and task scheduling in cloud computing. *IJETT*. 2020 Feb 11;5(1).

[10]. H. Mohapatra et al., “An Efficient Energy Saving Scheme Through Sorting Technique for Wireless Sensor Network,” *International Journal of Emerging Trends in Engineering Research*, vol. 8, no. 8, 2020.

[11]. Desai MT, Patel MR, Patel MP, Desai T, Patel R, Patel P. Cloud Computing in Education Sector. *IJIRST –International Journal for Innovative Research in Science & Technology*, Vol. 2, Issue. 10, March 2016, pp. 191 – 194.

[12]. Soni P, Upadhyay A, Maheshwari A, Lakkadwala P, Soni P, Upadhyay A, Maheshwari A, Lakkadwala P. Security Related Issues in Cloud Computing: A Survey. *IJIRST –International Journal for Innovative Research in Science & Technology*, Vol. 2, Issue. 11, April 2016, pp. 451 – 454.

[13]. Pujari VG, Khot SR. Novel Approach for Visual Cryptography Using Adaptive Local Filter. *International Journal of Advance Research in Computer Science and Management Studies*, Vol. 3, Issue 2, Feb 2015, pp. 128 – 132.

[14]. Spooorthy V, Mamatha M, Kumar BS. A survey on data storage and security in cloud computing. *International Journal of Computer Science and Mobile Computing*, Vol. 3, Issue. 6, June 2014, pg.306 – 313.

[15]. Ubhad SA, Chaube n N, Dubey SP. Advanced ASCII Based Cryptography Using Matrix Operation, Palindrome Range, Unique id. *International Journal of Computer Science and Mobile Computing*, Vol. 4, Issue. 8, June 2014, pg.66 – 71.

[16]. Dongyu Li. Research on Metadata Management in Cloud Computing. *International Journal of Database Theory and Application*, Vol.9, Issue. 2, 2016, pp. 1-8.

[17]. Qian Z, Yufei G, Hong L, Jin S. A load balancing task scheduling algorithm based on feedback mechanism for cloud computing. *International Journal of Grid and Distributed Computing*, Vol. 9, Issue. 4, 2016, pp.41-52.

[18]. Liu Q, Zhang H. Research on Multi-Dimensional Index in Cloud Computing System. *International Journal of Grid and Distributed Computing*, Vol. 9, Issue. 1, 2016, pp.1-12.

[19]. Jakimoski K. Security techniques for data protection in cloud computing. International Journal of Grid and Distributed Computing, Vol. 9, Issue. 1, 2016, pp.49-56

[20]. Hitesh Mohapatra et al., A Comparative Analysis of Clustering Protocols of Wireless Sensor Network, *International Journal of Mechanical and Production Engineering Research and Development (IJMPERD)* ISSN(P): 2249-6890; ISSN(E): 2249-8001 Vol. 10, Issue 3, pp:8371-8386, Jun 2020, DOI: 10.24247/ijmperdjun2020795.

[21]. Buha P, Sharma P. Secure Cloud Computing Mechanism for Enhancing: MTBAC. *International Journal of Information Sciences and Techniques (IJIST)*, Vol. 6, Issue. 1/2, March 2016, pp. 269 – 273.

[22]. Sharma V, Joshi NK. The Evolution of Big Data Security through Hadoop Incremental Security Model. *International Journal of Innovative Research in Science, Engineering and Technology*, Vol. 4, Issue. 5, May 2015, pp. 3489 – 3493.

[23]. Kamala K. A Study to Assess the Knowledge Regarding Blood Donation among Under Graduate Students Studying in Avvaiyar Government Arts and Science College for Women, Karaikal. *International Journal of Multidisciplinary Educational Research*, Vol. 5, Issue. 2(3), Feb 2016, pp. 1 – 6.
[24]. Radha K, Babu SM, Rao BT. A relative study on service level agreements in cloud computing. In 2015 Global Conference on Communication Technologies (GCCT) 2015 Apr 23 (pp. 66-71). IEEE.

[25]. Prajaka Jagtap, Prachi Jagdale, Smita Gawade, P. B. Javalkar. Online Healthcare System Using the Concept of Cloud Computing. International Journal of Scientific Research in Science, Engineering and Technology, Vol. 2, Issue. 2, April 2016, pp. 943 – 946.

[26]. SK. Jilani Basha, P. Anil Kumar, S. Giri Babu. Storage and Processing Speed for Knowledge from Enhanced Cloud Computing With Hadoop Frame Work : A Survey. International Journal of Scientific Research in Science, Engineering and Technology, Vol. 2, Issue. 2, March 2016, pp. 126 – 132.

[27]. Junho Jung, Jeong Byun and DongKeun Kim. NoSQL based Web Service System for Sharing of Emotion Information in Cloud Computing. International Journal of Software Engineering and Its Applications, Vol. 10, Issue. 3, 2016, pp. 149-162.

[28]. Ruchita Patel, Moni Chandel. Analysis of Various Task Scheduling Algorithms in Cloud Computing. International Research Journal of Engineering and Technology, Vol. 03, Issue. 03, March 2016, pp. 493 – 496.

[29]. Patricia M. Dove. Growing Pains in the Era of Big Data. Editorial – Washington State University, June 2015.

[30]. Yakooob S, Reddy VK, Dastagiraiah C. Multi User Authentication in Reliable Data Storage in Cloud. In Computer Communication, Networking and Internet Security 2017 (pp. 531-539). Springer, Singapore.

[31]. Youn TY, Chang KY, Rhee KH, Shin SU. Efficient Client-Side Deduplication of Encrypted Data with Public Auditing in Cloud Storage. IEEE Access. 2018 May 15;6:26578-87.

[32]. Shacham H, Waters B. Compact proofs of retrievability. In International conference on the theory and application of cryptology and information security 2008 Dec 7, pp. 90-107, Springer, Berlin, Heidelberg.

[33]. Ateniese G, Burns R, Curtmola R, Herring J, Kissner L, Peterson Z, Song D. Provable data possession at untrusted stores. In Proceedings of the 14th ACM conference on Computer and communications security 2007 Oct 28, pp. 598-609.

[34]. Dastagiraiah C, Reddy VK, Pandurangarao KV. Dynamic load balancing environment in cloud computing based on VM ware off-loading. In Data Engineering and Intelligent Computing 2018 (pp. 483-492). Springer, Singapore.

[35]. Kumar, Ranjan, Edalatpanah, SA, and Mohapatra, Hitesh. ‘Note on “Optimal Path Selection Approach for Fuzzy Reliable Shortest Path Problem”. 1 Jan. 2020: 1 – 4, DOI: 10.3233/JIFS-200923.