CLOUD COMPUTING STORAGE SECURITY & VARIOUS PROTOCOLS: A SURVEY

Meenu Tahilyani[1], Amit Dutta[2] and Dalima Parwani[1]
[1]Department of Computer Science, Sant Hirdaram Girls College, Bhopal, India
[2]Department of Computer Science & Application, Barkatullah University, Bhopal, India

ABSTRACT
Cloud computing has become known as an extensive vision of function computing pattern that makes available dependable and durable communications for consumers to remotely store data and personal information, and utilize on require applications and services. In the present scenario, many human beings and associations moderate the trouble of local data storage space and decrease the preservation expenditure by outsourcing data to the cloud users.

KEYWORDS: *Cloud Service Provider, Cloud computing, Third Party Auditor, Data Owner*

I. INTRODUCTION

The advancement in hardware, software, middleware, networking, and virtual machine equipment have led to an appearance of innovative, worldwide distributed computing proposals, that is to say cloud computing, that provides computation facilities and storage as services are accessible from anywhere via the Internet without significant investments in new infrastructure, training, or software is licensing. Cloud computing is one of the greatest platform which provides storage of data in very low cost and is available for all time over the internet. Cloud computing is Internet-based computing, whereby distributed storage, software and in sequence are provided to computers and devices on demand. For an increased level of scalability, ease of use and robustness, some clients may desire their data to be replicated on multiple servers across multiple data centers. Data replication varies according to the environment of data; more copies are required for critical data that cannot without any problems be replicated, while non-critical, reproducible data are stored at reduced levels of redundancy. When users outsource data files in a remote server, the substantial right of entry to the file is essentially lost and the administration of files is delegated to a cloud provider as an unreliable third party [1][2].

In cloud data storage system, the clients store their data in the cloud and no longer acquire the data locally. Data owners shift their data from their local computing systems to the cloud. After data goes into the cloud, the client loses control over it.

If such data storage is susceptible to attacks or Byzantine failures, in which the adversary can modify or delete the data or inject polluted data into the data storage servers or may access the data, these attacks or failures would bring irretrievable losses to the Clients since their data is stored in an uncertain storage pool outside the storage enterprises. These attacks prevent the Clients from accessing the original data correctly. In our current digital world, various organizations produce a large amount of sensitive data including individual in sequence, electronic health records, and financial data. The amount of digital data is increasing at a staggering rate; doubling almost every year and a half [3], and outpacing the storage ability of many organizations. This data often needs to be stored at multiple locations for a long time due to operational purposes and regulatory compliance.

The cloud data storage model in cloud computing consists of three entities namely Clients, Cloud

Service Provider (CSP) and Third Party Auditor (TPA) as illustrated in Figure 1.
Clients:- The Clients are those who have data to be stored, and access it with help of Cloud Service Provider (CSP). They are typically tablet, desktop computers, laptops, mobile phones, etc.
Cloud Service Provider (CSP):- Cloud Service Providers (CSPs) are those who have major resources and proficiency in construction, managing distributed cloud storage servers and provide applications, infrastructure, hardware, enabling technology to Clients as a service via internet.
Third Party Auditor (TPA):- Third Party Auditor (TPA) who has expertise and capabilities that Client may not have and verifies the Integrity of data stored in cloud on behalf of Clients. Based on the audit result, TPA could release an audit report to the Client.

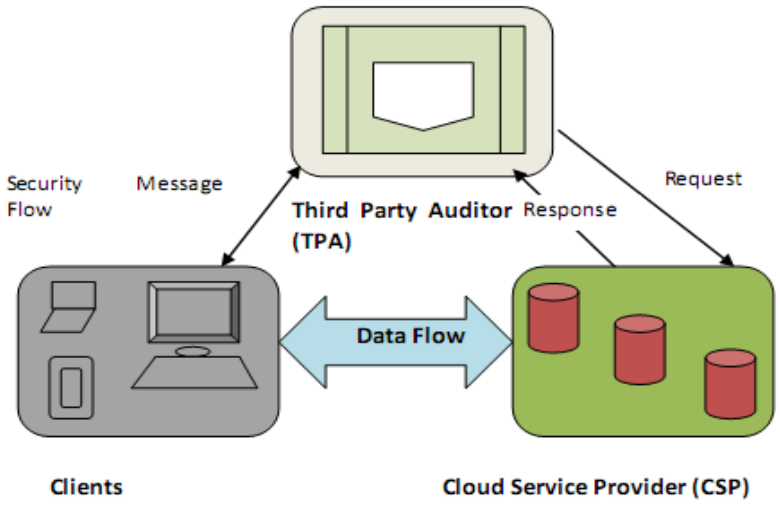

Figure 1. Cloud Data Storage Architecture

In cloud computing paradigm, the Clients store their data files in cloud and access them with the help of Cloud Service Provider (CSP) whenever and wherever they need. The cloud consists of a set of cloud servers, which are consecutively in a concurrent, cooperated and distributed approach. Data redundancy can be utilized with method of erasure-correcting code to additionally accept responsibilities or server crash as user‘s data and encrypting the data can prevents the data leakage. In addition, the Client can frequently verify the Integrity of data without having a local copy of data file. Unauthorized access and misuse of customers' confidential data are serious concerns regarding data outsourcing; hence, it is of significant importance to be aware of data administrators (CSPs) and extension of data access right.

II. THEORETICAL BACKGROUND

Cloud computing has made the extensive dreamed visualization of computing as an effective, authentic and will have the potential to benefit and shape the whole IT industry. When deciding whether or not to move into the cloud, potential cloud users would take into account factors such as service availability, security, system presentation and etc, in the middle of which safety measures is the most important apprehension. Wherever Cloud storage has the prospective of on condition that in nature distributed storage services since cloud can integrate servers and clusters that are allocated all over the world and recommended by special service providers into one virtualized environment. This can potentially resist disastrous failures and accomplish low right to use latency and to a great extent decreased network traffic by carrying data close to where they are required. Cloud data storage belongs to IaaS which allows the Clients to move their data from local computing systems to the remote Cloud. More and more the Clients start choosing to store their data in the Cloud. Thus the main reason for using cloud computing is cost efficiency, which is predominantly accurate for small and medium-sized businesses. Another reason is that the Clients can rely on the Cloud to provide more consistent services, with the intention that they can have the right to use data from anyplace and at any time. Individuals or small-sized business more often than not do not have the

source to keep their servers as dependable as the Cloud does Amazon's Elastic Compute Cloud (EC2) and Amazon Simple Storage Service (S3) [4] and apple icloud [5] are well known examples of cloud data storage.

III. RESEARCH CHALLENGE ON CLOUDS

In spite of its popularity, the cloud computing has evoked safety issues and confidentiality anxieties which hinder its adoption in sensitive environments. The transition to cloud computing model exacerbate security and privacy challenges, mainly due to its dynamic nature and the fact that in this model hardware and software components of a single service span multiple trust domains. In the cloud, data and services are not restricted within a single organization's perimeter. This dynamism and fluidity of data introduces more risk and complicates the problem of access control [6].

Therefore, compared with the traditional models, in cloud computing model ensuring confidentiality and integrity of the end-users' data is far more challenging. Moreover, cloud services are usually multi-tenancy services, meaning that a single communication, proposal, or software is made available for examination of multiple mutually entrusted parties simultaneously [7]. Therefore, confidentiality of these parties' data need to be protected alongside each other. On the other hand in some cases these parties may want to collaborate and share some data with each other in a controlled manner and thus there should be a mechanism that allows them to collaborate. These concerns are mainly devoted to missing or inadequate security and privacy related features, requiring customers to fully trust in the integrity of the cloud provider as well as the provider's security performs. On the other hand, alongside securing the cloud from inside the cloud communications between cloud servers, an alternative possibility is to clearly append appropriate security and privacy characteristics from the external without disturbing the cloud provider's interfaces and internal functioning's. Within the last few years several approaches to eliminate security deficiencies within state of the art cloud storage systems have been proposed [8]. To preserve data privacy, a basic solution is to encrypt data files, and then these encrypt data files upload on cloud server. Adversely, proposing a well-organized and secure data sharing scheme for groups in the cloud is not an easy task due to the following challenging issues.

IV. REMOTE DATA AUDITING

Today, a large amount of individuals and organizations are encouraged to decrease the cost and time occupied to get hold of and protection of local data storage communications by outsourcing the data to the cloud. Using cloud computing, the Cloud Service Provider (CSP) is in incriminating of organization the cloud storage services.

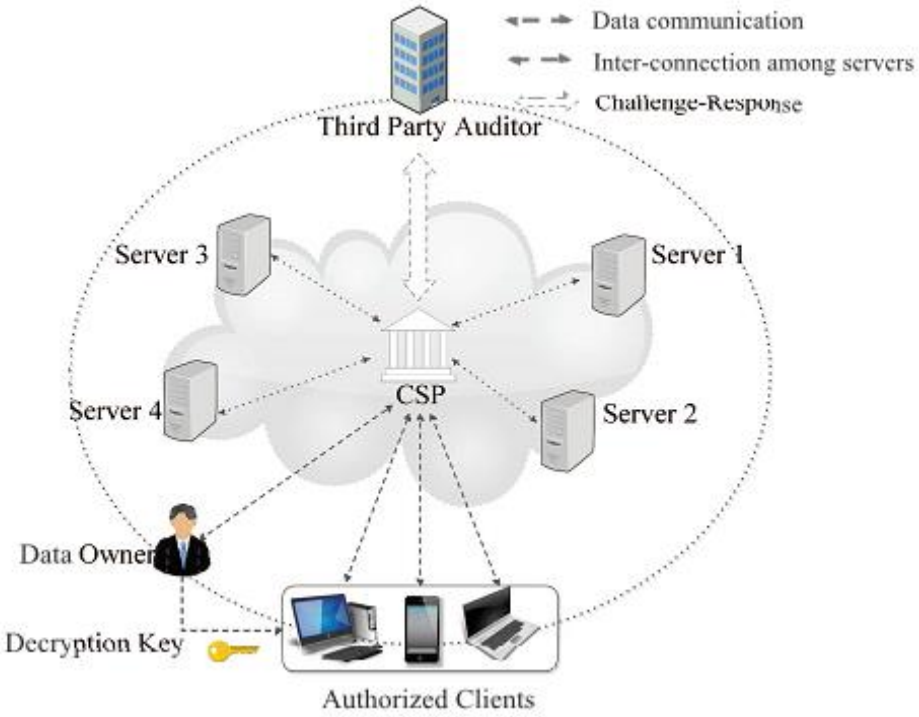

Figure 2. Single Cloud and Multi-Server Audit Architecture

Consequently, the DOs (Data Owners) are incapable to sustain their ownership and undeviating organizes over the uploaded information and as an alternative the data is entirely dealt with by an untrustworthy third party. Alternatively, the CSP or any insider opponent is able to unkindly influence data content lacking user approval or knowledge [9].

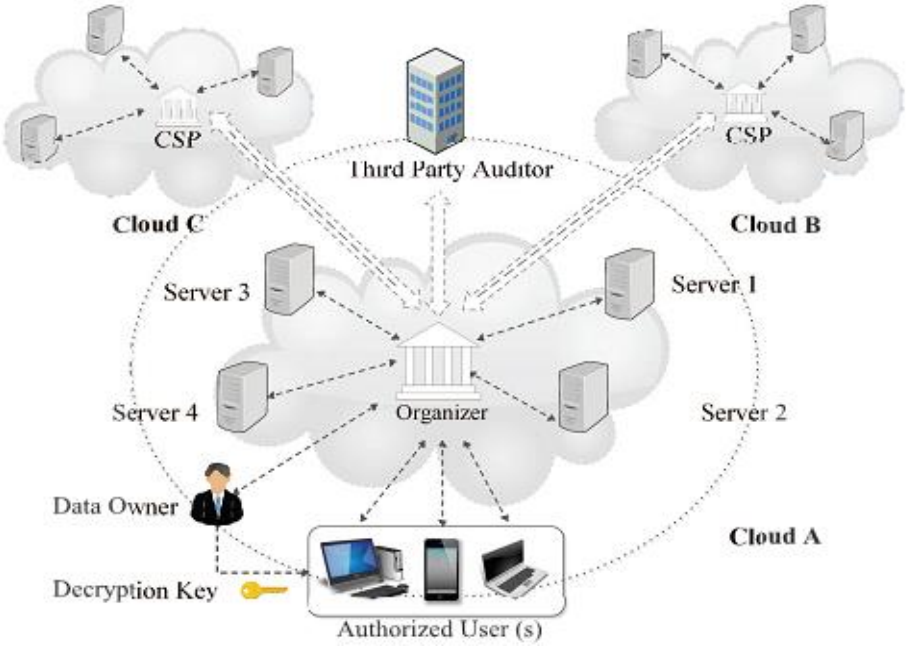

Figure 3. Multi-Cloud and Multi-Server Audit Architecture

Architecture of Remote Data Auditing for Distributed Servers: The RDA schemes for distributed cloud servers often consist of four main things. These things are: (1) Data Owner (DO): the person who uploads his/her data to the cloud space and later s/he might perform delete, insert and append operations on the outsourced data. (2) Cloud Service Provider (CSP): who has tremendous amount of computing resources and stores and manages DOs data. The CSP is also responsible for managing cloud servers. (3) Third Party Auditor (TPA): In order to alleviate the addition load on DOs side, the auditing process is frequently allocated to a Third Party Auditor with sufficient proficiencies and capabilities to achieve the auditing job on behalf of the Data Owner. The TPA's role is particularly important when DOs possess relatively poor computing device in terms of giving out command, storage space and bandwidth. While Third Party Auditor is regarded as a faithful & dependable and is intent at the same time.

As a result, one important countermeasure for the period of data auditing is to avoid Third Party Auditor acquiring knowledge of DO's data content and defend privacy of data. (4) User (individual or endeavor): Who is registered and validated by the DO and authorized to have encoded type of access on the outsourced data [10-12]. The RDA architecture for distributed storage systems are classified into three categories: multi-server, single cloud and multi-server, and multi-cloud and multi-server, which we detail below.

(1) Multi-server model: In this model, the DO allocates multiple copies of the data amongst a number of servers and separately checks each of them. Figure 1 shows the architecture of the multi-server data auditing model.

(2) Single-cloud and multi-server: In this model, all of the servers are distributed within a single cloud where the CSP is in charge of managing the servers. As it gives you an idea about in Figure 2, the DO and the TPA are openly connected to the CSP to a certain extent than all of the servers.

(3) Multi-cloud and Multi-server: Instead of a single cloud, the Data Owner outsources the data surrounded by multiple clouds. Comparable to the earlier representation, one of the CSPs, namely: the organizer is answerable for managing all of the servers and the other CSPs. As shown in Figure 3, the coordinator that is openly joined to the owner obtains data and confront from the DO to distribute among the clouds and the servers. Furthermore, the organizer comprehends the received confirmations from the servers and sends them to the Data Owner.

With the intention to confirm the integrity and accuracy of remote data be a feature of on the cloud, the Third Party Auditor decides on an accidental indices of the outsourced data as a confront message and undeviating that message to moreover the organizer or the CSP in case the TPA is not sustained by auditing service architecture, DO himself must produce the confront. When the coordinator or the CSP accepts the challenge, it is distributed amongst the servers, and then the controller calculates the equivalent answer by comprehending the received messages from the servers. After receiving a reply from the organizer or the CSP, the confirmation is accepted out by the examiner to make confident the consistent assignment of the file in the cloud storage space [13].

V. REPLICATION-BASED REMOTE DATA AUDITING

When DOs store data in an unpredictable storage structure, as is the case in the cloud and mobile cloud computing patterns, being without a job participate an essential responsibility to get better the reliability against data disappointments. The simplest and the major widespread method to accomplish the aforementioned objective is to use a duplication method in which multiple copies of data are outsourced within the distributed storage methods [14]. Even though the accomplishment of this technique is comparatively straightforward, there is no tough confirmation to confirm that the cloud essentially stores multiple copies of the data files [9]. The inexperienced technique to conquer the collusion attack is to be appropriate a single PDP method [15] t times for t different servers. On the other hand, as a matching copy of the file is stored on all of the t servers, the servers can unite simultaneously and make consider that t copies of files are accumulated while only a single copy is stored in authenticity. Additionally, the computational cost on the client for the pre-processing of the file is t times better than the computational weight when the single PDP technique is utilized. Consequently, such a technique is inappropriate for distributed storage schemes for the most part for larger values of t.

VI. LITERATURE SURVEY

Zhu et al. [16] proposed a new method to release the internal architecture of the organization to the consumers. The motivation is that consumers must effectively pre-process the files as a part of system process that requires a calculation transparency proportionate with the number of recommended servers. Consequently, an extreme amount of computational load is required on the client side. Other than the abovementioned method, there is no means to confirm whether the CSP in fact accumulates the accurate number of file copies as consented to the consumer. When the consumer is responsive of the internal architecture of the cloud, the CSP is not capable to improve proficiently the architecture without they required informing the client.

When DOs outsource data to the remote cloud or delegate the auditing task to the trusted third party, it is important for them that the auditors or cloud must not be given opportunity to gain the knowledge of data content or be able to make a copy of the original data [17]. That is to say that, most of the data auditing methods for the distributed cloud servers usually assumes that the TPA is a trustworthy agent. Though, such an illogical assumption further leads to data leakage. Randomization of data blocks and tags is a common method to address the privacy issue to prevent tag or data leakage during the entire verification phase.

In this paper [18] we take a closer look at distributed cloud storage systems approaches and provide an overview of different (security) features such systems support. We provide an overview of information dispersal strategies to realize reliable distributed cloud storage systems and provide a general idea of up to date cloud storage move towards. Subsequently, they examine them regarding security possessions. Additionally, we talk about the requirement of confidentiality features and in exacting features to make available access privacy in offered distributed cloud storage space methods, which is significant way for upcoming research on distributed cloud storage.

Table 1: Literature Review: The advantages and requirements of the proposed token

S. No.	Paper	Author	Advantages	Issues
1	An Efficient and Secure Dynamic Auditing Protocol for Data Storage in Cloud Computing. [19]	Kan Yang, Xiaohua Jia	To our proposed auditing protocols are make safe and proficient, mainly it reduce the overall expenditure of the auditor.	The proposed protocol at this time is to explain the data privacy problem, and it may generate an encrypted confirmation with the confront crush by using the Bilinearity property of the bilinear pairing
2	Replicated Data Integrity Verification in Cloud. [20]	Mukundan et al. 2012	This method to authenticate the reliability and inclusiveness of multiple copies and utilized probabilistic encryption algorithm, specifically the Paillier encryption method to produce distinctive copies of the original file	The performance of system is clear
3.	Multiple-File Remote Data Checking for cloud storage.[21]	Yan et al. 2012	An independent server or one of the offered CSP is presumed as a manager, who has dependability for supervision all of the CSPs, commencing and organizing the verification progression, and communicating in a straight line with the	Here challenge is issued by the consumer, the controller comprehensives all of the answers take delivery from the CPSs into one answer by using the HVR method, to be sent to the consumer.

			consumer.	
4.	Transparent, Distributed, and Replicated Dynamic Provable Data Possession. [22]	Etemad and Alptekin Kupcu 2013	Hides the internal structural design of the system to the consumers; Ensures that the CSP accumulates the right number of duplications;	To be transparent to the consumer point of view to get better scalability, ease of use, load balancing, and fault tolerance in RDA methods.
5.	Anonymous Cloud: A data ownership privacy provider framework in cloud computing. [23]	Khan, S. M. and K. W. Hamlen 2012	Anonymous Cloud obscures data origin from cloud nodes that calculate more than the data, and cover ups beneficiary distinctiveness in the form of IP addresses and possession tags.	Decentralizes expectation by decoupling billing information from suggested jobs. This work apprehensions the difficulty of privacy-preserving calculation.
6.	Enabling Public Audit-ability and Data Dynamics for Storage Security in Cloud Computing. [24]	Q. Wang, C. Wang, K. Ren, W. Lou, and J. Li	Here recommended a dynamic auditing protocol that can sustain the dynamic operations of the data on the cloud servers.	This technique may reveal the data content to the auditor because it necessitates the server to throw the linear combinations of data blocks to the auditor.

VII. CONCLUSION AND FUTURE WORK

Cloud computing is a promising standard with its cost-effectiveness and elasticity and hence provides a fabulous energy. On the other hand, there are many safety issues and challenges that, if not dealt with may obstruct its prompt implementation and development. Existing cloud computing stage accomplishments place centralized widespread trust over all the cloud nodes that enlarge the anxieties of data and computation reliability and protection.

It is obvious that Future work will involve further investigation to examine these challenges and provide a better understanding of the existing cloud computing security challenges. The further investigation will help to develop approaches/frameworks that capable to solve these challenges and overcome the identified obstacles.

This paper includes the survey of all the existing technique and protocols implemented for cloud storage and security. On the basis of various techniques issues and problems in the existing protocols various future directions can be planned such as 1) practical implementation of cloud data storage on EC2, Amazon etc. 2) implementation of some new and efficient protocol using Two factor authentication. 3) Scheduling of cloud resources for the minimization of energy.

REFERENCES

[1] Cong Wang, S.S.M. Chow, Qian Wang, Kui Ren, and Wenjing Lou. (2013), "Privacy-Preserving Public Auditing Protocol for Secure Cloud Storage" ,.

[2] CongWang, Kui Ren, Wenjing Lou, and Jin Li. (2010)., "Toward publicly auditable secure cloud data storage services*" IEEE Network* 24, , pp 19–24.

[3] Aameek Singh and Ling Liu. Sharoes(2008), "A data sharing platform for outsourced enterprise storage environments", *Proceedings of the 24th International Conference on Data Engineering, ICDE*, IEEE, pp 993-1002.

[4] Amazon.com, "Amazon Web Services (AWS)"(2008), Online at *http://aws.amazon.com.*

[5] Apple (2010), "ICloud" ,*Online at http://www.apple.com/icloud/what-is.html.*

[6] Andrzej M. Goscinski Rajkumar Buyya, James Broberg(2011), "Cloud Computing: Principles and Paradigms" *Wiley*.

[7] Jose M. Alcaraz Calero, Nigel Edwards, Johannes Kirschnick, Lawrence Wilcock, and Mike Wray(2010), "Toward a multi-tenancy authorization system for cloud services*", IEEE Security and Privacy*, pp 48-55.

[8] C. Basescu, C. Cachin, I. Eyal, R. Haas, A. Sorniotti, M. Vukolic, and I. Zachevsky(2012), "Robust Data Sharing with Key-Value Stores," in *Intl. Conference on Dependable Systems and Networks (DSN), IEEE.*

[9] Bo Chen, Reza Curtmola, Giuseppe Ateniese, and Randal Burns (2010), " Remote data checking for network coding-based distributed storage systems" .
[10] Dongyoung Koo, Junbeom Hur, and Hyunsoo Yoon (2013), "Secure and efficient data retrieval over encrypted data using attribute-based encryption in cloud storage." ,*Computers & Electrical Engineering*, Vol. 39, No. 1 , pp 34–46.
[11] Sandeep K. Sood. (2012), " A combined approach to ensure data security in cloud computing",. *Journal of Network and Computer Applications,* Vol. 35, No. 6 , pp 1831–1838.
[12] Shucheng Yu, Wnjing Lou, and Kui Ren (2012), "Data Security in Cloud Computing",. *Morgan Kaufmann/Elsevier, Book section* 15, pp 389–410.
[13] Xiaolan Zhang, Giovanni Neglia, and Jim Kurose (2012), "Chapter 10 - Network Coding in Disruption Tolerant Networks", *Academic Press, Boston*, pp 267–308.
[14] Liu Ying and V. Vlassov (2013), " Replication in Distributed Storage Systems: State of the Art, Possible Directions, and Open Issues", *International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery IEEE*, pp 225–232.
[15] Giuseppe Ateniese, Randal Burns, Reza Curtmola, Joseph Herring, Lea Kissner, Zachary Peterson, and Dawn Song. (2007), " Provable data possession " *untrusted stores*.
[16] Yan Zhu, Huaixi Wang, Zexing Hu, Gail-Joon Ahn, Hongxin Hu, and Stephen S. Yau. (2010), "Efficient Provable Data Possession for Hybrid Clouds", *Proceedings of the 17th ACM Conference on Computer and Communications Security (CCS '10)*. ACM, New York, NY, USA, pp 756–758.
[17] Cong Wang, S.S.M. Chow, Qian Wang, Kui Ren, and Wenjing Lou. (2013), "Privacy-Preserving Public Auditing for Secure Cloud Storage".
[18] Daniel Slamanig, Christian Hanser (2012), "On Cloud Storage and the Cloud of Clouds Approach" *IEEE*.
[19] Kan Yang, Xiaohua Jia, "An Efficient and Secure Dynamic Auditing Protocol for Data Storage in Cloud Computing" *IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS,* VOL. 24, NO. 9.
[20] Raghul Mukundan, Sanjay Madria, Mark Linderman, and NY Rome (2012), "Replicated Data Integrity Verification in Cloud.", *The Bulletin of the Technical Committee on Data Engineering* , pp 55–65
[21] Da Xiao, Yan Yang, Wenbin Yao, Chunhua Wu, Jianyi Liu, and Yixian Yang (2012), " Multiple-File Remote Data Checking for cloud storage", *Computers & Security*, Vol. 31, No. 2 , pp 192–205.
[22] Mohammad Etemad and Alptekin Kupcu. Transparent(2013), "Distributed, and Replicated Dynamic Provable Data Possession. IACR Cryptology", *ePrint Archive* , pp 225.
[23] Khan, S. M. and K. W. Hamlen (2012), " AnonymousCloud: A data ownership privacy provider framework in cloud computing.", *In Proceedings of the 11th IEEE International Conference on Trust, Security and Privacy in Computing and Communications (Trust-Com),* pp 170-176.
[24] Q. Wang, C. Wang, K. Ren, W. Lou, and J. Li(2011), "Enabling Public Auditability and Data Dynamics for Storage Security in Cloud Computing,", *IEEE Trans. Parallel Distributed Systems,* Vol. 22, No. 5, pp 847-859.

AUTHORS

Meenu Tahilyani working as Assistant Professor in the Department of Computer Science at Sant Hirdaram Girls College, Bhopal, Madhya Pradesh, India. She received the Bachelor of Computer Application degree from Govt. Geetanjali Girls College, Barkatullah University, India. She obtained the Master of Computer Application degree, from Indira Gandhi National Open University, India.

Amit Dutta working as Nodal Officer (Online Services) Barkatullah University Bhopal. He received the Bachelor of Science degree from Barkatullah University, India. He obtained the Master of Computer Application degree from BHOJ University and Master of Science in Mathematics degree from Barkatullah University. He is awarded doctorate in Mathematics and Computer Science. His area of specialization is Neural Networks and Artificial Intelligence.

Dalima Parwani working as Assistant Professor in the Department of Computer Science at Sant Hirdaram Girls College, Bhopal, Madhya Pradesh, India. She received the Bachelor of Commerce degree from Govt. Geetanjali Girls College, Barkatullah University, India. She obtained the Master of Computer Application degree, from Punjab Technical University, India.