RESEARCH PAPER | Computer Science | Volume : 3 | Issue : 11 | Nov 2013 | ISSN - 2249-555X

# TPA Added Service in Cloud Computing

KEYWORDS

**K. Palanisamy**

ASST.PROFESSOR, DEPARTMENT OF COMPUTER SCIENCE, SALEM SOWDESWARI COLLEGE(SFCW), SALEM-636010, TN.

**ABSTRACT** *Cloud computing involves three segments. They are cloud user(U), who has large amount of data files to be stored in the cloud , the cloud server(CS), which is managed by cloud service provider(CSP) to provide data storage service. The system adds the fourth layer the Third Party Auditor(TPA),who has expertise and capabilities that maintain the stability between cloud users, cloud data and cloud system. The TPA has the responsibilities to manage four things*

*1)Public audit ability : to allow TPA to verify the correctness of the cloud data on demand without retrieving a copy of the whole data or introducing additional on-line burden to the cloud users;*

*2)Storage correctness : to ensure that there exists no cheating cloud server that can pass the audit from TPA without indeed storing users data intact;*

*3)Privacy – preserving: to ensure that there exists no way for TPA to derive users data content from the information collected during the auditing process;*

*4)Batch auditing: to enable TPA with secure and efficient auditing capability to cope with multiple auditing delegations from possibly large number of different users simultaneously;*

*5)Lightweight: to allow TPA to perform auditing with minimum communication and computation overhead.*

*The U(Cloud User)sends the file F to Cloud server and disconnects the connections then U forms the META Data about F to TPA manages the META data without the whole storage data storage to avoid the duplication. User demands the TPA to audit with correctness of Cloud server about F. The TPA starts with auditing by connecting with cloud using META Data. For Cheating prevention META Data consists the Random Pin Approach.TPA uses the thread fashion to manage the auditing user demands with defined number of ways.*

## 1.1 INTRODUCTION

The third party auditor (TPA), who has expertise and capabilities that cloud users do not have and is trusted to assess the cloud storage service security on behalf of the user upon request. Users rely on the CS for cloud data storage and maintenance. They may also dynamically interact with the CS to access and update their stored data for various application purposes. The users may resort to TPA for ensuring the storage security of their outsourced data, while hoping to keep their data private from TPA.

We consider the existence of a semi-trusted CS as does. Namely, in most of time it behaves properly and does not deviate from the prescribed protocol execution. However, during providing the cloud data storage based services, for their own benefits the CS might neglect to keep or deliberately delete rarely accessed data files which belong to ordinary cloud users. Moreover, the CS may decide to hide the data corruptions caused by server hacks or Byzantine failures to maintain reputation.

We assume the TPA, who is in the business of auditing, is reliable and independent, and thus has no incentive to collude with either the CS or the users during the auditing process. TPA should be able to efficiently audit the cloud data storage without local copy of data and without bringing in additional on-line burden to cloud users. The Cloud Computing model of computing is a distributed application structure that partitions tasks or workloads between the providers of a resource or service, called Cloud servers, and service requesters, called clients.

Often clients and servers communicate over a computer network on separate hardware, butboth client and server may reside in the same system. A server machine is a host that is running one or more server programs which share their resources with clients. A client does not share any of its resources, but requests a server's content or service function. Clients therefore initiate communication sessions with servers which await incoming requests.

Downloading is the transmission of a file from one computer system to another, usually smaller computer system. From the Internet user's point-of-view, to download a file is to request it from another computer (or from a Web page on another computer) and to receive it.

Uploading is transmission in the other direction: from one, usually smaller computer to another computer. From an Internet user's point-of-view, uploading is sending a file to a computer that is set up to receive it. People who share images with others on bulletin board systems (BBS) upload files to the BBS.

The File Transfer Protocol (FTP) is the Internet protocol for downloading and uploading files and a number of special applications can furnish FTP services for you. (However, if you are downloading through a Web page, the FTP request is set up for you by the Web page. You are usually asked where you want the downloaded file placed on your hard disk, and then the downloading transmission takes place.)

Metadata describes other data. It provides information about a certain item's content. For example, an image may include metadata that describes how large the picture is, the color depth, the image resolution, when the image was created, and other data. A text document's metadata may contain information about how long the document is, who the author is, when the document was written, and a short summary of the document.

Metadata summarizes basic information about data, which

Volume : 3 | Issue : 11 | Nov 2013 | ISSN - 2249-555X

can make finding and working with particular instances of data easier. For example, author, datecreated and date modified and file size are examples of very basic document metadata. Having the abilty to filter through that metadata makes it much easier for someone to locate a specific document.

A computer network, often simply referred to as a network, is a collection of computers and devices interconnected by communications channels that facilitate communications among users and allows users to share resources. Networks may be classified according to a wide variety of characteristics. A computer network allows sharing of resources and information among interconnected devices.

Computer networking means connecting computers electronically to allow sharing of resources. Resources range from attached printers to files, disk space and secure information. There are three major categories of networks. Computers connect in the same building using Local Area Networks (LANs). Metropolitan Area Networks (MANs) connect buildings withing a city. Wide Area Networks (WANs) connect sites throughout a country or around the world.

On a single processor, multithreading generally occurs by time-division multiplexing (as in multitasking): the processor switches between different threads. This context switching generally happens frequently enough that the user perceives the threads or tasks as running at the same time. On a multiprocessor or multi-core system, the threads or tasks will actually run at the same time, with each processor or core running a particular thread or task.

It is easy to confuse multithreading with multitasking or multiprogramming , which are somewhat different ideas. Multithreading is the ability of a program or an operating system process to manage its use by more than one user at a time and to even manage multiple requests by the same user without having to have multiple copies of the programming running in the computer.

Each user request for a program or system service (and here a user can also be another program) is kept track of as a thread with a separate identity. As programs work on behalf of the initial request for that thread and are interrupted by other requests, the status of work on behalf of that thread is kept track of until the work is completed.

Multi-Threading is the ability of a CPU to execute several threads of execution apparently at the same time. CPUs are very fast at executing instructions. Modern PCs can execute nearly a billion instructions every second. Instead of running the same program for one second, the CPU will run one program for perhaps a few hundred microseconds then switch to another and run it for a short while and so on.It's also possible for a program to have multiple parts that run at the same time (or appear to!).

For example a background task could be responding to mouse input while a file is loaded intoRAM and another task updates a progress bar on screen.

Multithreading is similar to multitasking, but enables the processing of multiple threads at one time, rather than multiple processes. Since threads are smaller, more basic instructions than processes, multithreading may occur within processes.By incorporating multithreading, programs can perform multiple operations at once.

For example, a multithreaded operating system may run several background tasks, such as logging file changes, indexing data, and managing windows at the same time. Web browsers that support multithreading can have multiple windows open with JavaScript and Flash animations running simultaneously. If a program is fully multithreaded, the different processes should not affect each other at all, as long as the CPU has enough power to handle them.

File sharing is the public or private sharing of computer data or space in a network with various levels of access privilege. While files can easily be shared outside a network (for example, simply by handing or mailing someone your file on a diskette), the term file sharing almost always means sharing files in a network, even if in a small local area network. File sharing allows a number of people to use the same file or file by some combination of being able to read or view it, write to or modify it, copy it, or print it.

Typically, a file sharing system has one or more administrators. Users may all have the same or may have different levels of access privilege. File sharing can also mean having an allocated amount of personal file storage in a common file system.

File sharing has been a feature of mainframe and multi-user computer systems for many years. With the advent of the Internet, a file transfer system called the File Transfer Protocol (FTP) has become widely-used. FTP can be used to access (read and possibly write to) files shared among a particular set of users with a password to gain access to files shared from an FTP server site. Many FTP sites offer public file sharing or at least the ability to view or copy files by downloading them, using a public password (which happens to be "anonymous").

Most Web site developers use FTP to upload new or revised Web files to a Web server, and indeed the World Wide Web itself can be thought of as large-scale file sharing in which requested pages or files are constantly being downloaded or copied down to the Web user.

More usually, however, file sharing implies a system in which users write to as well as read files or in which users are allotted some amount of space for personal files on a common server, giving access to other users as they see fit. The latter kind of file sharing is common in schools and universities. File sharing can be viewed as part of file systems and their management.

Any multi-user operating system will provide some form of file sharing. Among the best known network file systems is (not surprisingly) the Network File System (NFS). Originally developed by Sun Microsystems for its UNIX-based systems, it lets you read and, assuming you have permission, write to sharable files as though they were on your own personal computer.

Files can also be shared in file systems distributed over different points in a network. File sharing is involved in groupware and a number of other types of applications.

It has now emerged as a methodology for sharing digital files across the web. Napster pioneered this, distributing music files. Others, such as gnutella are proving to be interesting distribution models for music etc.

Cloud computing is a general term for anything that involves delivering hosted services over the internet. These services are broadly divided into three categories: Infrastructure-as-a-Service (IaaS), Platform-as-a-Service (PaaS) and Software-as-a-Service (SaaS).

A cloud service has three distinct characteristics that differentiate it from traditional hosting. It is sold on demand, typically by the minute or the hour; it is elastic - a user can have as much or as little of a service as they want at any given time and the service is fully managed by the cloud service provider (the consumer needs nothing but a personal computer and Internet access).

The advantage of cloud is cost savings. The prime disadvan-

Volume : 3 | Issue : 11 | Nov 2013 | ISSN - 2249-555X

tage is security. Cloud computing is used by many software industries nowadays. Since the security is not provided in cloud, many companies adopt their unique security structure. For eg) Amazon has its own security structure.

Introducing a new and uniform security structure for all types of cloud is the problem we are going to tackle in this paper. Since the data placed in the cloud is accessible to everyone, security is not guaranteed. To ensure security, cryptographic techniques cannot be directly adopted. Sometimes the cloud service provider may hide the data corruptions to maintain the reputation. To avoid this problem, we introduce an effective third party auditor to audit the user's outsourced data when needed.

The security is achieved by signing the data blocks. We utilized public key based homomorphic authenticator with random masking to achieve privacy preserving auditing protocol.

TPA performs the auditing task for each user i.e. single auditing. This increases the auditing time and computation overhead. The technique of Bilinear Aggregate Signature is used to achieve batch auditing i.e.) multiple auditing tasks simultaneously. Earlier works performs auditing only for static data. We enhance the system with dynamic operations on data blocks i.e.) data update, append and delete.

Cloud computing has been envisioned as the next generation architecture of the IT enterprise due to its long list of unprecedented advantages in IT: on demand self-service, everywhere network access, location-independent resource pooling, rapid resource elasticity, usage-based pricing, and transference of risk. One fundamental aspect of this new computing model is that data is being centralized or outsourced into the cloud. From the data owners' perspective, including both individuals and IT enterprises, storing data remotely in a cloud in a flexible on-demand manner brings appealing benefits: relief of the burden of storage management, universal data access with independent geographical locations, and avoidance of capital expenditure on hardware, software, personnel maintenance, and so on.

While cloud computing makes these advantages more appealing than ever, it also brings new and challenging security threats to the outsourced data. Since cloud service providers are separate administrative entities, data outsourcing actually relinquishes the owner's ultimate control over the fate of their data. As a result, the correctness of the data in the cloud is put at risk due to the following reasons. First of all, although the infrastructures under the cloud are much more powerful and reliable than personal computing devices, they still face a broad range of both internal and external threats to data integrity.

Outages and security breaches of noteworthy cloud services appear from time to time. Second, for benefits of their own, there are various motivations for CSPs to behave unfaithfully toward cloud customers regarding the status of their sourced data. Examples include CSPs, for monetary reasons, reclaiming storage by discarding data that has not been or is rarely accessed or even hiding data loss incidents to maintain a reputation In short, although outsourcing data into the cloud is economically attractive for the cost and complexity of long-term Large scale data storage, it does not offer any guarantee on data integrity and availability. This problem, if not properly addressed, may impede successful deployment of the cloud architecture.

**1.2 Remote Parallel Computing System Based on Multi-Platform:**
Introduces the design and implementation of a Remote System for Parallel Computing.Based on Multi-platform. This system support remote operating and running Linux system parallel programes on Windows system and provide a kind of uniform mode for the applied customer to operation, The customer does not need to know the concrete structure of the parallel computing system,can operate the parallel system and complete the computation or simulation what they need.

This system reduces the complication on developing parallel programs and improves the efficiency on using parallel computing system, and solves the problem of real-time data transmission between server and client machine. It has highly practical value.

**1.3 Data Security Implementation for Real Time:**
In experimental field such as computer systems science, it is common to study real-world behavior as a means of gaining insight. One time-honored methodology is the collection of trace data, either as a snapshot or over a period of time, for later replay or analysis.

The sharing of Internet packet traces is very limited because real-world traces contain many kinds of sensitive information, such as host addresses, emails, personal web pages, and even authentication keys. The lack of such traces greatly limits research on application protocols. It is especially crippling for network intrusion detection research, forcing researchers to devise synthetic attacks. Inthis paper we describe an approach to transform and anonymize packet traces.

The paper elaborates on the anonymization of the internet packet traces and corresponding trace transformation The algorithm discussed can anonymize both packet headers and payloads, and can perform application-level transformations such as editing HTTP or SMTP headers, replacing the content of Web items with MD5 hashes, or altering filenames or reply codes that match given patterns.

Thus the paper aims to shed light on a new trace transformation & anonymization techniques with features for the future, coupled with reliability and frugal use of resources take technology to the masses as well as the researchers, making the world a truly global village. As such, we hope to help open up new opportunities in Internet measurement and network intrusion detection re-search.

**1.4 Dynamic Data Based on the Third-party Verifier in Cloud Computing:**
Cloud Computing has been envisioned as the next-generation architecture of IT Enterprise. More and more users store data in "clouds". As such, it has become crucial for an archive service to be capable of providing evidence to demonstrate the integrity of data for which it is responsible.

This work studies the problem of ensuring the integrity of data storage in Cloud Computing. In our work, we consider the task of allowing a third party verifier (TPV), on behalf of the cloud client, to verify the integrity of the dynamic data stored in the cloud.

In particular, to achieve efficient data dynamics, we improve the Proof of Retrievability model by manipulating the classic Merkle Hash Tree (MHT) construction for block tag authentication. Extensive performance analysis show that the proposed scheme is highly efficient.

**1.5 Third-Party Storage Security Protocol:**
With the trend of networking and the growth of data value, networked storage security becomes hotpot of research. Aiming at the scalability and security requirement of massive storage system, we proposed a security storage service model combining storage mechanism and security policy and designed a set of scalable third-party security protocols.

**1.6 Data Dynamics for Storage Security in Cloud Computing:**
Cloud Computing has been envisioned as the next-gener-

Volume : 3 | Issue : 11 | Nov 2013 | ISSN - 2249-555X

ation architecture of IT Enterprise. It moves the application software and databases to the centralized large data centers, where the management of the data and services may not be fully trustworthy.

This unique paradigm brings about many new security challenges, which have not been well understood. This work studies the problem of ensuring the integrity of data storage in Cloud Computing. In particular, we consider the task of allowing a third party auditor (TPA), on behalf of the cloud client, to verify the integrity of the dynamic data stored in the cloud.

The introduction of TPA eliminates the involvement of the client through the auditing of whether his data stored in the cloud are indeed intact, which can be important in achieving economies of scale for Cloud Computing. The support for data dynamics via the most general forms of data operation, such as block modification, insertion, and deletion, is also a significant step toward practicality, since services in Cloud Computing are not limited to archive or back-up data only. While prior works on ensuring remote data integrity often lacks the support of either public auditability or dynamic dataoperations, this paper achieves both. We first identify the difficulties and potential security problems of direct extensions with fully dynamic data updates from prior works and then show how to construct an elegantverification scheme for the seamless integration of these two salient features in our protocol design.

**1.7 Ensuring data storage security in Cloud Computing:**
Cloud computing has been envisioned as the next-generation architecture of IT enterprise. In contrast to traditional solutions, where the IT services are under proper physical, logical and personnel controls, cloud computing moves the application software and databases to the large data centers, where the management of the data and services may not be fully trustworthy.

This unique attribute, however, poses many new security challenges which have not been well understood. In this article, we focus on cloud data storage security, which has always been an important aspect of quality of service. To ensure the correctness of users' data in the cloud, we propose an effective and flexible distributed scheme with two salient features, opposing to its predecessors. By utilizing the homomorphism token with distributed verification of erasure-coded data, our scheme achieves the integration of storage correctness insurance and data error localization, i.e., the identification of misbehaving server (s).

Unlike most prior works, the new scheme further supports secure and efficient dynamic operations on data blocks, including: data update, delete and append. Extensive security and performance analysis shows that the proposed scheme is highly efficient and resilient against Byzantine failure, malicious data modification attack, and even server colluding attacks.

**1.8 Data Integrity on Remote Storage**
Nowadays, the cloud computing have engulfed not only the IT industry but also the general public's all around the world. Our daily life is now full of various cloud services such as Gmail or Google Document. Although the cloud services can provide on-line platforms for co-working between a group of collaborators, trust is always a hesitation for a user to adopt cloud services.

In this paper, we aim at the integrity issue for on-line co-working and seek for a proper solution. We develop a framework to enable the remote data integrity verification for on-line co-working scenarios. In addition to provide a framework, we also show the feasibility of our framework by providing a concrete example.

**1.9 Third-Party Logistics Information Management Model:**
By illustrating the classification rules and association rules of data mining and taking into consideration of the content and requirement of third-party logistics information management and analyzing the security problems exist in third-party logistics information management, the paper suggests a safety third-party logistics information management model based on data mining. This model could provide better solution to the security problems in logistics information management, and help ensure the security of third-party logistics information management.

**1.10 Result Verification Schemes for Map Reduce:**
Recent development in Internet-scale data applications and services, combined with the proliferation of cloud computing, has created a new computing model for data intensive computing best characterized by the Map Reduce paradigm. The Map Reduce computing paradigm, pioneered by Google in its Internet search application, is an architectural and programming model for efficiently processing massive amount of raw unstructured data.

With the availability of the open source Hadoop tools, applications built based on the Map Reduce computing model are rapidly growing. In this work, we focus on a unique security concern on the Map Reduce architecture. Given the potential security risks from lazy or malicious servers involved in a Map Reduce task, we design efficient and innovative mechanisms for detecting cheating services under the Map Reduce environment based on watermark injection and random sampling methods.

**REFERENCE** [1] Privacy-Preserving Public Auditing for Data Storage Security in Cluster Computing, IEEE INFOCOM 2010, San Diego, CA, March 2010 | [2] M. Armbrustet al., (Feb. 2009.) "Above the Clusters: A Berkeley View of Cluster Computing," Univ. California, Berkeley, Tech. Rep. UCBEECS-2009-28, | [3] Juels, J. Burton, and S. Kaliski, (07, Oct. 2007,) "PORs: Proofs of Retrievability for Large Files," Proc. ACM CCSpp. 584–97. | [4] P. Mell and T. Grance, (2009) "Draft NIST Working Definition of Cluster Computing" | | WEBSITES | • www.jsptutorial.com | • www.htmlreference.com | • www.universalteacher.com |