Survey Paper on Improved Transactional Security in Cloud Computing

Shilpi Chaubey, Dr.S.M.Ghosh
Research Scholar,M.Tech (Software Engineering), Associate Professor
Computer Science & Engineering Dept
Rungta College of Engg. & Tech
Bhilai(C.G.), India

***Abstract*— A transaction is an execution of a set of programs that access shared recoverable resources (e.g., data items).A resource is recoverable if its state, as viewed by the transaction when first accessing this resource, can be restored during the transaction, if has been modified by the transaction. A distributed Transaction involves several (more than one) processes that may access different recoverable resources. In distributed transactional database systems entities merge together to form proofs of authorizations which are justified by collections of certified credentials. Due to the risk of having the authorization policies or under the user credentials which are in the inconsistent states over extended time periods these proofs and credentials may be evaluated.**

Key words: Distributed transactions, consistency, two-phase commit protocol, atomic commit protocol.

I. INTRODUCTION

Cloud computing refers to different applications and services which run on a distributed network which are being using the virtualized resources and these are being accessed by common Internet protocols and networking standard[1]. The resources are virtual and limitless which are being differentiated by the notion and the different user becomes helpful for abstracting the details of the physical systems on which software runs.

Cloud computing took the technology, services, and applications those are same on to the Internet and which makes them to convert into a utility of self-services. The word "cloud" is referencing to the two essential concepts:

• *Abstraction*: From different users and developers the details of system implementation are abstracted from cloud computing. Different applications run on physical systems that are not specified, locations which are unknown data are stored there, others are being outsourced to the administration of systems, and access by users which are seem to be in all places.

•*Virtualization*: By pooling and sharing resources cloud computing virtualizes the systems. Storage and systems are provisioned as needed from a centralized infrastructure, on a metered basis costing are assessed, multi-tenancy is enabled, with agility resources are scalable.

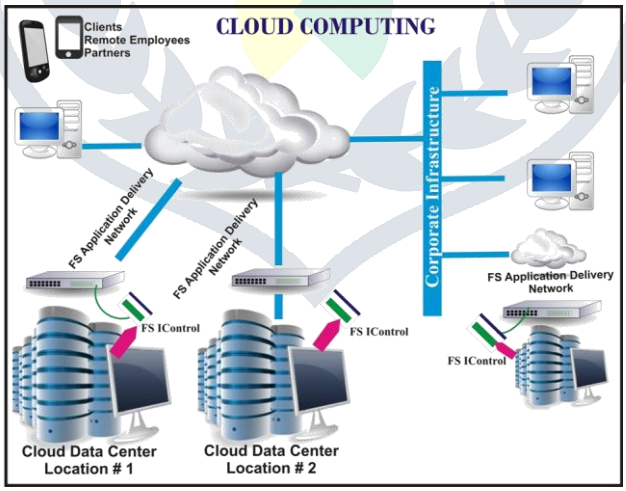

Figure1: Cloud Computing

Cloud Computing is used to add different number of computing concepts and technologies such as Service Oriented Architecture (SOA), other technologies and virtualization on the Internet, to satisfy computing needs of different users it provides common business applications which are online through different web browsers , while the server are being used to store their software and data. The Maturing of these technologies are represented by cloud computing and maturity is a marketing term to represent it and the different services which are being provide by them.

Transaction: Transaction processing systems, such as Web based services, which represents a large class of applications which are hosted by computing clouds; these types of applications run multiple instances of the service and it require reliable and in-order delivery of messages [2].A transaction is an execution of a set of programs that access shared recoverable resources (e.g., data items). A resource is recoverable if its state, as viewed by the transaction when first accessing this resource, can be restored during the transaction, if has been modified by the transaction. A transaction is characterized by its atomicity property either it is

committed, i.e., completes successfully bringing all modified resources to their new, final state, or it is aborted, returning all modified resources to their initial state.

Distributed Transactions: A distributed transaction [3] is typically initiated with one participant, the transaction's root, and propagated to other participants. The other participants are being invoked recursively, inducing a communication structure that can be modeled as a tree. The participants and communication connections comprise the invocation tree's nodes and edges respectively. Two nodes with a common edge axe neighbors (each one is a neighbor of the other). A node with a single neighbor is a leaf in the tree. A distributed transaction is terminated via an AC protocol. The voting and the decision notification (either commit, or abort) are communicated between participants and a special participant, the CC, also through tree like communication structures, termination trees.

Two Phase Commit Protocol: In computer networking and databases two-phase commit protocol (2PC) is a distributed algorithm that lets all nodes in a distributed system to commit a transaction. The rule implemented in these results to either committing or aborting of all the different nodes involved in the transaction, whether the case is for the network failures or node failures. Somewhat, at a time the protocol cannot handle the failures which occur on more than one random site. The two different phases of the algorithm are the commit-request phase, in which all the cohorts are being prepared by the coordinator attempts, and second is the commit phase, in which the coordinator completes the transactions.

Consistency: Consistency in database system says that the requirement to the any given database transaction can only change affected data in given allowed ways. Any these data should be written to the database according to all valid defined rules which includes constraints, cascades, and triggers. The application programmer who might have wanted the correctness of the transaction in all the ways does not always guarantee for this correctness but because of any programming errors cannot result for the violation of different defined rules. In computer science, distributed shared memory systems or distributed data stores are the different consistency models which are used in distributed systems.

II. LITREATURE REVIEW

Different paper has many different policies for maintaining transactional consistency in the cloud. Das.S et al [4] has given the concept of ElasTraS which addresses in a cloud computing environment the issue of elasticity and scalability of the data store and provide ACID guarantees for different transactions which are limited to a single partition.

Wei.Z et al [5] proposes Cloud TPS which is a scalable transaction manager even in the presence of server failures and network partitions, guarantees about the ACID properties for different multi-item transactions which are being issued by Web application. Bigtable[6] and SimpleDB[7] are the two main families of scalable data layers which are being implemented in this paper.

Khadilkar.V et al [8] gave the different overview about security in storage and retrieval of data in cloud environment on presenting the system which allows cooperating organizations for securely sharing large amounts of data. By using Hadoop it ensures that the organizations have a large common storage area. Further, it has used Hive which is used to give the structured view of the data to present different users of the system and used the SQL language which enables them for querying different types of the data.

For the authorization based secure data transactions in cloud computing Vamsikrishna.V et al [9] identify multiple consistencies issue that can emerge during cloud host transaction process by using weak consistency model. He used the algorithm for maintaining the security in different transaction.

Iskander.M et al [10] has given the concept for maintaining consistency in the cloud data transaction by using Two-Phase Validation commit protocol which is the enhancement of the Two- Phase-commit protocol.

Abadi.D[11] discussed the different limitations and opportunities of moving the data management issues which depends on the today's rising different cloud computing platforms such as Amazon Web services, Microsoft Windows Azure, etc. In this paper it has been discussed that cloud computing platforms has become very much useful for large scale data analysis tasks, decision support systems, and application specific data marts than transactional and operational database systems and also gives the list of features for large scale data analysis tasks which ever are running on an Amazon style for the designing of a Database management system should contain.

Kraska.T et al [12] has proposed the work which present a new transaction model that not only define the guarantee of consistency allowed by the designers on the data at the transaction level, but also allows to automatically switch on to the consistency guarantees at runtime. It gives different techniques which on monitoring the data and gathering temporal statistics of the data which gives the system which adapt the consistency level dynamically.

Guo.H et al [13] has proposed work to make the knowledge available to the Database Management System can guarantee that the constraints are satisfied by the explicit consistency and currency constraints in different queries and develop techniques. This paper describes the model for expressing consistency and currency constraints which is used to define the semantics, and also gives the new SQL syntax.

Lakshman.A et al [14] has proposed Cassandra which is distributed storage system for managing very big quantity of structured data which are spread around different commodity servers, which provides highly available services without any single point of failure. Cassandra's main work is to perform its task on top of an infrastructure of thousands of different nodes. At this scale, small and large components fail continuously.

Weihai.Y et al [15] analyses the well-known Two Phase Commit Optimizations presumed commit and presumed abort, and presents an improved 2PC which is suitable for the Web services based on the applications. Specifically, for any distributed transaction the protocol allows each individual service provider to choose dynamically the most appropriate presumption. The protocol which is used does not give the extra overhead to the 2PC variants in terms of number of messages and log records, and it is easy to realize and understand.

III. METHODOLOGY

In the previous papers I studied various methods and different applications which are used for security purposes in transaction of data in cloud Enviornment.

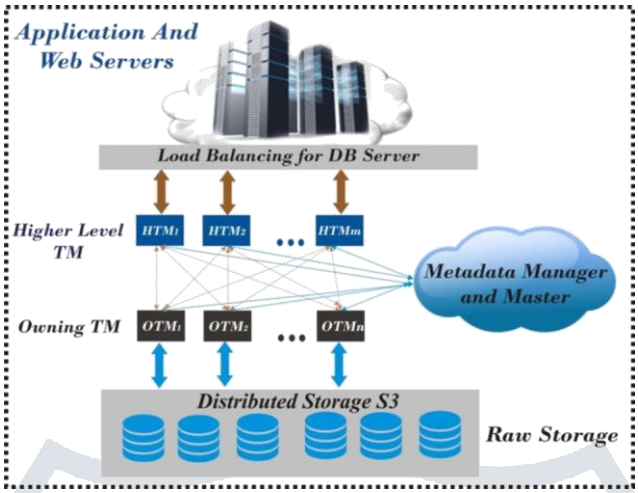

Figure 2: Overview of the ElasTraS system

Nowdays, some of the enhanced version of Indexed Sequential Access Methods are used for modern applications, and these are the basis for data model of modern scalable systems like Bigtable[6], Dynamo [16], PNUTS [17], SimpleDB[7] etc. In this paper [4] there is to choose a key-value based design which is as same as Bigtable whose values have application specified structure.

In the shown figure 2 systems has been designed according to the scalable manner which is used to provide transactional guarantees. The given above design of Elastras is used to provide ACID transactional guarantee which are limited to a single partition. They can provide efficient, scalable, and data store. Elasticity and scalability has been provided by the Elastras designing by reassignment of the dynamic partition which depends on the load of the system.

In the next paper there gives the concept of Cloud TPS [5] which shows how to support strict ACID transactions without compromising the scalability property of the cloud for Web applications. This work relies on few simple ideas.

- First, load of the data from cloud storage system into transactional layer.
- Second, splitting of the data across different number of LTMs, and replace them only for fault tolerance.

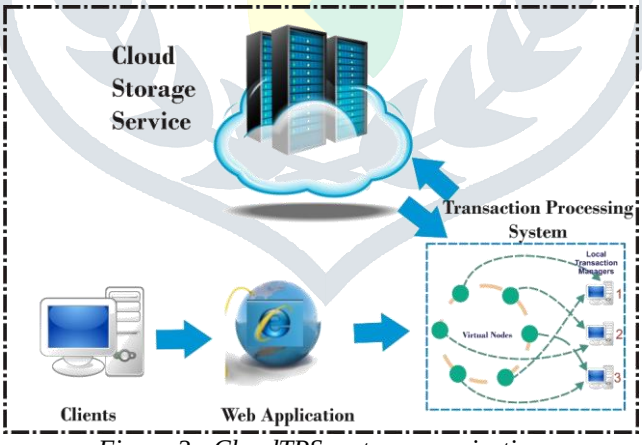

Figure 3: CloudTPS system organization

In the above given figure 3 shows the organization of cloud TPS. Here, Client gives HTTP request to web application, which as an output generates transactions to a Transaction Processing System. The Transaction Processing System (TPS) consists of different number of Load Transaction Managers (LTMs), where responsibility of each transaction depends for a subset of all data items.

In this paper is has been using the transactions which are implemented using the Two Phase commit protocol (2PC).In which in the first phase, the coordinator sends the requests to all the involved LTMs and asks to check that the execution of the operation can indeed been done correctly. If all the LTMs works properly, then the commitment of the second phase is done otherwise the transaction is aborted.

The next paper which is studied proposes the Two-Phase Validation Commit (2PVC) protocol which is executing on cloud servers which keeps the trustworthiness of transactions as a solution, which is a enhancement of the basic Two Phase Commit

(2PC) protocols. Here in the last analyzes the different approaches presented which uses analytical evaluation of the overheads and also uses simulations in order to guide the decision makers to which approach to use [10].

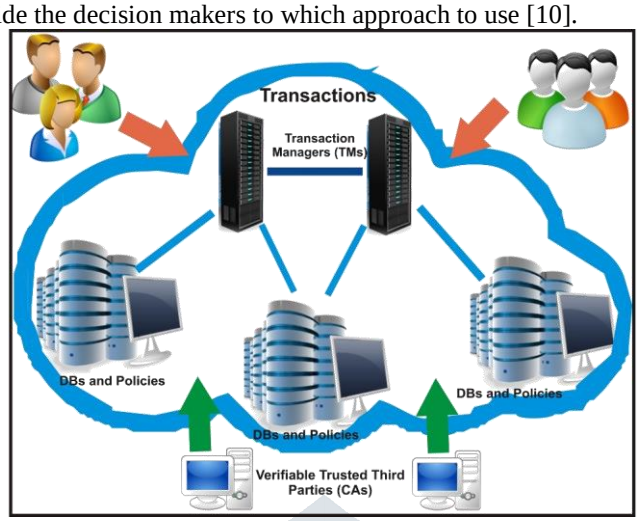

Figure 4: Interaction among the system components.

The given above figure 4 illustrates the interaction among the components in system assumes a cloud infrastructure which consists of set of servers, where each server's responsibility is to host a subset of all data items belonging to a specific application domain. By submission of different types of queries or different update requests which is encapsulated in ACID transactions through which users interact to the system. A Transaction Manager (TM) coordinates its execution on submitting transactions. As the workload of the system increases for balancing the load, multiple TMs could be invoked but only one TM is responsible for handling of the each transaction.

The next paper [9] which I had studied gives the information regarding data storage and retrieval of the data. Here, shows that the world produces large amount of data daily with the use of Internet or World Wide Web (w3) and the innovation of e-commerce applications and different social networking applications and organizations across the world. This data which comes from these different applications would results to be very much useful for cooperating organizations where sharing of the different data is possible. Two main obstacles in the process of sharing of data are on giving common spacing for storage and accessing securely the shared data. In this paper uses different layers each of the component layers as:

- The Web Application layer
- The ZQL Parser layer
- The XACML Policy layer
- The Basic Query Rewriting layer
- The Hive layer
- The Hadoop Distributed File System layer

The secured storage and retrieval of the data is achieved on combining these different layers. In this paper it has presented a system which allows large amount of data to be shared securely for different cooperating organizations. It has ensured that organizations have a large common storage area. Further, it uses Hive which is used to present a structured view of the data for different users of the system and with the usage of SQL language different users are able to query the data according to their wish and will.

The next paper [13] presents a number of techniques which on monitoring the data and gathering temporal statistics of the data which makes the system to follow the consistency level dynamically. Every service requests has a particular associated cost in cloud computing storage services. In cloud storage services, consistency is not only influences through the availability and per-formality of the systems but also through the overall operational cost.

To exhibit inconsistencies the optimization depends on allowing the database which helps in reducing the transactional costs but the penalty cost should not becomes higher. In this paper, different approaches are used to divides data into three different categories of consistencies: A, B, and C. In A category is used for ensuring strong consistency which guarantees and gives higher cost for each transaction. The C category ensuring for session consistency, which shows lower cost, but it will cause the result for inconsistencies. Depending on the particular specific policy data in the B category is used to handle by strong or session consistency.

IV. CONCLUSION

In distributed transaction database systems is deployed over cloud servers, in which entities cooperate to form proof of authorizations which are justified by collections of certified credentials.Cloud Computing uses the distributed transactions for their services. Many different techniques are yet been proposed for the secure transactions in the cloud which includes different methods and algorithm to maintain the security of the transaction of the data from the cloud databases. In this survey paper different transaction types and how to maintain their security are being studied.

REFERENCES

[1] Barrie.S,[2011], "Cloud Computing Bible", "Wiley Publishing",pp 24-34 .
[2] Marrinescu.D,[2013] "Cloud Computing: Theory and Practice", "Elsevier", Chapter 2,pp 21-28
[3] Raz.Y, "The Dynamic Two Phase Commitment (D2PC)Protocol", EMC Corporation,pp 162-169.
[4] Das.S et al, [2009] ,"Elastras: An Elastic Transactional Data Store in the Cloud," Proc. Conf. Hot Topics in Cloud Computing (USENIX HotCloud).
[5] Wei.Z et al, [2009], "CloudTPS: Scalable transaction for Web Applications in the Cloud" IEEE
[6] Chang.F et al ,[2006], "Bigtable: A Distributed Storage System for Structured Data," Proc. Seventh USENIX Symp.
[7] Amazon Web Services: Overview of Security Processes,http://aws.amazon.com/security/,June 2014.
[8] Khadilkar.V et al,"Secure Data Storage and Retrieval in the Cloud".
[9] Vamsikrishna.V et al,[2014] "Authorization Based Secure Data Transaction in Cloud Computing",International Journal of Engineering Trends and Technology (IJETT) .
[10] Iskander.M et al,[2014], "Balancing Performance, Accuracy, and Precision for Secure Cloud Transactions" ,IEEE.
[11] Abadi.D ,[2009], "Data Management in the Cloud: Limitations an Opportunities," IEEE , pp. 3-12.
[12] Kraska.T et al [2009], "Consistency Rationing in the Cloud:Pay only when it matters" ,ACM.
[13] H. Guo, P.-A. Larson, R. Ramakrishnan, and J. Goldstein, "Relaxed Currency and Consistency: How to Say "Good Enough" in SQL,"Proc. ACM Int'l Conf. Management of Data (SIGMOD '04), 2004.
[14] Lakshman.A et al ,[2010], "Cassandra- A Decentralized Structured Storage System," ACM , pp. 35-40.
[15] Weihai.Y et al[2004],"A Dynamic Two-Phase Commit Protocol forSelf-Adapting Services", Proc. of the IEEE International Conference on Services Computing.
[16] DeCandia.G et al[2007], "Dynamo: Amazons Highly Available Key-Value Store," Proc. 21st ACM SIGOPS Symp. Operating Systems Principles (SOSP).
[17] Cooper.B.F et al[2008], "PNUTS: Yahoo!'s Hosted Data Serving Platform," Proc. VLDB Endowmen, pp. 1277-1288.
[18] Lombardi.F et al [2012], "Secure virtualization for cloud computing, Elsevier.

AUTHOUR PROFILE

Shilpi Chaubey is currently an M.Tech Scholar in Computer Science Engineering Department specialization in Software Engineering Branch from Rungta College of Engineering & Technology. Her current research interests are in the areas of security, focusing on data transaction in Cloud Computing.