Fault Tolerance in Cloud Storage

Mrs. S. NITHYA ROOPA  
Assistant Professor  
Computer Science  
Kumaraguru College of Technology  
Coimbatore, India

G. SHANMATHI  
Student  
Computer Science  
Kumaraguru College of Technology  
Coimbatore, India

J. H. SURYEPRAKASH  
Student  
Computer Science  
Kumaraguru College of Technology  
Coimbatore, India

Abstract - Fault tolerance in cloud storage is to enhance data storage security during disaster, so that IaaS (Infrastructure as a Service) in NC Cloud methodology will be implemented here. NC Cloud computation power and storage capacity of cloud computing systems that allow scientists to deploy computation and data intensive applications without infrastructure investment, where big data application sets can be stored in the cloud. In pay-as-you-go model, storage strategies and benchmarking approaches have been mainly developed for cost-effectively storing large volume of generated application data sets in the cloud. However, they are either insufficiently cost-effective for the storage or impractical to be used at runtime. In this paper, toward NC Cloud achieving the minimum cost benchmark, we say highly cost-effective and practical storage strategy that can automatically decide whether a generated data set should be stored or not at the runtime in the cloud. The main aim of this strategy is the local-optimization for the tradeoff between computation and storage, while secondarily also taking users preferences on storage into consideration. At present, the remote monitoring system is growing very high due to the growth of supporting technologies and also in NC Cloud. The problem that may occur in remote monitoring such as the objects to be monitored and how quick how much amount of data to be transmitted to the data centre to be processed properly. This study proposes using a cloud computing infrastructure as processing centre in the remote sensing data. It focuses on the situation for sensing on the environment condition and disaster early detection. Where it has two important issue, especially in big cities that have many residents. This proposes to build the conceptual and also apotype model in a extensive manner from the remote terminal unit until development method for data retrieval. We propose using FTR-HTTP method to guarantee the delivery of the data from remote client to server. When the destruction occur the database architecture will transfer the database to the concern location assigned from the admin. So that data base can be saving exactly with the last fine transaction. Here data loss will not occur at any cost. This method is based on IP conflict procedure. So that roll backing process can also be possible. Using the same procedure of IP conflict method and this method will shows the data upto last minute transaction.

Keywords: Disaster Analysis, Highly Reliable, Dividing Data Into Fragments.

1. INTRODUCTION

There are several trends that are opening up the Cloud Computing, which is an Internet-based development and also use of computer technology. The cheaper and more powerful processors together with the software as a service (SaaS) computing architecture, are transforming data centers into large pools of computing service. The network bandwidth and the flexible network connections make it even possible that users now subscribe very high quality services from data and software that reside solely on remote data centers. Moving of the data into the cloud offers great friendly to user since they don’t have to care about the complexities of direct hardware management. The primary of Cloud Computing vendors, Amazon Simple Storage Service and Amazon Elastic Compute Cloud are both well known examples. While these internet based online services do providelarge amounts of storage space and customizable computing resources.

From the context of data security, which has always been an important aspect of quality of service, Cloud Computing gives new challenging security threats for number of reasons. Considering various kinds of data for each user stored in the cloud and the demand of long term continuous assurance of their data safety, the problem of verifying correctness of data storage in the cloud becomes even more challenging. Cloud Computing is not just a third party data warehouse. The data stored in the cloud can be frequently updated by the users, including insertion, deletion, modification, appending, recording. To ensure storage correctness under dynamic data update is importance. This dynamic feature also makes traditional integrity techniques futile and entails new solutions. The deployment of Cloud
Computing is powered by datacenters that are running in a simultaneous, cooperated and distributed manner. Individual users data is redundantly stored in multiple physical locations to further reduce the data integrity threats. Therefore, distributed protocols for storage correctness assurance will be of most importance in achieving a robust and secure cloud data storage system in the real world.

Most recently the importance of ensuring the remote data integrity has been highlighted by the following research works. These techniques are useful to ensure the storage correctness without having users possessing data, it cannot address all the security threats that are in cloud data storage, because they are all focusing on single server scenario and most of them do not consider dynamic data operations. As a complementary approach, researchers have also proposed distributed protocols for ensuring storage correctness across multiple servers to peers. But none of these schemes are aware of dynamic data operations. As a result, their applicability in cloud data storage can be drastically limited.

II. WORKING OF PROPOSED FRAMEWORK

A. Cloud Formation with NC Cloud.

The public cloud environment is the IaaS/PaaS Infrastructure or Platform as a Service that we rent from Linux (IaaS) or Microsoft (PaaS). Both are enabled for web hosting. Then, your SaaS stack will run under your Internet environment most likely in a virtualized one on your own equipment which would make it private. In this project we specialize in private cloud technology. Here we execute in a cloud environment. If strict security requirements go public or hybrid and if not, try the public or community cloud environment. So that here we are implementing a web services for the output purpose as well as the environment will be shown in actual while hosting the application. So finally SaaS can be fully utilized in cloud environment as IaaS/PaaS. Thus we formed cloud environment.

B. Disaster Analysis

Due to global warming our earth may face many types of disasters like earthquake, tsunami, storm, flood and etc. This disaster can be analysed through cloud remote monitoring. This module has main function to capture data from sensor both in digital or analog input. Package of specific sensors with Remote Terminal Unit will be placed in some places or objects prone to disasters. Cloud computing could be proposed as central of data processing to run service like service listener. It has function to capture and store information sent from the remote client. Otherwise, it could be used for the central data storage and application server to display the processed results to the user.

C. Data preservation using cloud service provider

This module deals with the software architecture of the cloud service provider, which is inter related with the remote disaster tool, so that when ever disaster will occur the cloud service provider will trigger out the malware process. This process may execute through Intranet, Internet and also through GPS. So that global communication will be possible here. This architecture should be assigned during the server configuration.

D. Executing IP conflict for TPA

The Cloud service provider will the triggering function with the TPA(Third Party Auditing). This module will take cares the database migration process. So that when ever disaster will occur the CSP will trigger through the IP conflict and the data base will be restored in the concern location assigned by the admin. Admin can customize the database by providing priority to the table sets. So that the transfer will works according to the assigned priority. This saves the database from data loss.

E. Roll over data

This module will execute after the disaster and CSP trigger out process. The roll back process too needs IP conflict procedure for analysing the failure calculation as the location of the database. According to the admin request original database can betransfer to the default location and also transfer of duplicate database also possible.

III. EXPERIMENTS AND DISCUSSIONS

We conducted experiments to evaluate our framework’s performance and simulated in larger scale networks using the bandwidth of 890Mbps. We deployed servers at four different locations. When the disaster is about to occur, In one method, we used the traditional way of performing backup, where we sent the entire database to a fixed location which was set by the user. In another method, we divided the data into fragments which was done based on priority set by the user and sent
those data to the locations as assigned by administrator. When comparing both the methods, data reliability was higher when using the first method. This was because, time taken by the first method was large since we transferred the entire database to a single location. This became even worse when we increased the size of the database. But while using second method, time taken to transfer entire database was drastically reduced since all the data were divided and sent to assigned locations parallelly. Even with increasing the size of the database, time taken was much better when compared with the first method. By doing this way, time taken to perform backup of database could be increased and also Data reliability can be promised.

<table>
<thead>
<tr>
<th></th>
<th>First Method</th>
<th>Second Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>500MB</td>
<td>342ms</td>
<td>296ms</td>
</tr>
<tr>
<td>1GB</td>
<td>1.1s</td>
<td>870ms</td>
</tr>
<tr>
<td>500GB</td>
<td>5 minutes 27secs 870 ms</td>
<td>3minutes 8 secs 620 ms</td>
</tr>
<tr>
<td>1TB</td>
<td>11 minutes 41 secs 789 ms</td>
<td>4minutes 53 secs 045ms</td>
</tr>
</tbody>
</table>

**IV. RELATED WORKS**

Some researchers proposed solutions for achieving higher reliability in dynamic networks. Dimakis et al. proposed several erasure coding algorithms for maintaining a distributed storage system in a dynamic network [1]. Leong et al. proposed an algorithm for optimal data allocation that maximizes the recovery probability [2]. Aguilera et al. proposed a protocol to efficiently adopt erasure code for better reliability [3]. These solutions, however, focused only on system reliability and do not consider energy efficiency.

[7], [8] distribute data and process the distributed data in a dynamic network. Both the distributed data and processing tasks are allocated in an energy-efficient and reliable manner, but how to optimally schedule the task to further reduce energy and job makespan is not considered. Compared with the previous two works, this paper propose a framework that reduces the job completion time and minimizes the energy wasted in executing duplicated tasks on multiple processor nodes. Further more, the tradeoff between the system reliability and the overhead, in terms of more storage space and redundant tasks, is analyzed.

**V. CONCLUSION**

In this paper, we propose a scheme to ensure the correctness of users’ data in the cloud. We distribute the files at various locations to provide redundancies and guarantee the data dependability. Through this way, we drastic reduce the communication and storage overhead as compared to the traditional replication-based file distribution techniques. Whenever data gets corrupted, it will be detected during the storage correctness verification, our scheme can almost guarantee the simultaneous localization of data errors, i.e., the identification of misbehaving servers. It is concluded that the proposed system works well and satisfy the owner and customers. The system is properly implemented and is tested very well and errors were properly fixed. The implemented site can be simultaneously accessed from more than one system. The framework can be also be added with other data mining techniques such as nearest neighbours, neural network and general algorithm to find some interesting patterns in the data base.

**VI. REFERENCES**


