ISSN: 2277-9655
Scientific Journal Impact Factor: 3.449
(ISRA), Impact Factor: 1.852

IJESRT
INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

Peer to Peer Distributed Data Storage with Security in Cloud Computing

Dilip Reddy.B,[*1] Dr N.Kasiviswanath[2] ,Dr S.Zahoor Ulq Huq[3]
[*1]Assistant Professor, Department of CSE,G.Pulla Reddy Engineering College, Kurnool, Andhra Pradesh, India.
[2]Professor & Head of CSE. G.Pulla Reddy Engineering College, Kurnool, Andhra Pradesh, India.
[3]Professor Dept of CSE, G.Pulla Reddy Engineering College, Kurnool, Andhra Pradesh, India
b.dilipkumarreddy@gmail.com

Abstract

Cloud computing is a new computing paradigm that attracted many computer users, business, and government agencies. Cloud computing brought a lot of advantages especially in ubiquitous services where everybody can access computer services through internet. In recent years, the technology of cloud computing has been widely applied in e-business, e-education and etc.Cloud computing platform is a set of Scalable large-scale data server clusters, it provide computing and storage services to customers. The cloud storage is a relatively basic and widely applied service which can provide users with stable, massive data storage space. Our research shows that the architecture of current Cloud Computing System is central structured one, all the data nodes must is central structured one, all the data nodes must be indexed by a master server which may become bottle neck of the system.So to over come bottle neck problem we are implementing the number of servers and each server is said to be an chunk. Each file may be partitioned into several parts and stored in chunk Server. Each chunk is stored in Remote machine. Meanwhile, the security of the system must be ensured with algorithm called TWOFISH. Confidentiality, availability and integrity are the main keys for a secure system. We propose new cloud storage architecture based on P2P and design a prototype system. By these we are going to overcome the problem bottle neck. The system based on the new architecture has better scalability and fault tolerance. This research paper is about a cloud storage architecture based on P2P with fault tolerance. Here the central entity is removed and all servers are interconnected to form a ring structure. When one of the servers fails, the work will be taken over by any of the best performing servers.

Keywords: Cloud computing, p2p, Cloud storage, Security, Secure data transmissions.

Introduction

A cloud computing platform dynamically provisions, configures, reconfigures, and provisions servers as needed. Servers in the cloud can be physical machines or virtual machines. Advanced clouds typically include other computing resources such as storage area networks (SANs), network equipment, firewall and other security devices. [1] This paper will focus on the storage service from cloud. Some typical cloud systems, such as GFS of Google[2], Blue Cloud of IBM[1], Elastic Cloud of Amazon[3], have a similar architecture for storage. In the system architecture, there is a central entity to index or manage the distributed data storage entities. It is effective to simplify the design and maintenance of the system by a central managed architecture, but the central entity may become a bottleneck if the visiting to it is very frequent. Although systems in practice have used some technique as backup recovery to avoid the probably disaster from the central bottleneck,the flaw come from the architecture has not resolved essentially.We propose a cloud computing architecture based on P2P which provide a pure distributed data storage environment without any central entity. The cloud based on the proposed architecture is self-organized and self-managed and has better scalability and fault tolerance. Rest of the paper is organized as follows; we will introduce some related work about cloud storage system and P2P storage system. In section 3 of this paper, we describe a typical scenario to explain the architecture of our proposed cloud computing storage environment. There is an introduction on our prototype about the P2P cloud system.

ISSN: 2277-9655
Scientific Journal Impact Factor: 3.449
(ISRA), Impact Factor: 1.852

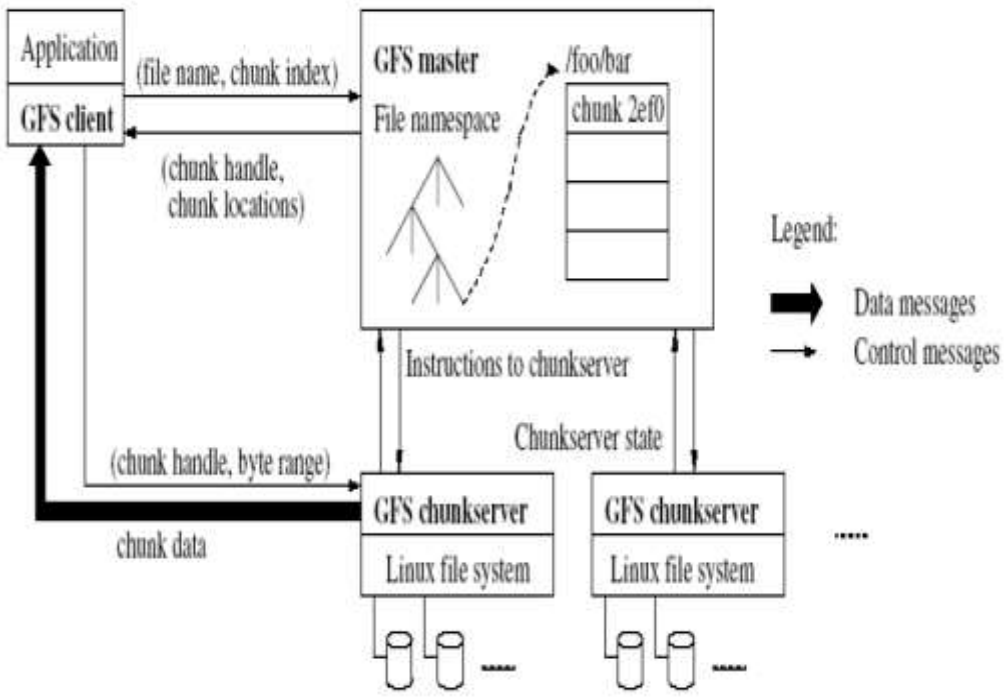
Literature Survey

In this section,we will introduce some related work about cloud system and P2P products for storage.

The first to give prominence to the term cloud computing (and maybe to coin it) was Google's CEO Eric Schmidt, in late 2006[4]. Google Inc. has a proprietary cloud computing platform [5] which was first developed for the most important application of Google search service [6] and now has extended to other applications. Google cloud computing infrastructure has four systems which are independent of and closely linked to each other.

They are Google File System for distributed file The GFS above is actually a central indexed distributed storage system GFS master work as an index server which can provide the global information about each chunk server for clients. The flaw of central index architecture is that the GFS master may become bottle neck of the system since all the request to the target data chunk must be originated from the index server which burdens the master. The literature identifies three different broad service models for cloud computing:

a) Software as a Service (SaaS), where applications are hosted and delivered online via a web browser offering traditional desktop functionality for example Google Docs, Gmail and MySAP. b)Platform as a Service (PaaS), where the cloud provides the software platform for systems (as opposed to just software), the best current example being the Google App Engine.c) Infrastructure as a Service (IaaS),where a set of virtualizedcomputing resources, such as storage and computing capacity, are hosted in the cloud.



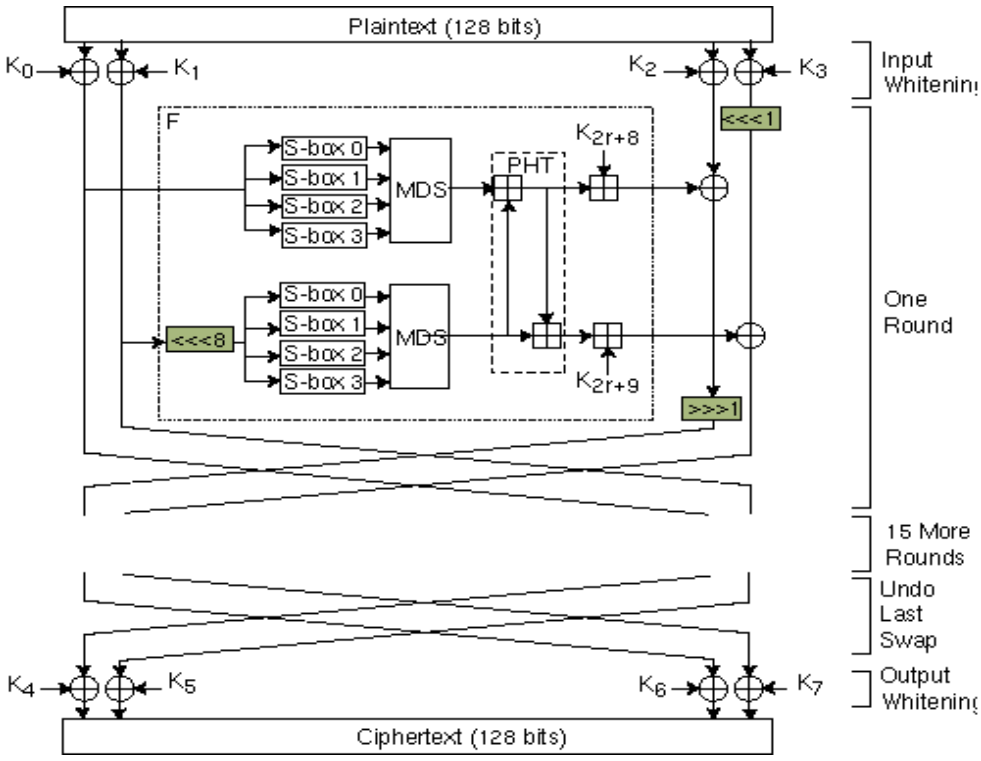
Architecture of Google File System

The distributed P2P network indexed by storage, MapReduce program model for parallel Google applications[7], Chubby for distributed lock mechanism[8] and BigTable for Google large-scale distributed database[9].shows the architecture of Google file system. A GFS cluster consists of a single master and multiple chunk servers and is accessed by multiple clients. Chunk servers store chunks on local disks as Linux files and read or write chunk data specified by a chunk handle and byte range. The master maintains all file system metadata. This includes the namespace, access control information, the mapping from files to chunks, and the current locations of chunks. When a client wants to visit some data on a chunk server, it will first send a request to the Master, and the master then replies with the corresponding chunk handle and locations of the replicas. The client then sends a request to one of the replicas and fetch the data wanted.

ISSN: 2277-9655
Scientific Journal Impact Factor: 3.449
(ISRA), Impact Factor: 1.852

Proposed method

Main aim of Proposing these paper is to reduce the bottle neck of our current distributed system. Through these we improve our transmissions, accessibility, and performance.The main changeling issue is too adopted is Security. Totally Security is a lack of issues in Distributed System.Security measures assumed in the cloud must be made available to the customers to gain their trust. There is always a possibility that the cloud infrastructure is secured with respect to some requirements and the customers are looking for a different set of security. The important aspect is to see that the cloud provider meets the security requirements of the application and this can be achieved only through 100% transparency. We are implemented some of Security Requirements with help of TWOFISH Algorithm.Twofish is a symmetric key block cipher with a block size of 128 bits and key sizes up to 256 bits.



Proposed Architecture for Twofish Algorithm

In order to have secure cloud system, the following aspect must be considered:

Authentication: Authentication is the process of verifying a user or other entity's identity. This is typically done to permit someone or something to perform a task. There is variety of authentication system, some are stronger than others. A strong authentication system ensures that the authenticators and messages of the actual authentication protocol are not exchanged in a manner that makes them vulnerable to being hijacked by an intermediate malicious node or person. That is, the information used to generate a proof of identity should not be exposed to anyone other than the person or machine it is intended for.

Authorization: Authorization is when the system decides whether or not a certain entity be allowed to perform a requested task. This decision is made after authenticating the identity in question. When considering an authentication system for a particular application, it is crucial to understand the type of identifier required to provide a certain level of authorization. \

ISSN: 2277-9655
Scientific Journal Impact Factor: 3.449
(ISRA), Impact Factor: 1.852

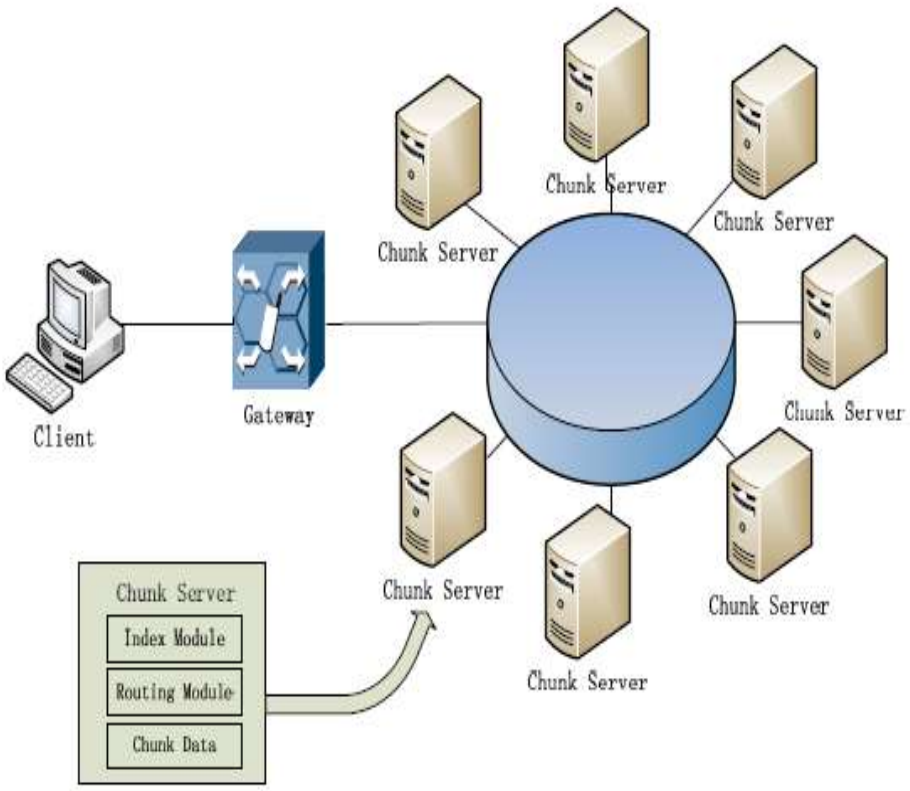
Confidentiality: Confidentiality is needed when the message sent contains sensitive material that should not be read by others and therefore must not be sent in a comprehensible format. A loss of confidentiality is the unauthorized disclosure of information. Confidentiality, as it relates to security and encryption techniques can be obtained by

encrypting messages such that only intended recipient are able to read them.

Integrity: Integrity is ensuring that the data presented are true and valid master source of the data and includes guarding against improper information modification or destruction to ensure information non-repudiation and authenticity. A loss of integrity is the unauthorized modification, insertion, or destruction of information. One way of ensuring of data integrity is by using simple checksums which prevent an attacker from forging or replaying messages.Checksum is usually implemented when the channel between communication parties is not secure and ensure that the data has reached its destination with all bits intact, if bits have been modified that the modification will not go unobserved.

Proposed system

Architecture: We design a new system of P2P storage for cloud platform, which can take advantage of the P2P distribute architecture and do well in concurrent update. Following shows the architecture of the system.



Figrue 2 Cloud Storage Based on P2P

Roles involved in our architecture can be defined as follows and illustrated as below.

Client App: The client application is designed to get the data from the platform. Here the client sends user name and password for getting authentication. The authentication for client access is given, if and only if both the user name and password matches to one of the details in database. Else access is denied. After authentication the client send request to gateway.

Gateway: This entity can transfer the request or

ISSN: 2277-9655
Scientific Journal Impact Factor: 3.449
(ISRA), Impact Factor: 1.852

response between the Client App with the network and can lead the request to the nearest node in the network. This is the important module which acts as an intermediary between the client and the Chunk server. It receives the client's request and forward the request to the nearest chunk server and then it receives the response messages from the chuck server and orward that message to corresponding client/requester.

Chunk Server: This entity is served as the data resource node and P2P node. Different with the function of pure data storage in GFS, the chunk server here has three function modules with separated interfaces.

Conclusion and future work

We propose a cloud computing architecture based on P2P which provide a pure distributed data storage environment without any central entity for controlling the whole processing. The advantage of this is architecture is that it prevents the bottleneck problem that arises in most of the client server communications.

The proposed system does its operation based on the performance of the system. It does the monitoring operation to find out the best chunk servers within the P2P network. It does this operation in order to perform efficient resource utilization and load balancing of the servers.

The future work of this proposed system could to modify the system performance by reducing the number servers present in the network. It a tough job to manage a lot number of servers. The enhancement says that if the operation is performed, for example, with the help of 100 servers, then reduce the number of servers to 50 servers by increasing the capacity of each server. Then pipelining concept could also be introduced within this P2P network in order to provide faster access. By enabling all these concepts the architecture provides better scalability, manageability, fault tolerance, better performance.

References

1. *Boss G, Malladi P, Quan D, Legregni L, Hall H.Cloud computing. IBM White Paper, 2007.*
2. *Ghemawat S, Gobi off H, Leung ST. The Google file system. In: Proc. of the 19th ACM Symp.Operating Systems Principles. New York: ACM Press, 2003. 29_43.*
3. *Amazon. Amazon elastic compute cloud (Amazon EC2). 2009. 4 Francesco Maria Aymerich, Gianni Fenu,Simone Surcis. An Approach to a Cloud Computing Network.*
4. *Barroso LA, Dean J, Hölzle U. Web search foraplanet: The Google cluster architecture. IEEE Micro,*
5. *Brin S, Page L. The anatomy of a large-scale hypertextual Web search engine. Computer Networks,*
6. *Ben Y.Zhao, John Kubiatowicz, and Anthony Joseph,"Tapestry: An Infrastructure for Fault-tolerant Wide-area Location and Routing", Technical Report No.UCB/CSD - 01-1141, University of California Berkeley.*
7. *Francesco Maria Aymerich, Gianni Fenu, Simone Surcis.An Approach to aCloud ComputingNetwork.978424426249/08/$25.0 0 ©2008 IEEE conference.*
8. *Antony Rowstron and Peter Druschel, "Pastry: Scalable, decentralized object location and routing for large-scale peer-to-peer systems", IFIP/ACM International Conference on Distributed SystemsPlatforms.*

Author Biblography

Dilip Reddy
Workinas an Assistant Professor in Departmentof CSE,G.Pulla Reddy Engineering College, Kurnool.
Email:b.dilipkumarreddy@gmail .com

DR N.Kasiviswanath
Head Of Departmentof CSE,G.Pulla Reddy Engineering College,Kurnool.
Email:gprechodcse@gmail.com

Dr S.Zahoor Ulq Huq

Professor in Department of CSE,G.Pulla Reddy Engineering College,Kurnool**.**

Email:gprecacoe@gmail.com