

Analysing the Cloud Data Storage Technology and Its Architecture Applications

Rahul – PG Scholar

Mahendra Kumar – Professor

Department of Master of Application

Dayananda Sagar College of Engineering

Abstract

The idea of using the internet to store data in the cloud is becoming more popular these days. Cloud computing data storage is the most crucial and significant research topic. The article begins with an introduction to cloud computing, storage in the cloud, and architecture for storage in the cloud. GFS (Google towards genuine corporate examples) and HDFS (Hadoop Distributed File System) are the cloud data storage technologies that we examine next (Hadoop Distributed File System) With that, we come to the end of our series. We show how to enhance the conventional method of storing files using the eyeOS Web operating system. Distributed storage and fault-tolerant control are provided by HDFS, a component of Hadoop.

Keywords: “Cloud Computing; Cloud Storage; Web Operating System; Distributed File System”

Introduction

Currently, the notion of cloud computing is gaining in popularity at an unprecedented rate. It is common practise in cloud computing to employ a distributed approach for business model.

The likes of Google, Amazon, IBM, Microsoft, and Sun are all working on cloud computing solutions at the same time as well. GFS, MapReduce, and BigTable are all examples of this, as is Google's marketing of application engines based on the GFS, MapReduce, and BigTable technologies. A wide range of terms and ideas are discussed, such as the terms "cloud computing," "cloud storage," and "cloud storage architecture." The eyeOS Web operating system on our PC will be used to develop cloud storage architecture once we've studied cloud storage technologies like GFS and HDFS (Hadoop Distributed File System).

Cloud computing and cloud storage

Cloud computing definition

Grid computing, distributed computing, parallel computing,

Examples of cloud computing technologies include utility computing and virtualization, both of which use conventional computer and network technology together. As processing power in the cloud grows, cloud computing reduces the amount

of processing users must do on their own devices. In the end, user terminals will be nothing more than a collection of input and output devices.

Cloud storage definition and Architecture

The phrase "cloud storage" refers to a service that provides both data storage and access to the cloud for both businesses and consumers. Cluster applications, grid approaches, distributed file systems, and other comparable features are used to connect a broad variety of storage devices. Both a little amount of storage in the cloud and a huge amount of cloud storage may be called "cloud storage" under the same umbrella word. Storage, the basic management interface, and the access interface make up the architecture of a cloud storage system.

Cloud storage technology and Enterprise

GFS -System Architecture

As seen in the diagram, a GFS cluster comprises of a single master, several chunk servers, and numerous clients. Each one of these machines runs a variant of the Linux operating system [1].

- Managing the directory structure and the file system is the responsibility of the Master. GFS utilises a single master policy to save coordination costs, implying that only one master supplies services at any one time. Clients and masters only exchange metadata; all other data is transmitted to the chunk servers directly. With a default block size of 64M per chunk server, each chunk server stores GFS files in chunks of a predetermined size. When a chunk is formed, the master immediately assigns it an immutable and globally unique 64-bit chunk handle to use going forward. There are three separate chunk servers for each block.
- Various sections of the file namespace may be replicated at different levels. Several pieces and servers are displayed, including five chunks. To store a piece, you'll need three separate chunkservers.
- Each application's GFS client code connects to the master and chunkservers using a separate implementation of the file system API. It's only that chunkservers are responsible for all data-related communications.

Work flow

Clients and chunkservers communicate with each other through dashed lines, whereas chunkservers communicate with each other using thick lines of data.

Clients execute a computation based on the file structure and chunk size before providing the file name and chunk index. A customer's master provides them with chunk handles and chunk locations. When the procedure is complete, the chunk handle and byte range are sent to the nearest chunk server. The client receives the data from the chunk server. There is no way to permanently save the mapping of chunk server to chunk in the Master.

Instead, when the master is first started or a new chunk server is added, it asks each chunk server about its parts. To exchange instructions and

gather data, each chunk server gets frequent Heart Beat signals from the master.

HDFS

The Apache Software Foundation, which houses Hadoop, provides assistance for open source software projects. Complementary services or higher-level abstractions may be provided by other Hadoop subprojects, although MapReduce and its distributed file system (HDFS) are best known. The contents of [4] are laid down in this document. Hadoop Distributed File System (HDFS) is its full name. HDFS runs on commodity hardware in the same way that Google's GFS does. Multiple data nodes in an HDFS cluster are connected to one namenode (the name server). The "master" server in GFS is referred to as Namenode. HDFS files may be stored in one or more blocks on a datanode. It is possible to run Namenode and datanodes on a low-cost Linux machine.

Cloud storage architecture based on Hadoop - EyeOS

Office applications and personal information management systems are included in the EyeOS web desktop environment, which allows users to store data online and work from any location. Because of the lack of fault-tolerant backups and insufficient reliability, EyeOS' document management is confined to a single server environment. The performance of file access is poor since it is done in a single thread. HDFS technology is used in this study to improve the basic file storage strategy and enable distributed storage and fault-tolerant management of files.

System implementation -Architecture

EyeOS, a cloud server (NameNode), and an eyeOS cloud storage centre are all shown in this diagram of the storage system we built (DataNode). Users access the cloud storage using a web browser that is pre-installed on each device. Clients are the means through which cloud storage systems and their users interact.

- Web Operating System: The web operating system receives user access requests, checks the legitimacy of the users, and communicates directly with clients. It is built on eyeOS, which provides

users with a huge variety of apps. Users can download the software they need to create a customised system. EysOS also serves as a file access interface for users, allowing them to store data to cloud storage clusters.

- Cloud server (Cloud NameNode): A Hadoop-based cloud storage cluster contains a cloud server (NameNode) and a cloud storage centre (DataNode). In Hadoop, the cloud server is the namenode that administers the file system namespace, computes the file-to-datanode mapping, allocates datanodes to save file blocks, and restricts external client access.
- As the name implies, a Hadoop datanode is a cloud storage centre. In addition to preserving files, it implements file distributed storage, distributes load balancing, and ensures that files are fault-tolerant.

Operation Process

Two of the most often used eyeOS operations are file writing and file reading. You must first download the file to your local computer before you can manage or view it in an application. software that runs on the internet When a file has been edited and saved, the web operating system uploads it to the server.

PC-to-cloud data backup and synchronisation Reading a file is done as follows: 1 Double-clicking a file icon on the web OS brings up the web OS in the user's browser. EyeOS then requests the file from the namenode of Hadoop.

Namenode uses a search for file-related information to locate a specific file. The datanodes that hold the file blocks send them to the clients. 3Clients download and merge individual file blocks into a single file. 4 Programs linked to a file open and display the file in a web-based operating system. A file's creation procedure: 1 In order to edit and save files on the web OS, users must log in using a web browser on their client computer. Files must be posted to the Hadoop namenode in order for EyeOS to use them. 2Namenode allocates storage space to datanodes depending on the file size and storage condition of the datanodes after receiving an uploading request. A file is sent

from the client's computer to the server. A Namenode divides the data into blocks, which are then stored on the designated datanodes.

Experiments

Five computer systems are used in these investigations. Using datanodes instead of clients, eyeOS, and the namenode is the preferred ,method. The datanodes are assumed to be Da and Db. Da and Db contain files X and Y. Da and Db are always normal throughout trials. Table 1 shows that when File1 is produced, it is simultaneously stored in Da and Db.

Db is always OK, but Da is always invalid, as seen in Table 1's Line 3. Table 1 shows the findings on Line 4 when Da is back to normal. During the generation of the file, it is possible to locate File1 in Db, but not in Da. We may also look at File1 to check whether Da is back to normal. If you delete a file from Db and then delete it again from Db, it will reappear in Da. It will be too late if everything goes back to normal.

Conclusions

As the internet's reach grows, so does the need for increasingly powerful online applications, which in turn necessitates the rise of cloud computing. The area of cloud data storage, which deals with how data is stored on the cloud, is an important one. In this work, we cover cloud computing and cloud storage principles. We then set up a cloud storage architecture on our PCs using eyeOS, an online operating system. The system has been tested and shown to be effective.

References

- [1] "Cloud storage made easy: Securely backup and share your files" , James Bernstein ,2018
- [2] Cloud Computing Principle and Paradigm , James Broberg,2010
- [3]Cloud Computing :Architect and Application,2019
- [4] Cloud computing Theory and Practicle ,Dan Cristian Marinescu,2013
- [5] Cloud computing : concept, Theory and technology,Ricardo Putini,2013