

Cloud Storage Based on Key-Value Store

^[1]Abhishek Pandhare, ^[2]Mayur Patil, ^[3]AkshayPandure, ^[4]Yogesh Pawar

^[1]abhishek.pandhare54@gmail.com, ^[2]mayur750patil@gmail.com, ^[3]akshay.pandure321@gmail.com,
^[4]pyogesh924@gmail.com

BansilalRamnath Agarwal Charitable Trust's
Vishwakarma Institute of Information Technology

Abstract :- — Nowadays, cloud-based storage services are rapidly growing and becoming an emerging trend in data storage field. There are many problems when designing an efficient storage engine for cloud-based systems with some requirements such as big-file processing, lightweight meta-data, low latency, parallel I/O, deduplication, distributed, high scalability. Key-value stores played an important role and showed many advantages when solving those problems. This paper presents about Big File Cloud (BFC) with its algorithms and architecture to handle most of problems in a big-file cloud storage system based on key value store. It is done by proposing low-complicated, fixed-size meta-data design, which supports fast and highly-concurrent, distributed file I/O, several algorithms for resumable upload, download and simple data deduplication method for static data. This research applied the advantages of ZDB - an in-house key value store which was optimized with auto-increment integer keys for solving big-file storage problems efficiently. The results can be used for building scalable distributed data cloud storage that support big-file with size up to several terabytes.

Keywords: - Cloud Storage, Key-Value, NoSQL, Big File, Distributed Storage

I. INTRODUCTION

Cloud based storage services are rapidly growing and becoming an emerging trend in data storage field. The main basic idea of this project is to handle most of problems in a big file cloud storage system based on key value store. It is done by proposing low-complicated, fixed size Meta data design which supports fast and highly concurrent, distributed file I/O for resumable uploading and downloading the file. We propose new big-file cloud storage architecture and a better solution to reduce the space complexity of meta-data. Key-Value stores have many advantages for storing data in data intensity services. They often outperform traditional relational databases in the ability of heavy load and large-scale systems.

II. MOTIVATION

Motivation of this project is to use cloud computing which handles all big file cloud and store the file in cloud and also deduplication method is used so that no same file exists in the cloud. Also Key value processing is done so that file can be easily handle on cloud. Also Distributed Storage system is idea provided in this project. The purpose of distributed file system is to provide backup and retrieve data functions. The Main motive is to form a simple Meta data to create a high performance cloud storage based on key value store. Every big-file stored in BFC is

spited into fixed size chunks. The chunks have a contiguous ID range.

III. EXISTING SYSTEM

People use cloud storage for the daily demands, for example backing-up data, sharing file to their friends via social networks such as Face book [3], Zing Me [2]. Users also probably upload data from many different types of devices such as computer, mobile phone or tablet. After that, they can download or share them to others. System load in cloud storage is usually really heavy. Thus, to guarantee a good quality of service for users, the system has to face many difficult problems and requirements.

IV. TECHNOLOGY USED

3-Tier Client-Server Architecture

- 1) **Server:** Apache Tomcat [5], often referred to as tomcat server is open-source java servlet container developed by the ASF Apache Software Foundation. Tomcat implements several Java EE specifications including Java Servlet, Java Server Pages (JSP) and provides a pure Java HTTP web server environment in which Java code can run.
- 2) **Database:** MySQL is an open-source relational database management system (RDBMS). MySQL is a central component of

the LAMP open-source web application software stack. MySQL is also used in many high-profile, large-scale websites, including Google Facebook, Twitter, Youtube.

- 3) **Web Technologies:** HTML, CSS, JSP, ServletHyperText Markup Language (HTML) is the standard markup language for creating web pages and web applications. HTML describes the structure of a web page semantically and originally included cues for the appearance of the document.

Cascading Style Sheets (CSS) is a stylesheet language used for describing the presentation of a document written in a markup language. JavaServer Pages (JSP) is a technology that helps software developers create dynamically generated web pages based on HTML, XML or other document types. Released in 1999 by Sun Microsystems, JSP is similar to PHP and ASP, but it uses the Java programming language.

- 4) **IDE:** Eclipse/Netbeans: Eclipse is an integrated development environment (IDE) used in computer programming, and is the most widely used Java IDE. Integrated development environments are designed to maximize programmer productivity by providing tight-knit components with similar user interfaces. IDEs present a single program in which all development is done.
- 5) **Apache Tomcat Sever.**

V. FUTURE SCOPE

The predictions of cloud in the future are as follows: Developers are trying to build many new software for the cloud, by this it is predicted that there will be no differentiation between public and private clouds. We can also predict that in the Coming years. Almost a quarter of all applications will be accessible on the cloud. This can be estimated when you look at the fact that approximately 60 percent of enterprises will spend More than 10 percent of their budget on cloud services (predicted by IBM). Nowadays, cloud based services are rapidly developing and turned into a rising pattern in information technology field. Due to this, enterprises are relying more and more on cloud service in order to develop, market and sales of products, supply chain management. It is also estimated that Software as a Service (SAAS) would grow at an outstanding rate with yearly growth rate of 20.2 percent. This implies that it will be growing from \$18.2 billion in 2012 to \$45.6 billion in 2017.

In the future, there will be increase in implementation of hybrid clouds. It is estimated that 50 percent of the enterprise will implement hybrid clouds by 2017. The hybrid cloud which is a combination of on and off premises will provide the best of both worlds: integration of strengths permitting organizations so as to attain the performance of on premises solutions further also the management assistance of the cloud business model . One such example of dynamic cloud is connection of new apps to already existing systems like operation new mobile social apps on clouds.

There will be increase in the development for the cloud going forward. According to Evans Data Corporation at present there are more than 18 million software developers globally so far less than 25 percent are developers for cloud. It is estimated that as cloud continues to develop rapidly, more developers will develop for cloud it becomes evident when you look at the fact that 85 percent of the new software being built today is for cloud computing. Going forward, it is recommended that there will be growth in third party, commercial and enterprise developers and contributions to cloud application systems.

It is estimated that with rise in competition in the cloud storage systems will lead to better products, services and innovation. When developers work collectively as a group, focusing on the quality of work they should deliver, innovative cloud services will come into existence because of code collaboration and dynamic developer ecosystems. End users are trying to drift utilizing separate physical servers which are connected to storage area networks to more composed reference architectures which would include every component essential in order to run their applications. Previously, servers and storage are sold as independent IT infrastructure elements. The shift to the cloud has changed all of that discussed above. Since everything is converged into one single solution and virtualized on a converged network, there might not be any need for a conventional storage only network.

VI. REFERENCES

- [1] Dropbox tech blog. <https://tech.dropbox.com/>. Accessed October 28, 2014.
- [2] Zing me. <http://me.zing.vn>. Accessed October 28, 2014
- [3] Facebook. <http://facebook.com>, 2014
- [4] Federal Information Processing Standards Publication 197 ADVANCED ENCRYPTION STANDARD (AES) November 26, 2001
- [5] I. Drago, E. Bocchi, M. Mellia, H. Slatman, and A. Pras. Benchmarking personal cloud storage. In Proceedings of the 2013

- conference on Internet measurement conference, pages 205–212. ACM, 2013.
- [6] I.Drago, M.Mellia, M. M Munafo, A.Sperotto, R.Sadre, and A. Pras. Inside dropbox: understanding personal cloud storage services. In Proceedings of the 2012 ACM conference on Internet measurement conference, pages 481494. ACM, 2012.
- [7] S. Ghemawat, H. Gobioff, and ST. Leung. The google file system. In ACM SIGOPS Operating Systems Review, volume 37, pages 2943. ACM, 2003.
- [8] Thanh Trung Nguyen, Tin Khac Vu, Minh HieuNguyen ,LeQuy Don Technical University, Ha Noi, Distributed Big File Cloud Storage based on key value store Viet Nam †VNG Research, Viet Nam
- [9] International Journal of Advanced Research in Computer Science and Software Engineering Volume 5, Issue 12, December 2015.
- [10] Yan Kit Li, Xialofeng Chen, Patrick P.C.Lee Secured Authorized Deduplication Based Hybrid Cloud Approach.
- [11] Oracle<http://www.oracle.com/technetwork/articles/java/compress-1565076.html>
- [12] SDLC [http:// www.veracode.com/ security/ softwaredevelopmentlifecycle](http://www.veracode.com/security/softwaredevelopmentlifecycle)

