

Storage and Backing Up Of Data in a Cloud Environment Using File System

Shruti Sigi

M.tech Student

Dept of Computer Science and Eng LAEC Bidar

EMAIL ID: shrutisigi35@gmail.com

Nirupama Sheelwant

Assistant Prof

Dept of Computer Science and Eng LAEC Bidar

EMAIL ID: nirupama.sheelwant@gmail.com

ABSTRACT

-Today's computational advances have created numerous applications for I/O operations have the capacity to store more data. The authentic need is accessible of information from anyplace. This gets to be issue for some clients who use applications, for example, databases, media, and hard disks, and so forth. The I/O data solicitations of these applications get higher as they get greater. To beat these issues we present a cloud domain for support and data stockpiling utilizing parallel document framework. That have the capacity to store and reinforcement of information through removed server cap can be open through web association. The usage intends to expand the availability of data and reduce in loss of information.

Keywords-Cloud, backup, Storage, PVFS2, SAN.

1. INTRODUCTION

We have many applications to store our data due to the growth of the technologies .When we store our data to local devices we cannot access data remotely or distantly. But main genuineness is to access data remotely it means anywhere or anytime. This has given rise to seek alternatives to solve this problem. One of the alternatives is known by the term cloud computing:Cloud computing is a universal term that deliver the hosted services through internet. These services are generally divided into three categories: Software-as-a-Service (SaaS) Infrastructure-as-a-Service (IaaS), Platform-as-a-Service (PaaS) and. The name cloud was inspired by the symbol that is often used to represent the Internet in flowcharts or diagrams. The most recent and accepted standardized definition of Cloud Computing is the one by the National Institute of Standards and Technology (NIST) [1]: “Cloud computing is a model for enabling ubiquitous, convenient, on demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. This cloud model promotes Availability and is composed of five important characteristics, three service models, and four deployment models.” Five important characteristics of cloud computing are: Services on demand, Pay per use,

Broad network access, Resource pooling and rapid elasticity. Some Internet service providers like Google, Amazon AWS, Microsoft and others provide services as a public. They build their own infrastructure to offer services to user but these services are managed and operated by third parties not related to the organization. In this case data will not be confidential but companies need high data security. To overcome this problem another cloud is Private cloud. This type of cloud is a good alternative. Administration is carried by a client that switches which applications should run. Servers, network, and storage devices are the property of the organization. Hybrid cloud is combination of these two clouds.

We have many applications to load and store large amount of data. For example hard disk, database etc. Earlier, to improve the performance of nodes of disk are used storage area network (SAN) to Perform I/O operations. But cost of SAN is much. In order to improve the performance and cost, we introduce a file system called parallel virtual file system (PVFS2) version 2 on a private cloud. In its second version it reuses the existing infrastructure such as server, cluster etc. This file system stripes data of file over multiple servers so that it allows performing I/O operations in a parallel manner. So that It reduces the access time of a data and loss of information. This option avoids the usage of expensive SAN.

2. LITERATURE SURVEY

Cloud computing and cloud storage have become more useful for delivering online services. While there are some cloud services that offers free storage services up to some gigabyte to terabyte or some type of subscription based. Goggle provides services to store or load data. Dropbox [2] allows user to share entire folder to other user and it is viewable by all members in team. It offers only 2GB of free storage space for further usage user has to pay for it. We can say ‘pay per use’ basis. In cloud environment there are some file systems for backup and storage of files within input/output [3, 4, 5, 6, 7].Network File System [3] it is a ‘ubiquitous’ in nature. And it is an open source. NFS has performance problem .here single server manages data as well as metadata. Due to this reason performance of the system will reduce.This enables system administrators to consolidate resources onto centralized servers on the network. Bluesky file system [4], it is a network file system,BlueSky gives standard POSIX document framework semantics, including nuclear renames and hard connections. Additionally utilizes 32 KB hinders rather than normal circle record framework size like 4 KB to lessen overhead. One goal of this work is to add to a multiplatform application that serves as capacity and reinforcement environment in the cloud. Cloud File System Oracle [5] is a document framework for private cloud situations, intended to oversee universally useful record store outside of a Prophet database over various agent framework stages with one administration

interface. Excessively it is firmly coordinated with the programmed stockpiling administration highlights of the prophet database.

Panzura CloudFS document System [6] is a document framework created from no place to give combination with cloud and NAS situations. It offers usefulness straightforward to clients, as everybody can see the same document from any area. It additionally permits information offering, without needing to erase the first document. Shared cloud-backed file system [7] gives solid consistency to the client. The vast majority of the document frameworks don't bolster the controlled record imparting among users. But this document framework does. We chose as a record framework PVFS2 for our cloud environment; to be free and open source, we have the flexibility to utilize and change as indicated by our needs. It likewise offers accessibility, adaptability and general awesome execution when composing to or perusing from the I/O servers. In the following section describes details of PVFS2.

3. PVFS2 OVERVIEW

The guardian catalog of the document can be on another metadata server. PVFS2 recognizes three unique segments as it is indicated in Fig.1.

1) Data servers or I/O servers constitute the PVFS store subsystem. Keeping in mind the end goal to ascertain the accessible store space the framework first decides the accessible space in every information server and after that it duplicates the littlest got esteem by the quantity of information servers.

2) Metadata servers store all the metadata data of the PVFS namespace records (metafiles).

3) Clients or transforming hubs that get to the record framework through one of three application developer interfaces (APIs): a part Linux module that permits access through the Linux virtual document framework (VFS) with the POSIX I/O Programming interface; a client space library that gives access through what it is known as the PVFS framework interface (pvfs2lib); and an abnormal state library, MPI-IO. A hub can be arranged as a metadata server, an information server, or both without a moment's delay.

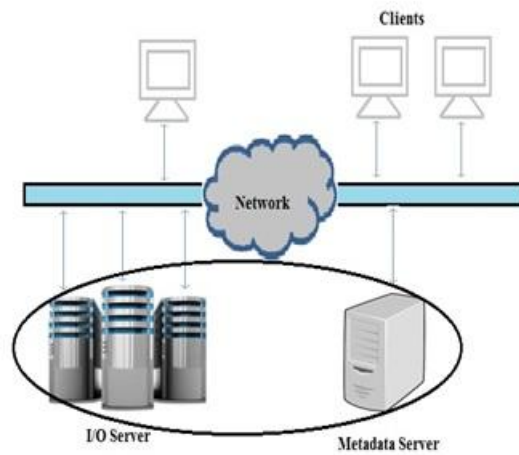


Figure 1:PVFS2 Structure

4.EXPIREMENTAL RESULT

The expiremental results shows storage an backup in cloud environment



Figure 2:User login

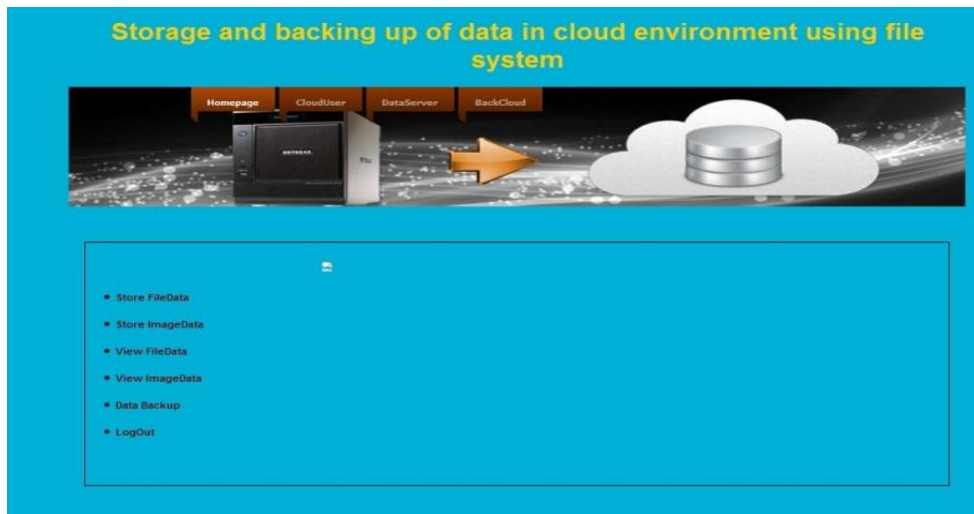


Figure 2:Colud User



Figure 4:User Operations



Figure 5:File backup id sent to user email id

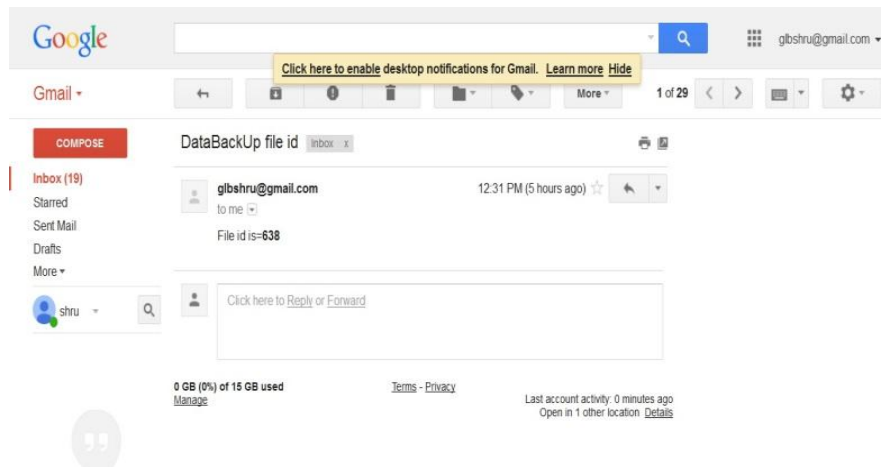


Figure 6: user email id



Figure 7:Send File back to user account

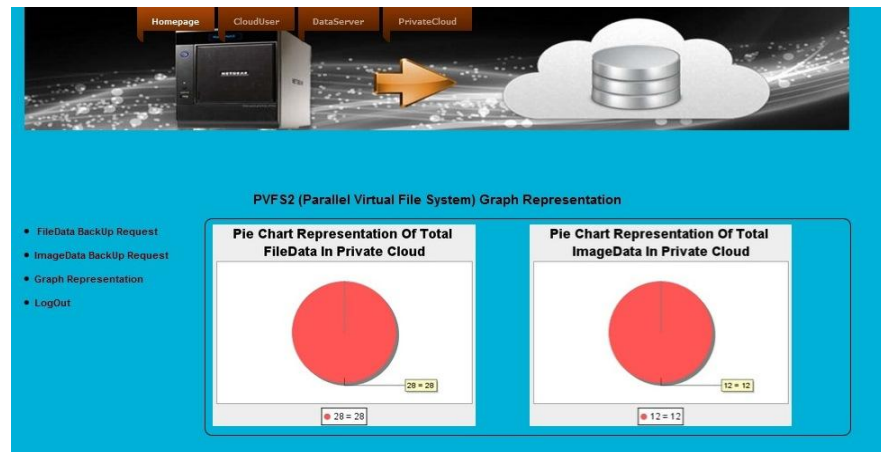


Figure 8:Graph Representation

5.CONCLUSION

The file system PVFS2 for a cloud environment for backup and data storage. The aim is to obtain better performance with the inclusion of PVFS2, because it decreases the data access latency, reducing network traffic and the data is distributed across different I/O servers. This allow data be distributed rather than be centralized, preventing complete loss of data. Also in the future we may use implementations made by Camacho and Nieto [11, 12] in order to further improve its performance and fault tolerance. Finally this implementation is being developed in order to give users the experience of implementing and managing a private cloud environment that facilitates the backup and data storage, using infrastructure already available or low cost. Avoiding payment of fees or memberships required to contract this service. This allows full control of those who access the information, so maintaining the confidentiality of the ddata

ACKNOWLEDGMENT

I thank to my guide Mrs. Nirupama Sheelawant Assistant.Prof, Department of computer science and Engineering, LAEC, Bidar for providing me with enough technical guidance and all the support needed throughout my work.

REFERENCES

- [1] Mohammad Hamdaqa, Ladan Tahvildari, Cloud Computing Uncovered: A Research Landscape, In: Ali Hurson and Atif Memon, Editor(s), Advances in Computers, Elsevier, 2012, Volume 86, Pages 41-85, ISSN0065-2458,ISBN9780123965356, <http://dx.doi.org/10.1016/B978-0-12-396535-6.00002-8>.
- [2] Dropbox: <https://www.dropbox.com>
- [3] Lei Chai Xiangyong Ouyang Ranjit Noronha Dhabaleswar K. Panda, "pNFS/PVFS2 over In_niBand: Early Experiences [4] Michael Vrable, Stefan Savage, and Geoffrey M. Voelker, "BlueSky: A Cloud-Backed File System for the Enterprise" Proceeding of th 10th USENIX Conference on File and Storage Technologies, February 14-17, 2012. ISBN 978-1-931971-91-1
- [5] Oracle Cloud File System (White paper): <http://www.oracle.com/us/products/database/cloud-file-system/Overview/index.html>
- [6] Panzura CloudFS file system (White paper) <http://panzura.com/products/global-file-system/>
- [7] Alysson Bessani, 1Ricardo Mendes,1Tiago Oliveira,1 Nuno Neves,1 Miguel Correia, 2Marcelo Pasin,3Paulo Verissimo1," SCFS: A Shared Cloud-backed File System"

- [8] Philip Carns, Sam Lang, Robert Ross, Murali Vilayannur, Julian Kunkel and Thomas Ludwig. (2009, 04-2009). Small-file access in parallel file systems.
- [9] R. Latham, N. Miller, R. Ross and P. Carns, "A Next-Generation Parallel File System for Linux Clusters," *LinuxWorld*, vol. 2, January, 2004.
- [10] J. M. Kunkel and T. Ludwig, "Performance evaluation of the PVFS2 architecture," in 2007, pp. 509-516.
- [11] Camacho, H.E.; Nieto, E.; Anguita, M.; Díaz, A.F.; Ortega, J., "Client cache for PVFS2," *Parallel Distributed and Grid Computing (PDGC), 2010 1st International Conference on*, vol., no., pp.38, 43, 28-30 Oct. 2010. Doi: 10.1109/PDGC.2010.5679607
- [12] Nieto, E.; Camacho, H.E.; Anguita, M.; Díaz, A.F.; Ortega, J., "Fault tolerant PVFS2 based on data replication," *Parallel Distributed and Grid Computing (PDGC), 2010 1st International Conference on*, vol., no., pp.107,112, 28-30 Oct. 2010 doi: 10.1109/PDGC.2010.5679880