

# Server Virtualization using Cloud Environment for Data Storage & Backup

Mayur S. Agrawal<sup>1</sup>, Nilesh Vani<sup>2</sup>

<sup>1,2</sup>North Maharashtra University, GF's G.C.O. Engineering, NH-6, Jalgaon (MH), India

**Abstract:** *The use of the disks of the node of a cluster as a worldwide stockpiling framework is a reasonable answer for the cloud environment. The usage for the availability of the data or information from wherever is increasing; this became a problem for many clients or users who use application such as databases, media, private files, documents etc. The Input-Output data difficulties of this application get higher as they get larger. In order to avoid this problem and improve the performance of this application here we use parallel file systems 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, Storage, backup, File System, Cluster, Parallel Virtual File System

## 1. Introduction

Cloud computing can be defined as that service (software, platform or infrastructure) located on the Internet and is accessed from a mobile device or desktop computer, giving users a wide variety of applications (databases, middle office software, storage, etc.). The different cloud deployment strategies public cloud services are characterized as being available to clients from a third party service provider via the Internet. The term "public" does not always mean free, even though it can be free or fairly inexpensive to use. A public cloud does not mean that a user's data is publicly visible; public cloud vendors typically provide an access control mechanism for their users. The difference between a private cloud and a public cloud is that in a private cloud-based service, data and processes are managed within the organization without the restrictions of network bandwidth, security exposures and legal requirements that using public cloud services might entail. In addition, private cloud services offer the provider and the user greater control of the cloud infrastructure improving security and resiliency because user access and the networks used are restricted and designated. A hybrid cloud is a combination of a public and private cloud that interoperates. A Cloud is a type of parallel and distributed system consisting of a collection of interconnected and virtualized computers that are dynamically provisioned and presented as one or more unified computing resources based on Service Level Agreements (SLAs) established through negotiation between the service provider and consumers" [1] Major goal of this work is to develop such application that serves as storage and backup through remote servers via cloud computing, we selected as a file system PVFS2 for our work; to be free and open source, we have the freedom to use. It also offers availability, flexibility and overall great performance when writing to or reading from the I/O servers. Cloud storage is the data is backed up and stored by using an internet connection on remote servers.

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. Related Work

Cloud Computing and cloud storage have become the preferred method for delivering information and online functionality. While some cloud services focus on providing consumers a wide range of services and functionalities. Others provide cloud storage to consumers for free or charge some type of subscription-based fee.

- A. Dropbox
- B. Google Drive
- C. Windows Azure
- D. Cloud File System Oracle
- E. Panzura CloudFS file System

Dropbox [2] gives users the capability of sharing entire folders with other Dropbox account users, which allows updates to be viewable by all collaborators. Users can download shared documents directly from Dropbox's web interface without having to install the Dropbox desktop client. Storing files in the Dropbox "Public" folder allows links to files to be sent to Dropbox and non-Dropbox users; however non-Dropbox link recipients must download the file to access/edit it, and any changes or revisions made to the file by the link-recipients will not be reflected in the Dropbox version of the file. Dropbox allows users to create a special folder on each of their computers, which Dropbox then synchronizes so that it appears to be the same folder (with the same contents) regardless of which computer is used to view

Volume 5 Issue 6, June 2016

[www.ijsr.net](http://www.ijsr.net)

Licensed Under Creative Commons Attribution CC BY

it. In Another approach of Google drive [3] Users of Google Drive documents must have a Google Drive account. All updates and editing by collaborators will be synced to Google Drive. For documents that you have permission to access, you can receive notifications when changes are made. You can share files with people by sending them a link to your file [3]. Windows Azure [4], is an open cloud platform in a global network of data centers run by Microsoft. Let's compile applications in any language, tool or framework for integrate your public cloud applications. It is important to know how to actually manage the backup and storage of files within the input / output. Currently there are file systems for cloud environments [5,6]. Panzura CloudFS file System [5] is a file system developed to provide integration with cloud and NAS environments. It offers functionality transparent to users, as everyone can see the same file from any location. It also allows data sharing, without having to delete the original file. Cloud File System Oracle [6] is a file system for private cloud environments, designed to manage general purpose file store outside of an oracle database across multiple operative system platforms with one management interface. Too it's tightly integrated with the automatic storage management features of the oracle database. The Parallel Virtual File System project is a multi-institution collaborative effort to design and implement a production parallel file system for HPC applications [7]

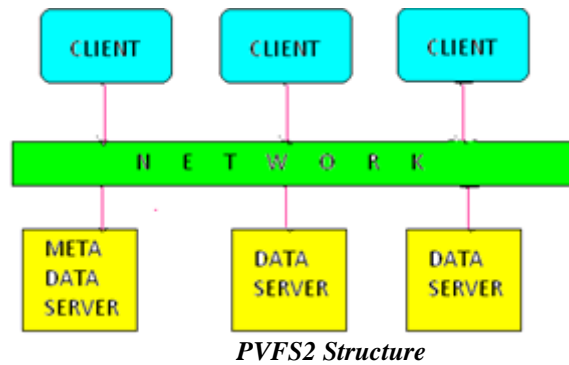
### 3. Proposed System

Virtualization of server using cloud environment is archived for data storage and backup by developing a client side application. That application is multiplatform simple application allows data transfer fast and simple way, to store data on cloud using file system like PVFS comparing with local file system in order to increases the performance of applications that requires high I/O need.

#### 3.1 Server for storage and Backup File System

Cloud is a network model in which the data is backed up and stored using an internet connection on remote servers. The service providers (hosting companies) provide large data centers and users who want to access their data or to store it buy or lease storage space from the servers. Virtualization is used by the data center managers often to reserve resources based on the requirements of the customer and storage spaces are allotted to back up their files or data objects. Physically, these resources can protract along pooled servers and locations. The confidentiality provided to the user data purely depends on the service provider. PVFS2 file system had been selected for our data store on server; as it is open source, we have the flexibility to use and modify according to our compulsion. Let us know how these file systems actually manage the backup and storage of files within the input / output.

#### 3.2 PVFS2 Overview



PVFS2 distinguishes three different components as

**3.2.1 Data servers or I/O servers** aggregates the PVFS store subsystem. The storage space available in individual data servers is obtained and the smallest obtained value is multiplied by the number of data servers to know the available data storage.

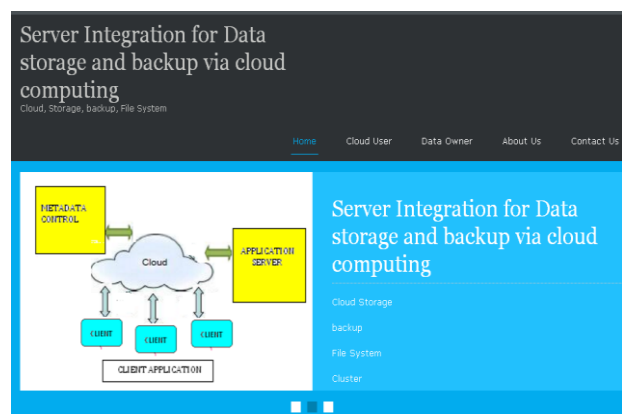
**3.2.2 Metadata servers** store all the metadata information of the PVFS namespace files (metabytes).

**3.2.3 Clients or processing nodes** that approach the file system through one of three application programmer interfaces (APIs): a kernel Linux module that allows approach through the Linux virtual file system (VFS) with the POSIX I/O API; a user-space library that provides access through what it is called the PVFS system interface (pvfs2lib); and a high level library, MPI-IO.

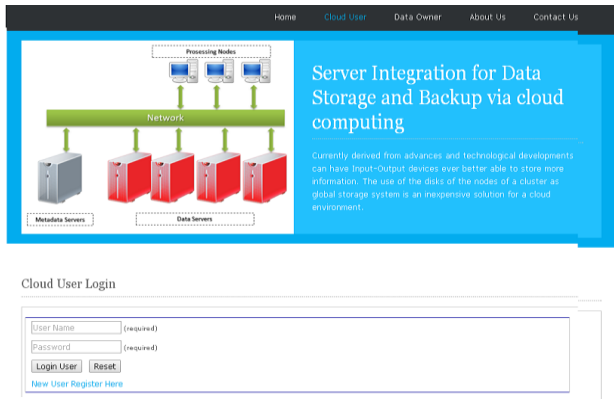
#### 3.3 Operation of Server Application

Each user is provided with a thread through which Sockets application get in contact with the client for sharing the data, based on the described rules in the precedent section. Data and metadata are received by application server as string of bytes encrypted. A web application had been designed for users to view and download their data by creating the user account and thus maintaining personal information. File name, size, path, last date modified, deleted mark, this information is stored in a MySQL table.

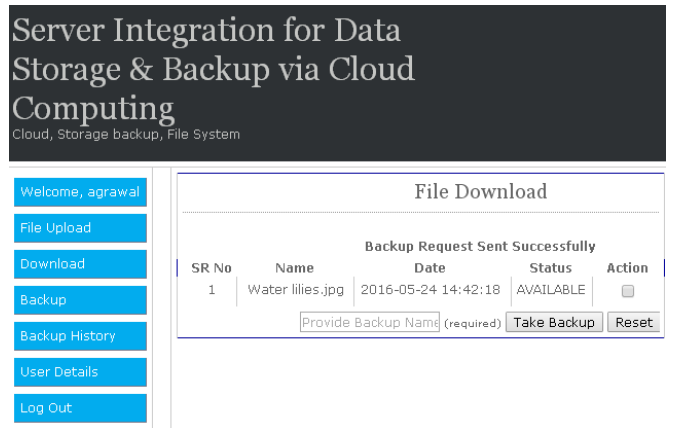
### 4. Experimental Results



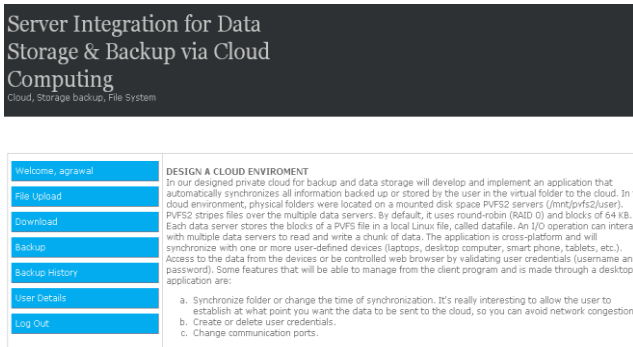
**Figure 1:** Home Page for Server Integration for Data Storage & Backup via Cloud Computing



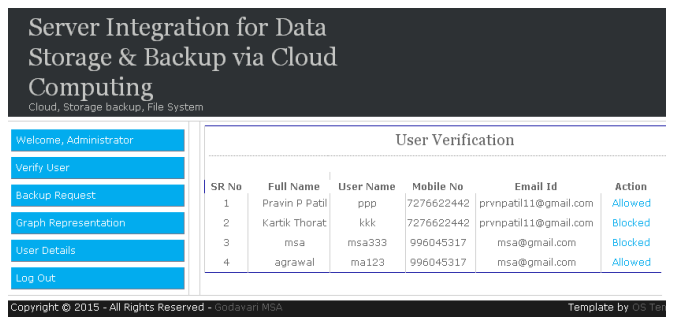
**Figure 2: User Login**



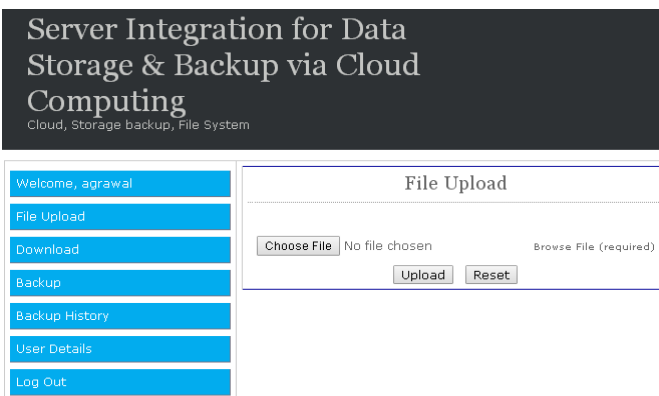
**Figure 6: Backup request sent to data owner**



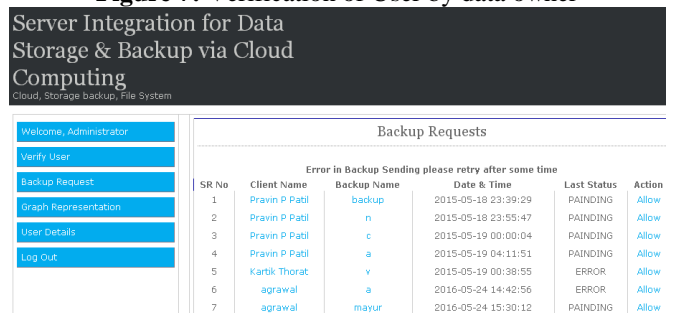
**Figure 3: User Action**



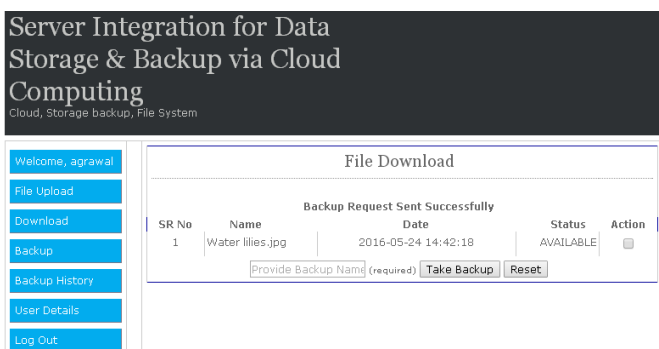
**Figure 7: Verification of User by data owner**



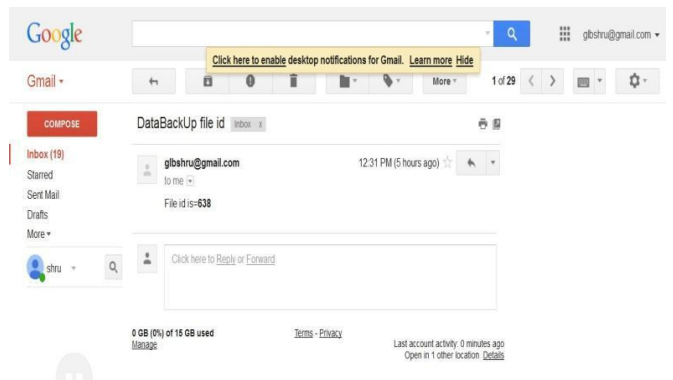
**Figure 4: File Upload**



**Figure 8: Verification & provision of backup by data owner**



**Figure 5: File Download**



**Figure 9: User getting back up on Email Id**

## 5. Conclusion

We integrate server using private cloud model for backup and data storage will develop and application is implemented in such way that automatically synchronizes all information backed up or stored by the user in the virtual folder to the cloud We use PVFS2 file system for storage data in order to increase the performance of applications that requires high

I/O data demands. PVFS2 is file system which is open source, will reduce access time to data as it allows input / output parallel. On the client-side, an application is developed that allows data transfer fast and simple way. The advantages of this implementation are that it can reuse existing infrastructure (servers, cluster, and other devices) that reduces the cost and increases the throughput.

## References

- [1] Rajkumar Buyya, Chee Shin Yeo, and Srikumar, Venugopal. "Market-oriented cloud computing: Vision, hype, and reality for delivering it services as computing utilities". CoRR, (abs/0808.3558), 2008.
- [2] Dropbox: <https://www.dropbox.com>
- [3] [https://www.google.com/intl/en\\_US/drive/start/index.html](https://www.google.com/intl/en_US/drive/start/index.html)
- [4] Windows Azure: <http://www.windowsazure.com/es-es/>
- [5] Oracle Cloud File System (White paper): <http://www.oracle.com/us/products/database/cloudfilesystem/overview/index.html>
- [6] Panzura CloudFS file system (White paper) <http://panzura.com/products/global-file-system/>
- [7] Samuel Lang, Philip Carns, Robert Latham, Robert Ross, Kevin Harms, William Allcock, "I/O Performance Challenges at Leadership Scale," Proceedings of Supercomputing, 2009
- [8] Philip H. Carns, Walter B. III, Robert Ross, Pete Wyckoff, "BMI: a network abstraction layer for parallel I/O," Proceedings of IPDPS '05, 2005

## Author Profile

**Mayur S Agrawal** received the B.E. degree in Computer Engineering from S S G B College of Engineering in 2008 He was worked with S N P P W Polytechnic affiliated to Maharashtra State Board of Technical Education, Mumbai.

**Prof. Nilesh Vani** received the B.E. & M.Tech Degrees in Computer Engineering He is Asst professor in GF'S G.C.O.Engineering affiliated to North Maharashtra University, Jalgaon lephone, Inc.