An architecture of a Cluster based Storage Server for Cloud Storage

Tin Tin Yee, Thinn Thu Naing
University of Computer Studies, Yangon
tintinyee.tty@gmail.com, thinnthu@gmail.com

Abstract

Cloud storage server design and services play a vital role in the cloud computing infrastructure. Cloud storage focused at the enterprise to that focused on end users, as well as it providers offer huge capacity cost reductions, the elimination of labor required for storage management and maintenance, and immediate provisioning of capacity at a very low cost per terabyte. This paper presents the architecture of low cost PC cluster based storage server which can be activated to store large amount of data and provide cost-effective, high-performance machines.

Keywords: *Cloud Computing, Cloud storage, PC Cluster*

1. Introduction

Cloud computing is a new concept of computing. Cloud computing is Internet (cloud) based development and use of computer technology (computing), whereby dynamically scalable and often virtualized resources are provided as a service over the Internet. In generally, the concept of cloud computing incorporate web infrastructure, Web 2.0, virtualization technologies and other emerging technologies. Cloud computing has become a significant technology trend, and many experts expect that cloud computing will reshape information technology (IT) processes and the IT marketplace. With the cloud computing technology, users use a variety of devices, including PCs, laptops, smartphones, and PDAs to access programs, storage, and application-services offered by cloud computing providers. Advantages of the cloud computing technology include cost savings, high availability, and easy scalability.

The main objective of cloud computing is to provide ICT services over the cloud. The service models are divided in cloud as Software-as-a-Service (SaaS), which allows users to run applications remotely from the cloud. Infrastructure-as-a-service (IaaS) refers to computing resources as a service. This includes virtualized computers with guaranteed processing power and reserved bandwidth for storage and Internet access. Platform-as-a-Service (PaaS) is similar to IaaS, but also includes operating systems and required services for a particular application. In other words, PaaS is IaaS with a custom software stack for the given application. The Data storage-as-a-Service (DaaS) provides storage that the consumer is used including bandwidth requirements for the storage.

From the end user point of view, cloud computing services provide the application software and operating system from the desktops to the cloud side, which makes users will be able to use anytime from anywhere and utilize large scale storage and computing resources. In cloud computing, disk storage is the one of the biggest expenses. There is strong concern that cloud service providers will drown in the expense of storing data, especially unstructured data such as documents, presentations, PDFs, VM Images, Multimedia data, etc. Cloud service providers offer huge capacity cost reductions, the elimination of labor required for storage management and maintenance and immediate provisioning of capacity at a very low cost per terabyte. Therefore, the storage and computing on massive data are major key challenge for a cloud computing infrastructure. In this paper, we focus on PC cluster based storage server which provides the hardware and software devices for large amount of data storage.

The rest of this paper is organized as follows. Section 2 describes cloud storage and related works. In section 3 presents academic based private cloud architecture. In section 4 discuss how to build physical topology of CCPS (Cloud based Cluster PC System). Finally section 5 future works and concludes the paper.

2. Cloud Storage and Related Works

The cloud has become a new vehicle for delivering resources such as computing and storage to customers on demand. Cloud storage is not a new concept. Cloud Storage delivers virtualized storage on demand, over a network based on a request for a given quality of service (QoS). There is no need to purchase storage or in some cases even provision it before storing data. We only pay for the amount of storage we data are actually consuming. There are so many ongoing researches in cloud storage.

Nowadays, a huge variety of cloud storage systems is available, all with different functionality, optimizations and guarantees. That is, solutions focus on a particular scenario and the provided services are tailored accordingly. As a result, cloud storage systems vary in the data format (e.g., Key/Value vs. row store), access-path optimization (e.g., read vs. write, one-dimensional vs. multi-dimensional access), distribution (e.g., single vs. multi-data center distribution), query language, transaction support, availability. Typical cloud storage system architecture includes a master control server and several storage servers. Many cloud storage architecture focus on multiple tenants.

Jiyi [6] introduced cloud storage reference model. This model designed scalable and easy to manage storage system but aren't designed to be high performance. This paper presented the key technologies, several different types of clouds services, the advantages and challenges of cloud storage. Hussam [1] described RAID (Redundant Arrays of Inexpensive Disks)-like techniques at the cloud storage and showed that reduce the cost and better fault tolerant. Amazon S3 provides a simple web services interface that can be used to store and retrieve any amount of data, at any time, from anywhere on the cloud. It gives any developer access to the same highly scalable, reliable, secure, fast, inexpensive infrastructure that Amazon uses to run its own global network of web sites. The service aims to maximize benefits of scale and to pass those benefits on to developers. Donald [4] presented Cloudy, a modular cloud storage system and provided a highly flexible architecture for distributed data storage and is designed to operate with multiple workloads. André [2] presents analytical considerations on the scalability of storage clusters and presents a storage cluster architecture based on peer-to-peer computing that is able to scale up to hundreds of servers and clients.

3. Academic based Private Cloud Architecture

The main focus of our work is to build storage server using pc cluster for academic based private cloud. Our Cloud implementation is implements by Ubuntu Enterprise Cloud (UEC) architecture. The UED is powered by Eucalyptus, an open source implementation for the emerging standard of the Amazon EC2 API. Our physical architecture is shown in figure 1.

Physical network architecture is based on the backbone network. This architecture consists of front-end infrastructure, back-end infrastructure and storage server. The front-end infrastructure contains Cluster Controllers (CC), Cloud Controller (CLC). The back-end infrastructure consists of Node Controllers (NC). The storage server is makes up of many of PC cluster node and which is the main focus of our work. Cloud testbed systems configuration is shown in table 1 and table 2 present allocated IP address.

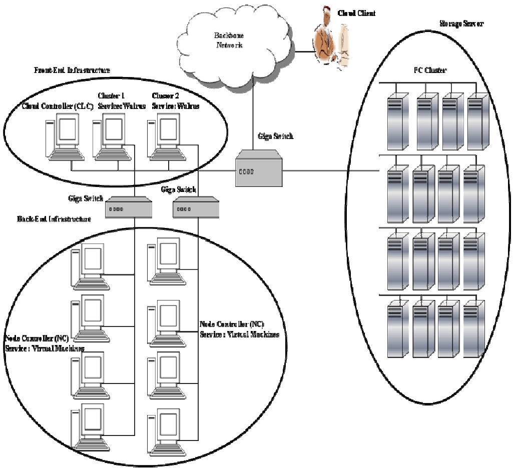

Figure 1. Physical architecture for academic based private cloud

3.1 Front-end Infrastructure

The front-end infrastructure contains cluster Controllers (CC), Cloud Controller (CLC). There are two Cluster Controllers: cluster 1 and cluster 2 and each cluster configure Walrus Storage Controller (WSC) services. And then, Elastic Block Storage Controller (EBS) runs on the same machines as the Cluster Controller and is configure automatically when the Cluster Controller. It allows to create persistent block devices that can be mounted on running machines in order to gain access to virtual hard drive.

The Cluster Controller (CC) operates as the go between the Node Controller and the Cloud Controller. It will receive requests to allocate virtual machine images from the Cloud Controller and in turn decides which Node Controller will run the virtual machine instance. All the virtual machine images were stored in the Walrus Storage service. The actual storage behind the Walrus was implemented as iSCSI volumes hosted on Storage Server. The Cloud Controller (CLC) is providing the interface with which users of the cloud interact. The CLC contacts with the Cluster Controller (CC) and makes the top level choices for allocating new instance.

Table 1. Cloud testbed system configuration

System	Process Configuration	Other Information
Cloud Controller (CLC)	Intel ® Core™ 2 Duo	Processor: Intel® Core ™ 2 Duo CPU E7500 @ 2.93GHz Memory:2GB RAM Storage: 250GB HDD
Cluster Controller (CC)	Intel ® Core™ 2 Duo	Processor: Intel® Core ™ 2 Duo CPU E7500 @ 2.93GHz Memory:2GB RAM Storage: 250GB HDD
Node Controller (NC)	Intel ® Core™ 2 Duo	Processor: Intel® Core ™ 2 Duo CPU E7500 @ 2.93GHz Memory:4GB RAM Storage: 250GB HDD
Storage Server Pools	Pentium® Dual-Core	Processor: Pentium® Dual-Core CPU T4200 @2.00GHz Memory:1GB RAM Storage: 250GB HDD x 20

Table 2. Allocated IP address

	IP range	Subnet mask
Cloud Controller (CLC)	192.168.10.4	255.255.255.0
Cluster Controller (cluster1)	192.168.10.5	255.255.255.0
Cluster Controller (cluster2)	192.168.10.6	255.255.255.0
Node Controller (NC)	192.168.20.0/29	255.255.255.0
Node Controller (NC)	192.168.30.0/29	255.255.255.0
Virtual Machine Network on Node Controller (NC)	192.168.40.0/16	255.255.255.0
Storage Server Pools	192.168.50.0/16	255.255.255.0

3.2 Back-end Infrastructure

Back-end infrastructure consists of Node Controllers (NC). The Node Controllers (NC) software runs on the physical machines on which virtual machine instances will be instantiated. These Virtual machine instances has own IP address on Node Controller. Each of Node Controller has a local storage device. The local storage was only used to hold virtual machine image at run time and for caching virtual machine instances. When a virtual machine is terminated, the storage on the virtual machine instance is released and storage in the storage server pool as virtual machine image.

4. Storage server

Storage server is main focus of our work. We build terabyte storage server as iSCSI storage using many of PC cluster node.

4.1 Physical Topology of CCPS

In this section describes how to build physical topology of CCPS. The system framework of CCPS is shown in figure 2.

The overall framework of CCPS consists of three layers which are web based application services layer, Cloud base Cluster File System (CCFS) layer and PC Cluster layer. The web based application layer provides the services to the users which are Virtual Machine (VM) images, dataset and multimedia data, etc. The CCFS layer supports the cluster file system for PC cluster layer. The PC cluster layer provides to store large amount of data as the iSCSI storage server.

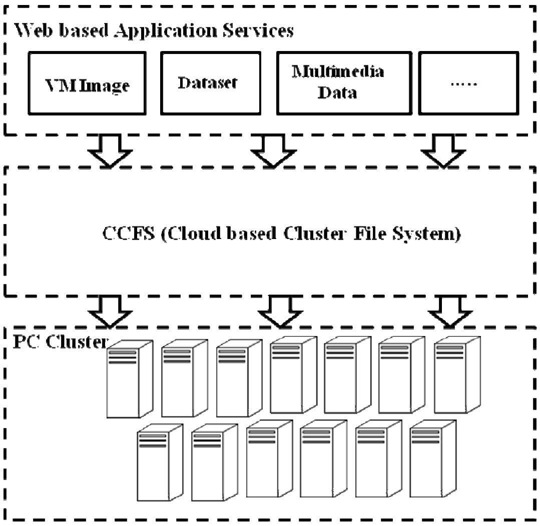

Figure 2. System framework of CCPS

4.2 PC Cluster

Nowadays, many organizations need terabytes storage systems but are very expensive and require higher degree of skills for their operations and maintenance. Usually a PC contain more than 100 GB Hard Disk Drive (HDD), at least 256 MB or greater RAM and 2GHz or higher processor. A typically installation of an operating system and other software installation do not use more than 30 to 40 GB of HDD storage. This leaves on the average about 70% of the storage space to be unused. Therefore, I propose PC cluster based storage server that is inexpensive and easy to maintain.

PC cluster become popular in the 1990 for batching processing of high performance application. A PC cluster is a collection of computer nodes, which is interconnected by a high-speed switching network, all nodes can be used individually or collectively as a cluster. PC cluster system which achieve high performance using many PC.

4.2.1 System overview of PC Cluster based storage server

PC cluster based storage server tries to transfer to the cluster computing to storage server. The storage server uses inexpensive PC components, offers an inherently scalable aggregate I/O bandwidth, and can take advantage of existing cluster installations through double-use of hardware. Moreover, large files can be stored in a scalable fashion by striping the data across multiple nodes. Further, the availability of CPU and memory on each node offers the flexibility of additional data manipulations such as caching. The system configuration of PC cluster is shown in figure 3.

In PC cluster, each individual machine of a cluster is referred to as a node. Our system is based on client-server architecture and consists of one server node and many clients nodes in order to make them work as a single machine. Server node has two networks cards and one is connect to the front-end infrastructure using giga switch and other is connect to the client nodes using 100 Mbps Fast Ethernet switch. Each of the nodes has 2.0 GHz Pentium P4 processors, 250 GB Hard disks and 1 GB RAM. All of the nodes are installed Ubuntu Desktop 10.10 (32 bits) Linux distribution operating system which is an open source operating system built around the Linux kernel. In our system configuration used the number of 20 PCs and consumed about 30 GB Hard disks storage of each PC for installation of an operating system and other software installation. The available storage capacity of these PCs is combined together, and then can provide 20 x 220 = 4400 GB of storage capacity. This surplus multi-Terabytes storage capacity remains unused and can be utilized if combined to store huge amount of data.

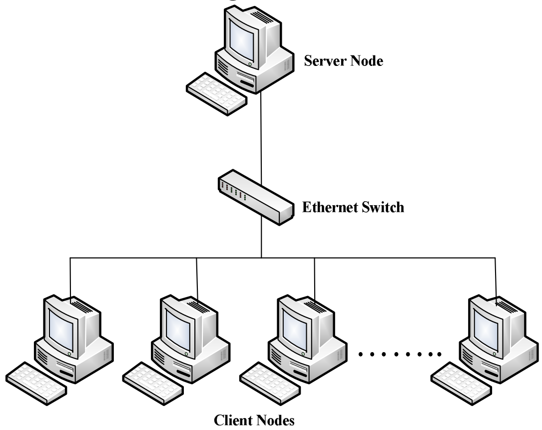

Figure 3. System configuration of PC cluster

CCPS (Cloud based Cluster PC System) uses CCFS (Cloud based Cluster File System) to store data files in the collection of the computer nodes. In CCFS has name node and data node. Name node works as the global controller and maintains the information of whole CCFS system. Data node stores the physical storage of the file. In pc cluster, name node runs on the server node that maintains the metadata of CCFS which contains the information of blocks, the current locations of blocks, etc. Moreover, master node provides the monitoring of all client nodes states in system. Files are divided into multiple chunks and distributed cross client nodes. The client nodes are responsible to store data. As large numbers of nodes are used in this system, the probability of some nodes will be failure. Therefore, improving fault tolerant is a major challenge.

PC cluster has lower implementation cost and store large amount of data but major bottleneck is I/O operation for data processing. Due to I/O bottleneck problem, our design achieves a single I/O space for all of the blocks of data in the cluster. This allows any node to remotely access any I/O peripheral or disk devices without the knowledge of their physical location.

4.2.2 Single I/O space for PC cluster

Single I/O space could be implemented in the user level, file system level and device driver level. User level designs are believed to have higher portability and lower implementation cost. However, they usually cannot provide a complete Single I/O space while users still have to use specific application programming interfaces (APIs) and identifiers in order to exploit full functionality of the packages. Furthermore, using system calls and kernel facilities to perform file and network I/O may decrease the performance. File system level designs can have full control in data distribution while providing a complete Single I/O space to the user. However, changing the file system does not guarantee strict compatibility with current applications and the required development cost is high. Our design is based on device driver level. The main objective is a single I/O space for all data blocks in the cluster by constructing a single device driver (SDD) for each node in the cluster. The CCFS in each node will perceive an illusion that it is using a large disk, and this illusion is provided by SDD. Figure 4 illustrates the Single I/O space in pc cluster. Therefore, Single I/O space improves I/O performance processing capacity.

High availability features can be supported by implementing some established technologies directly in the SDD. We have implemented RAID-6 architectures in our prototype. In RAID 6, two different parity calculations are carried out and stored in separate blocks on different disks. Thus, a RAID 6 whose user data require N disks consists of N+2 disks. This makes it possible to regenerate data even if two disks containing user data fail. The advantage of RAID 6 is that it provides extremely high data availability.

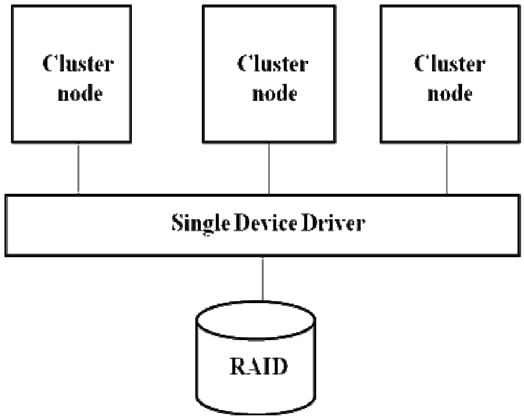

Figure 4. Single I/O space

5. Conclusion and Future Works

In this paper, we design PC cluster based storage server which can be used to store large amount of data and provide high performance processing capacity. As our future works, the greatest challenges of our system are high availability and fault tolerance.

References

[1] A. Hussam, P. Lonnie, W. Haki , "RACS: A Case for Cloud Storage Diversity", In Proceedings of the SoCC'10, 2010.

[2] B. André and E. Sascha, "Inter-node Communication in Peer-to-Peer Storage Clusters", In Proceedings of 24th IEEE Conference on Mass Storage Systems and Technologies, 2007.

[3] G. Garth, "Cloud Storage and Parallel File Systems", In Proceedings of SNIA Storage Developer Conference, 2009.

[4] K. Donald , M.Stephan, K. Tim, M.Raman L.Simon , P. Flavio, "Cloudy: A Modular Cloud Storage System", In Proceedings of the 36th International Confernce on Very Large Database, 2010, pages 1533-1536.
[5] P. Bo, C. Bin and L. Xiaoming, "Implementation Issues of A Cloud Computing Platform", Bulletin of the IEEE Computer Society Technical Committee on Data Engineering, 2009.
[6] W. Jiyu, Z. Jianlin, L. Zhije, J. Jiehui, "Recent Advances in Cloud Storage". In Proceedings of the Third International Symposium on Computer Science and Computational Technology, 2010, pages 151-154.