KISM Smart Media Journal
Vol. 4, No. 3 pp. 62~67 ISSN:2287-1322

Prototype Design of Mass Distributed Storage System based on PC using Ceph for SMB

ByungRae Cha*,YongilKim**

Abstract

The trend keywords in ICT sector will be Big Data, Internet of Things, and Cloud Computing. The rear end to support those techniques requires a large-capacity storage technology of low-cost. Therefore, we proposed the prototype of low-cost and mass distributed storage system based on PC using open-source Ceph FS for SMB.

Keywords : Ceph | Distributed File System (DFS)|Mass Distributed Storage | Internet of Things (IoT)

I. INTRODUCTION

The trend keywords in ICT sector will be Big Data, Internet of Things, and Cloud Computing in Gartner's top 10 strategic IT trends of Fig. 1. [1] The inexorable march of technology has the computing and software community focusing on a large-scale integration challenge by exploiting cloud computing, machine-to-machine (M2M) applications, and Big Data solutions. The convergence of M2M, Big Data, and the cloud will provide key capabilities for building next-generation systems and expanding the Internet of Things. This Internet of Things (IoT) will sustain communication and data sharing for billions of connected devices and systems. The number of connections and endpoints is so large it requires IPv6 addresses because the IPv4 addressing scheme cannot accommodate the volume of sensors, smart phones, smart factories, smart grids, smart vehicles, controllers, meters, and other devices that will be transmitting data over the Internet.

Building the IoT will be an exercise in integrating disparate devices and carrier networks, multiple communication protocols, and a wide variety of applications. It will often require integrating legacy networks and applications. The new M2M applications are and will continue to be complex, using geographically dispersed devices and services, a mix of connectivity, and logic in the data center and edge devices.

	2010	2011	2012	2013	2014	2015
1	Cloud Computing	Cloud Computing	Media Tablets & Beyond	Mobile Device Battles	Mobile Device Diversity and Management	Computing Everywhere
2	Advanced Analytics	Mobile Applications and Media Tablets	Mobile Centric Applications & Interfaces	Mobile Applications & HTML 5	Mobile Apps and Applications	The Internet of Things(IoT)
3	Client Computing	Social Communications & Collaboration	Contextual and Social User Experience	Personal Cloud	Internet of Everything	3D Printing
4	IT for Green	Video	Internet of Things	Enterprise App Stores	Hybrid Cloud and IT as Service Broker	Advanced, Pervasive, Invisible Analytics
5	Reshaping the Data Center	Next Generation Analytics	App Stores and Marketplaces	Internet of Things	Cloud/Client Architecture	Context-Rich Systems
6	Social Computing	Social Analytics	Next-Generation Analytics	Hybrid IT & Cloud Computing	Era of Personal Cloud	Smart Machines
7	Security-Activity Monitoring	Context-Aware Computing	Big Data	Strategic Big Data	Software Defined Anything	Cloud/Client Architecture
8	Flash Memory	Storage Class Memory	In-Memory Computing	Actionable Analytics	Web-Scale IT	Software Defined Infra. & Applications
9	Virtualization for Availability	Ubiquitous Computing	Extreme Low-Energy Servers	In-Memory Computing	Smart Machines	Web-Scale IT
10	Mobile Applications	Fabric-Based Infra. & Computers	Cloud Computing	Integrated Ecosystems	3-D Printing	Risk-based Security & Self-Protection

Fig. 1. Gartner's top 10 Strategic IT trends between 2010 and 2015

The rear end to support those techniques requires a large-capacity storage technology of low-cost. Therefore, we proposed the draft of low-cost and mass distributed storage system using open-source Ceph for SMB.

* Member, Dept. of Telecommunication Engineering, GIST, Korea
** Member, Dept. of Internet Contents, Honam Univeristy, Korea

Manuscript : 2014.12.30
Revised : 2015.09.23
Confirmation of Publication : 2015.09.25
Corresponding Author: Yongil Kim,
e-mail: yikim@honam.ac.kr

Ⅱ. Storage Technology & Ceph

A. Storage Technology

The trend of storage technology is extensible through the network technology, continually being evolved into high performance / high capacity.

- DAS/SAN/NAS

DAS (Direct-attached storage) refers to a digital storage system directly attached to a server or workstation, without a storage network in between. It is a retronym, mainly used to differentiate non-networked storage from the concepts of SAN (storage area network) and NAS (network-attached storage). NAS is file-level computer data storage server connected to a computer network providing data access to a heterogeneous group of clients. NAS not only operates as a file server, but is specialized for this task either by its hardware, software, or configuration of those elements. NAS is often manufactured as a computer appliance – a specialized computer built from the ground up for storing and serving files – rather than simply a general purpose computer being used for the role.

- CFS (Cluster File System)

A CFS is a file system which is shared by being simultaneously mounted on multiple servers. There are several approaches to clustering, most of which do not employ a clustered file system (only direct attached storage for each node). CFSs can provide features like location-independent addressing and redundancy which improve reliability or reduce the complexity of the other parts of the cluster. Parallel file systems are a type of clustered file system that spread data across multiple storage nodes, usually for redundancy or performance.

- SDS (Software Defined Storage)

SDS is a term for computer data storage technologies which separate storage hardware from the software that manages the storage infrastructure.[3] The software enabling a software-defined storage environment provides policy management for feature options such as deduplication, replication, thin provisioning, snapshots and backup. By definition, SDS software is separate from hardware it is managing. That hardware may or may not have abstraction, pooling, or automation software embedded. This philosophical span has made software-defined storage difficult to categorize. If it can be used as software on commodity servers with disks, it suggests software such as a file system. If it is software layered over sophisticated large storage arrays, it suggests software such as storage virtualization or storage resource management, categories of products that address separate and different problems. In March 2014 SNIA began a Draft Technical Work available for public review on Software Defined Storage.[4]

B. Ceph

Ceph is a free software storage platform designed to present object, block, and file storage from a single distributed computer cluster.[5] Ceph's main goals are to be completely distributed without a single point of failure, scalable to the exabyte level, and freely-available. The data is replicated, making it fault tolerant. Ceph software runs on commodity hardware. The system is designed to be both self-healing and self-managing and strives to reduce both administrator and budget overhead.

Ceph uniquely delivers object, block, and file storage in one unified system. Ceph is highly reliable, easy to manage, and free. The power of Ceph can transform your company's IT infrastructure and your ability to manage vast amounts of data. Ceph delivers extraordinary scalability – thousands of clients accessing petabytes to exabytes of data. A Ceph Node leverages commodity hardware and intelligent daemons, and a Ceph Storage Cluster accommodates large numbers of nodes, which communicate with each other to replicate and redistribute data dynamically.

Fig. 2 and Fig. 3 show the architecture and storage typeof object, block, and file in one unified system of Ceph. Ceph provides an infinitely scalable Ceph Storage Cluster based upon RADOS (Reliable Autonomic Distributed Object Store).

Ceph implements distributed object storage. Ceph's software libraries provide client applications with direct access to the RADOS object-based storage system, and also provide a foundation for some of Ceph's features, including RBD (RADOS Block Device), RADOS Gateway, and the Ceph File

System. The RADOS Gateway also exposes the object store as a RESTful interface that can present as both OpenStack Swift APIs and native Amazon S3.

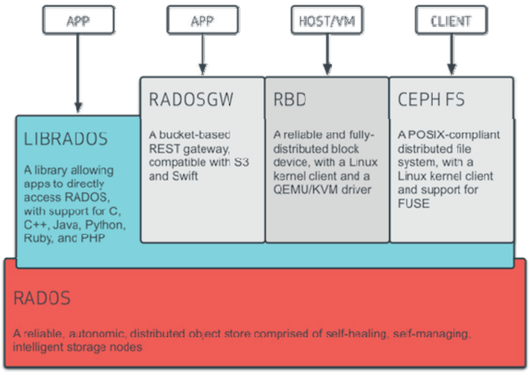

Fig. 2. Architecture of Ceph

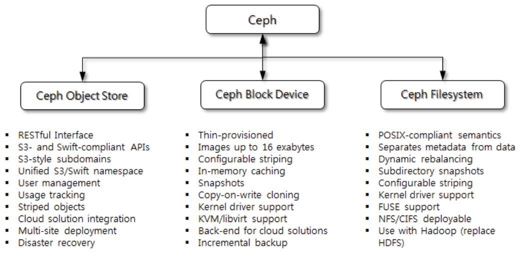

Fig. 3. Storage types of Ceph

Ceph's object storage system allows users to mount Ceph as a thinly provisioned block device. When an application writes data to Ceph using a block device, Ceph automatically stripes and replicates the data across the cluster. Ceph RBD also integrates with KVMs (kernel virtual machines). Ceph RBD interfaces with the same Ceph object storage system that provides the librados interface and the Ceph FS, and it stores block device images as objects. Since RBD is built on top of librados, RBD inherits librados's capabilities, including read-only snapshots and revert to snapshot. By striping images across the cluster, Ceph improves read access performance for large block device images.

Ceph FS runs on top of the same object storage system that provides object storage and block device interfaces. The Ceph metadata server cluster provides a service that maps the directories and file names of the file system to objects stored within RADOS clusters. The metadata server cluster can expand or contract and it can rebalance the file system dynamically to distribute data evenly among cluster hosts. This ensures high performance and prevents heavy loads on specific hosts within the cluster. Clients mount the POSIX-compatible file system using a Linux kernel client. And an older FUSE-based client is also available. The servers run as regular UNIX daemons.

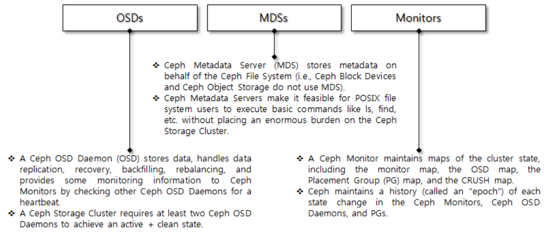

Fig. 4. Component of Ceph

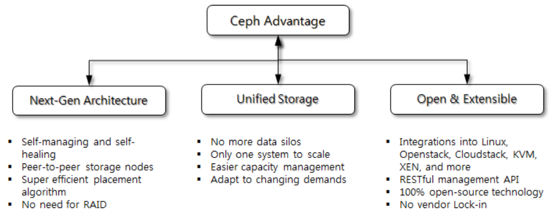

Fig. 5. Advantages of Ceph FS

A Ceph storage cluster requires at least one Ceph monitor and at least two Ceph OSD daemons. The Ceph metadata server is essential when running Ceph file system clients as shown in Fig. 4. Ceph stores a client's data as objects within storage pools. Using the CRUSH algorithm,[6] Ceph calculates which placement group should contain the object, and further calculates which Ceph OSD daemon should store the placement group. The CRUSH algorithm enables the Ceph storage cluster to scale, rebalance, and recover dynamically. Fig. 5 presents the advantages of Ceph: next-gen architecture, unified storage, open, and extensible. [7]

Ⅲ. Draft of Mass Distributed Storage System using Ceph FS

A. Proposed mass distributed storage

We proposed the draft of mass distributed storage system using open-source Ceph, concept diagram of Ceph infrastructure, and S/W & H/W Spec. based on PC in Fig. 6, Fig. 7, and Table 1. The proposed mass distributed storage consists of type A and type B, it is designed to enable the scale out by adding type B. Specially, type A is mounted to the security module. And Fig. 8 shows the dashboard Kraken to check the state of Ceph storages for storage manager.

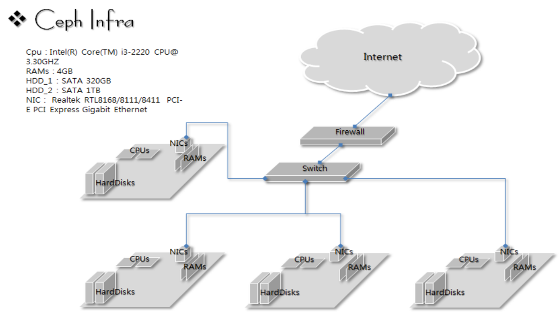

Fig. 6. Concept diagram of proposed prototype system

Table 1. H/W Spec. of proposed mass distributed storage.

Item	Spec.	Count
CPU	Intel i3, 3.3GHz, Dual Core	4
RAM	32GB	4
Storage	512GB, 1TB	4,4
NIC	RealtPCI-E Gigabit Ethernet	4
Network	NetGear 24 port Switching Hub	4

Fig. 7. . Prototype of mass distributed storage system based on PC using Ceph

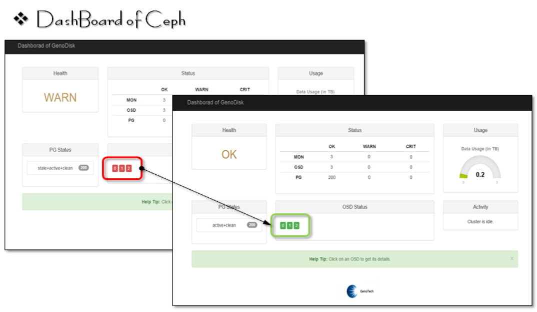

Fig. 8. Dashboard kraken to check the state of Ceph storages

B. Security and Monitoring Module

As shown in Fig. 9, the Security Onion is a Linux distro for IDS (Intrusion Detection System), NSM, and log management. It's based on Ubuntu and contains Snort. [8] Snort is a free and open source network intrusion prevention system (NIPS) and network intrusion detection system (NIDS) created by Martin Roesch in 1998. Snort's open source network-based intrusion detection system (NIDS) has the ability to perform real-time traffic analysis and packet logging on Internet Protocol (IP) networks. Snort performs protocol analysis, content searching, and content matching. These basic services have many purposes including application-aware triggered quality of service, to de-prioritize bulk traffic when latency-sensitive applications are in use. The program can also be used to detect probes or attacks, including, but not limited to, operating system fingerprinting attempts, common gateway interface, buffer overflows, server message block probes, and stealth port scans.

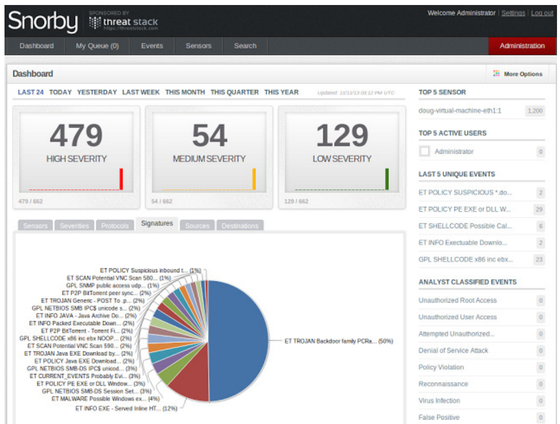

Fig. 9 Dashboard of Security Onion

Snort can be configured in three main modes: sniffer, packet logger, and network intrusion detection. In sniffer mode, the program will read network packets and display them on the console. In packet logger mode, the program will log packets to the disk. In intrusion detection mode, the program will monitor network traffic and analyze it against a rule set defined by the user. The program will then perform a specific action based on what has been identified.

Ganglia are a scalable distributed monitoring system for high-performance computing systems such as clusters and Grids. It is based on a hierarchical design targeted at federations of clusters. It leverages widely used technologies such as XML for data representation, XDR for compact, portable data transport, and RRDtool for data storage and visualization. [11]

Ⅳ. Performances of Prototype System

Snort can be configured in three main modes: sniffer, packet logger, and network intrusion detection. In sniffer mode, the program will read network packets and display them on the console. In packet logger mode, the program will log packets to the disk. In intrusion detection mode, the program will monitor network traffic and analyze it against a rule set defined by the user. The program will then perform a specific action based on what has been identified.

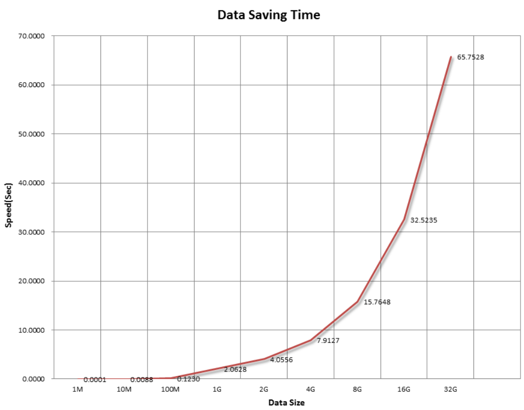

Fig. 10. Data saving time by data size

Ⅴ. Future Appliances

A. OpenStack with Ceph FS

OpenStack software controls large pools of compute, storage, and networking resources throughout a datacenter, managed through a dashboard or via the OpenStack API. [9] OpenStack works with popular enterprise and open source technologies making it ideal for heterogeneous infrastructure.

B. Hadoop with Ceph FS

The Hadoop has become a hugely popular platform for large-scale data analysis. [10] The Ceph file system can be used as a drop-in replacement for the Hadoop File System (HDFS). The HDFS has a single metadata server that sets a hard limit on its maximum size. Ceph, a high-performance distributed file system under development since 2005 and now supported in Linux, bypasses the scaling limits of HDFS.

Ⅵ. Conclusion

The trend keywords in ICT sector will be Big Data, Internet of Things, and Cloud Computing. Building the IoT will be an exercise in integrating disparate devices and carrier networks, multiple communication protocols, and a wide variety of applications. It will often require integrating legacy networks and applications. The new M2M applications are and will continue to be complex, using geographically dispersed devices and services, a mix of connectivity, and logic in the data center and edge devices. The rear end to support those techniques requires a large-capacity storage technology of low-cost. Therefore, we have proposed the prototype of low-cost and mass distributed storage system based on PC using open-source Ceph, and tested the scale out function and performance of data saving speed for SMB.

REFERENCES

[1] Gartner www.gartner.com/

[2] U2N, http://www.u2n.kr/

[3] Margaret Rouse, "Definition: software-defined storage," SearchSDN. Tech Target. Retrieved November 7, 2013.

[4] SNIA (March 2014). "Technical Whitepaper: Software Defined Storage".

[5] Ceph, http://ceph.com

[6] Sage A. Weil, Scott A. Brandt, Ethan L. Miller, and Carlos Maltzahn, "CRUSH: Controlled, Scalable, Decentralized Placement of Replicated Data," 2006.

[7] InkTank, http://www.inktank.com/what-is-ceph/

[8] Snort, https://www.snort.org/

[9] OpenStack, http://www.openstack.org

[10] Hadoop, http://hadoop.apache.org

[11] Ganglia, http://ganglia.sourceforge.net/

Authors

ByungRae Cha

He is a research professor in school of Information and Communication Engineering and Super Computing & Collaboration ENvironment Technology (SCENT) Center, GIST, Korea. He received the Ph.D. degree in Computer Engineering from National Mokpo Univ. in 2004 and the M.S. degree in Computer Engineering from Honam Univ. in 1997. Prior to becoming a research professor at GIST, he has worked as a research professor in Department of Information and Communication Eng., Chosun Univ., and professor in Department of Computer Engineering, Honam Univ., Korea. His research interests include Computer Security of IDS and P2P, Neural Networks Learning, Cloud Computing, and Future Internet.

Yongil Kim

He received a B.S. in computer science from Chonnam University in 1984, and M.S. degree in computer science from Korea Advanced Institute Science and Technology (KAIST), Korea in 1986. From March 2002, he has joined as an Associate Professor at the Honam University, Gwangju, Korea. His research interests in data mining, big data, and intelligent agents. He is a member of IEEE.