Real-time and Flexible Management of Storage Service Provider in Distributed Storage

Longbo Ran and Hai Jin

Services Computing Technology and System Lab
School of Computer Science and Technology
Huazhong University of Science and Technology, Wuhan, 430074, China
hjin@hust.edu.cn

Abstract

In distributed storage system, storage service provider (SSP) is a critical component to receive and response these requests from clients. We propose a real-time and flexible management of SSP in this paper. With the advantages from virtualization and cloud computing, the states of SSP main VM are observed by another isolated virtual machine, and the administrator can inspect the storage service provider through the uniform management interface in real-time.

Keywords: *Distributed Storage, Global Distributed Storage System, Storage Service Provider, State Management*

1. Introduction

Data from scientific research, information service and digital media increase dramatically, and it is the most significant resource. Since the magnitude of data is from gigabyte to terabyte or petabyte, it brings tremendous challenge for the capacity and speed of storage system [1]. Until now, researchers have devised different types of data storage approach, such as *direct attached storage* (DAS), *network attached storage* (NAS) [2], *storage area network* (SAN) [3], *data grid* [4], *virtual storage* [5], *cloud storage* [6]. DAS refers to the storage device connected to computer directly. NAS refers to a computer connected to a network that only provides file-based data storage. SAN is used to make storage devices accessible to servers so that the devices appear like locally attached to it. DataGrid [7] is a project supported by European Union, it builds the next generation computing infrastructure providing intensive computation and analysis of shared large-scale databases. Virtual storage provides a transparent method to access the storage devices without considering the storage media. Cloud storage is a novel concept proposed by researchers in recent years, and data is stored in the virtualized resource pool generally hosted by the third party. It transforms the usage pattern of data, and provides extremely flexible for cloud users.

We had proposed a global distributed storage system, called *Global Distributed Storage System* (GDSS) [8]. It implements management and sharing of storage resource in *wide area network* (WAN), besides it shields the heterogeneity of the underlying data resource. *Storage Service Provider* (SSP) is the entrance of GDSS, and it provides the access point of all other modules. SSP is the critical component to provide storage service for all terminal users, and the running state of SSP should be guaranteed during the execution procedure of storage system. In order to implement transparence for GDSS, we propose a real-time state management approach in this paper. It can implement real-time and flexible management of SSP for GDSS.

The rest of this paper is organized as follows: section 2 introduces the related work and background. Section 3 introduces the architecture of GDSS and the architecture of SSP. Section 4 describes the implementation and the experiments of this structure. Conclusion is presented in section 5.

2. Related Works

Distributed storage has been studied for several years, and many papers are published to build a distributed storage system with high scalability, or high efficiency, or high security. The representative systems are as follows: DataGrid, *Storage Resource Broker* (SRB), *Distributed Parallel Storage System* (DPSS), GridFTP, and GridNFS.

DataGrid [7] project develops a grid infrastructure which links the various science Grids. The initial motivation of the DataGrid project is to aggregate the large scale of computing and data management grid.

SRB [9] is a data grid management system operating in computational science research projects. SRB middleware builds on top of various storage systems, real-time data sources, or a relational database management system. It has callable library functions that can be utilized by higher-level software.

DPSS [10] is a data storage system originally developed as one part of the DARPA-funded MAGIC test bed with the characteristics of scalability and high-performance. It is a data block server, which provides high-performance data handling and architecture for building high-performance storage systems from low-cost commodity hardware components.

GridFTP [11] is an extension of the standard *File Transfer Protocol* (FTP) for grid computing. It is defined as one part of the Globus toolkit, which is under the organization of the Global Grid Forum. The aim of GridFTP is to provide a more reliable and high performance file transfer approach for grid computing applications. This is necessary because of the increased demands of transmitting data in grid computing.

GridNFS [12] is a middleware solution, which extends distributed file system technology and flexible identity management techniques. GridNFS combines the NFSv4 protocol and a collection of supporting middleware services configured to run in Globus environment. GridNFS provides file system name space, which spans a virtual organization.

3. Design of Real-time Management Architecture

3.1 Brief Introduction of Global Distributed Storage System

Global Distributed Storage System (GDSS) is built atop SAN, NAS, or other storage systems. Its purpose is to manage and share storage resource in wide area network through the technology of storage virtualization. GDSS masks the heterogeneity of the underlying data resource, and provides transparent logical view to reduce the management overhead. Otherwise, it provides a data management service with the characteristics of high scalability, high efficiency, and high security in the data grid environment.

The architecture of GDSS is indicated in Figure 1. GDSS implements storage virtualization from three aspects: SSP provides uniform access interface and storage view for the clients; SA shields the concrete detail of storage system to adapt to the heterogeneous storage environment; GNS ensures the map relationship between logic storage pool and physical storage device.

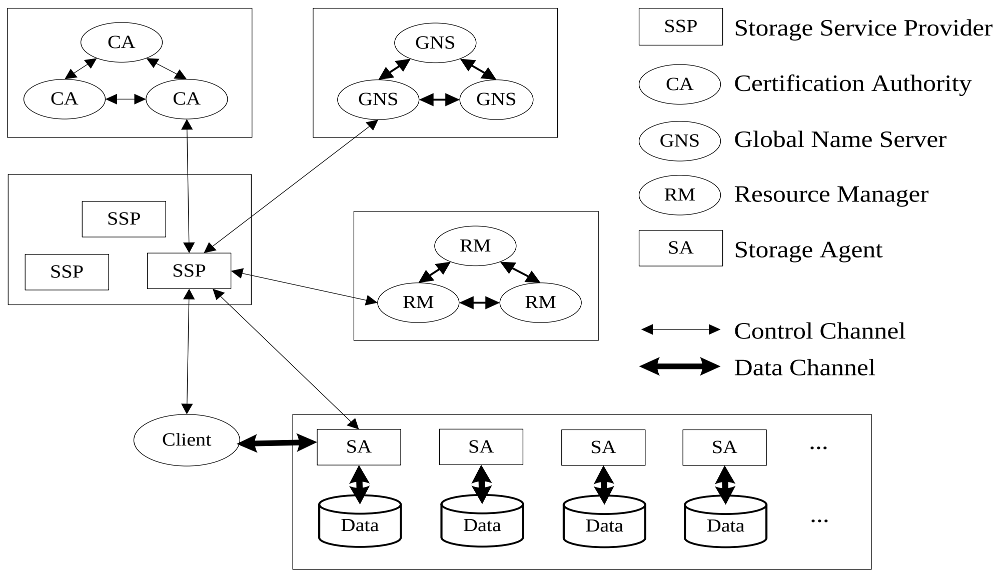

Figure 1. Architecture of GDSS

The basic function of each component is listed as follows:

(1) Storage service provider: SSP is the remarkable component of GDSS, and it was integrated to metadata server or resource manager in previous systems. However, when thousands of clients connect to GDSS, the time to deal with these connections will consume mass system resource. At the same time, the maximum connection is limited to operating system. On the other hand, metadata server and resource manager is the bottleneck of distributed storage system. In order to enhance the system capacity and mitigate the burden of metadata server or resource manager, SSP is separated with the metadata server. SSP is the entrance of GDSS, and the clients access other components through it. It provides 4 types of service interfaces: FTP interface, CA interface, GNS interface, and RM interface. SSP is not unique, and it can be loaded on demand. The method alleviates the burden of GNS, and improves the scalability of GDSS.

(2) Certification authority: CA certificates these components of GDSS, and build secure communication channel of each component. Furthermore, it ensures data operation under the control of access control policy.

(3) Global name server: GNS takes charge of metadata management, including the operating interface of metadata, the search of metadata and fault tolerant of metadata.

(4) Resource manager: RM implements the management of resource and copy. Resource management includes the map between logic metadata and physical metadata, the schedule and application of resource. Copy management includes copy creation, selection, distribution, and reclaiming.

(5) Storage agent: SA shields the underlying storage systems, including different operating systems (Windows, Linux, etc.) and diverse file systems (FAT32, NTFS, EXT2, EXT3, etc.). It provides interface to access data and file for the clients. It also provides interface to resource registration, information gathering, and garbage collection for SSP.

3.2 Real-time Management for Storage Service Provider

Storage service provider is the critical component in GDSS. On one hand, it receives these requests from the clients. On the other hand, it communicates with other components (CA, GNS, RM, and SA) to response user's requests. However, there are few papers to focus on the state of storage service provider. For this reason, we discuss the issue that will happen in the system, and give a real-time monitoring architecture for the system.

In the old versions of GDSS the monitor services are running on the SSP nodes. But there are some shortcomings. One is when the SSP node is overloaded the monitor service will become slowly because of competition of CPU or IO resources. On the other hand this is the most important time when the monitor service need to do something like load balance control.

With the development of virtualization and cloud computing technologies, the advantages of using these two technologies become more and more attractive.

In Figure 2 we provide a new structure of SSP. There are three basic virtual machine instances, one is SSP main VM, the other is SSP backup VM, and the third is SSP monitor VM. SSP main VM provides the most of the functions like old versions in GDSS. SSP backup VM is a hot backup of SSP main VM. They work at active/active mode or active/standby mode. So if the main VM breakdown the backup VM can take over the jobs. SSP monitor VM monitors the status of SSP main VM and SSP backup VM.

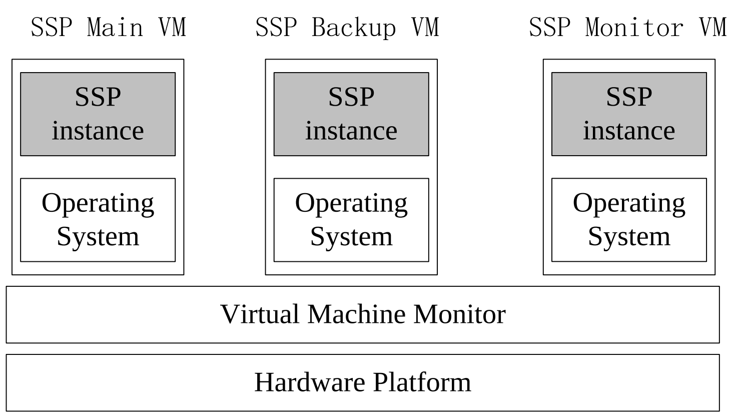

Figure 2. Structure of SSP

SSP monitor VM is the most important part of the system. It provides functions including: 1) monitoring the live status of SSP main VM and SSP backup VM, if one VM is dead, SSP monitor VM instance will start a corresponding VM instance; 2) monitoring the load of the other VM instances, if someone is overloaded, SSP monitor VM instance will trigger the load balance mechanism; 3) monitoring the SSP status, such as users, connections, CPU load and memory used.

With the development of virtualization, lots of hardware support virtualization directly, such as Intel and AMD newly released CPUs. The bottleneck of virtualization system often occurs on network. So on the SSP server there are at least two network cards, the fast network card is exclusively used by SSP main VM.

4. Implementation and Experiments

For a monitor system, the response time is a key value. In this structure the monitor service is running on an isolated VM, SSP monitor VM just runs the monitor service and the host machine can give enough CPU and memory to guarantee the service running fluently. If we can give the SSP monitor VM one exclusively network card, the whole monitor response time can be reduced to a very low level.

In order to prove our idea, we have built a new SSP with Xen [13][14][15]. There are 3 nodes: the old GDSS SSP is deployed on one node, the new GDSS SPP with three VMs is deployed on another node, the test client is running on the third node. All nodes are two Intel Xeon CPUs with the frequency of 2GHz, 8GB of memory, and 1TB of disk capacity. All nodes are pre-installed with Fedora 8 as the host operating system.

Figure 3 displays the experiment result. In the old GDSS SSP, if the CPU is overloaded the response time grows obviously. But in new GDSS SPP the response time is relatively stable with the CPU load changes. That means the response time is predictable and we can consider it close to real-time in new structure.

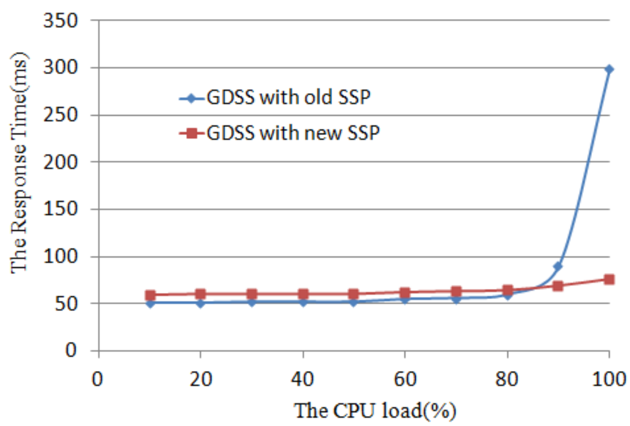

Figure 3. Response Time

4.2. Flexibility

As a global distributed storage system, the SSP need to be deployed on different servers with all kinds of hardware and software. It is really a very hard work. But with the advantage of virtualization and cloud computing, we can deploy the VM on different servers and the VM can provide uniform interfaces of storage services and monitor services.

5. Conclusions

Since the data from every walk of life increase dramatically, distributed storage is one of the most significant research areas of computer science. GDSS is a representational system, which manages and share data for isolated storage resource. Storage service provider is the remarkable component for GDSS, and its state has effect on the quality of service. A real-time state management method is proposed in this paper, and SSP monitor VM is deployed on the same host to monitor the SSP main VM independently, so the states of SSP main VM are observed by another isolated virtual machine, and the administrator can inspect the storage service provider through the uniform management interface. This method can supervise SSP with the characteristics of real-time and flexibility.

References

[1] S. A. Weil, S. A. Brandt, E. L. Miller, D. D. E. Long and C. Maltzahn, "Ceph: A Scalable, High-Performance Distributed File System", Proceedings of 7th Conference on Operating Systems Design and Implementation, **(2006)** November 6-8; pp.307-320. Seattle, USA.
[2] G. A. Gibson and R. V. Meter, "Network Attached Storage Architecture", Communications of the ACM, Vol. 43(11), pp. 37-45. ACM Press **(2000)**.
[3] J. Ward, M. O. Sullivan, T. Shahoumian and J. Wilkes, "Appia: Automatic Storage Area Network Fabric Design", Proceedings of Conference on File and Storage Technologies, (**2002**) January 28-30; Montery, USA.
[4] A. Chervenak, I. Foster, C. Kesselman, C. Salisbury and S. Tuecke, "The Data Grid: Towards an Architecture for the Distributed Management and Analysis of Large Scientific Datasets", Journal of Network and Computer Applications, Vol. 23, pp. 187-200 **(2000)**.
[5] The Wikipedia of Virtual Storage, http://en.wikipedia.org/wiki/Virtual_storage.
[6] W. Y. Zeng, Y. L. Zhao, K. R. Ou and W. Song, "Research on cloud storage architecture and key technologies", Proceedings of the 2nd International Conference on Interaction Sciences: Information Technology, Culture and Human, **(2009)** November 24-26; pp. 1044-1048. Seoul, Republic of Korea.
[7] The DataGrid Project, http://eu-datagrid.web.cern.ch/eu-datagrid/.
[8] L. Ran, H. Jin, Z. Wang, C. Huang, Y. Chen and Y. Jia, "Architecture Design of Global Distributed Storage System for Data Grid", High Technology Letters, Vol. 9, pp. 1-4. **(2003)**.
[9] C. Baru, R. Moore, A. Rajasekar and M. Wan, "The SDSC Storage Resource Broker", Proceedings of the Centre for Advanced Studies on Collaborative Research, **(1998)** November 30-December 3, pp.189-200. Toronto, Canada.
[10] The Distributed-Parallel Storage System (DPSS) Home Page, http://www-didc.lbl.gov/DPSS/.
[11] W. Allcock, J. Bresnahan, R. Kettimuthu, M. Link, C. Dumitrescu, I. Raicu and I. Foster, "The Globus Striped GridFTP Framework and Server", Proceedings of the 2005 ACM/IEEE Conference on Supercomputing, **(2005)** November 12-18; pp.1-11. Washington, USA.
[12] P. Honeyman, W. A. Adamson and S. McKee, "GridNFS: Global Storage for Global Collaborations", Proceedings of Local to Global Data Interoperability – Challenges and Technologies, **(2005)** June 20-24; pp. 111-115. Ann Arbor, USA.
[13] P. Barham, B. Dragovic, K. Fraser, S. H. T. Harris, A. Ho, R. Neugebauer, I. Pratt and A. Warfield, "Xen and the Art of Virtualization", Proceedings of the 19th ACM Symposium on Operating Systems Principles, **(2003)** October 19-22; pp. 164-177. New York, USA.
[14] B. Clark, T. Deshane, E. Dow, S. Evanchik, M. Finlayson, J. Herne and J. N. Matthews, "Xen and the Art of Repeated Research", Proceedings of the 2004 USENIX Annual Technical Conference, **(2004)** June 27-July 2; pp.135-144. Boston, USA.
[15] I. Pratt, K. Fraser, S. Hand, C. Limpach, A. Warfield, D. Magenheimer, J. Nakajima and A. Mallick, "Xen 3.0 and the Art of Virtualization", Proceedings of the 2005 Linux Symposium, (**2005**) July 20-23; pp.65-85. Ottawa, Canada.