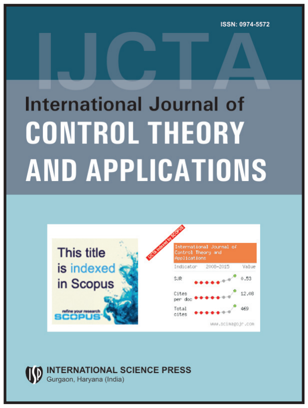

International Journal of Control Theory and Applications

ISSN : 0974-5572

Volume 9 • Number 44 • 2016

A Survey on IN-SITU Metadata Processing in Big Data Environment

Anuja Kulkarni[a] and Naveenkumar Jayakumar[b]

[a]*M.Tech, Department of Computer Engineering, Bharati Vidyapeeth Deemed University, Pune, India. Email: anukul2009@gmail.com*
[b]*Associate Professor, Department of Computer Engineering, Bharati Vidyapeeth Deemed University, Pune, India. Email: naveenkumar@bvucoep.edu.in*

Abstract: Big data environment judged on characteristic - storage performance. Flawless storage performance is needed to accelerate processes in respect to gain business characteristics such as innovation, profitability, productivity, and greater discovery. Data throughput can be increased to some extent using storage virtualization and parallel data paths. In present architecture of out-of-band storage virtualization system Agents on host/server machine indirectly increase the load on the server as it has to perform functions of path control, data access, data preservation, data presentation. Hypothesis of research is that Server/Host should primarily focus on executing or processing the business/user centric applications rather than spending their CPU cycles on various storage management services and operations.

Keywords: Big data, Metadata, Lustre, Active Storage, Out-of-band, Virtualization.

1. INTRODUCTION

The World today is generating a huge amount of data through different systems and devices. This large data is known as Big Data. High influence of metadata processing is thereupon Big Data processing. Metadata gives us information about data objects. It provides information underlying any artifact, such as - author, location, date of creation and modification, title, subject, and its relative information as per the need of business enterprise [1].

Everytime a file is accessed, some portion of the metadata is also accessed. Thus metadata operations are critical in any applications point of view. Research shows that more than 75% calls to file systems need access to file metadata. Thus, management of this metadata is crucial in the context of system performance. Several tasks in metadata management activity are to ensure – create, store, and control of metadata to remove redundancy and inconsistencies [1].

Big data environment judged on characteristic - storage performance. It is highly noted when the client is waiting for input and output from the storage system [2]. Flawless storage performance is needed to accelerate processes in respect to gain business characteristics such as innovation, profitability, productivity, and greater discovery. Data throughput can be increased to some extent using Storage virtualization and Parallel data paths. Storage Virtualization provides a logical abstraction of physical storage resources to the hosts.

Along with data parallelism, it would be an advantage to boost up performance if we can restrict burgling metadata access in the data path. Legacy architectures usually do not contain this [2].

Processing of metadata outside the data path and enabling client to access file data directly in object storage devices (OSDs) can expel most of the metadata related performance issues, and its effects.

To separate data path and control path can be done by using out-of-band architecture. In this, the virtualized environment configuration stored external to the data path. Thus, virtualization engine achieves a split of the 'access path to metadata' to 'access path to data.' By splitting data path and control path, virtualization does not create a congestion to data transport.

Virtual Storage Manager (VSM) is responsible for address mapping in out-of-band virtualization, and its location is outside of data path [3]. Thus, performance and scalability get increased. For this, it needs different agents on different platforms. The role of VSM is to maintain metadata and use it during address mapping. In each of the I/O request, the agent has to query VSM for address mapping. Experimental studies show that this method brings communication latency and hence reduces the throughput of disks [3]. Agents on host/server machine indirectly increase the load on the server as it has to perform different functions of path control, data access, data preservation, data presentation.

The proposed architectural design shows that the load on server/host can be reduced by adding intelligence to the virtualization appliance or metadata appliance so that metadata consistencies and data management will be handled by the Metadata Appliance freeing the Server from initiating and monitoring these activities.

2. RELATED WORK

The literature reviewed suggests various challenges encountered due to increase in the quantity and distributed pattern of data. In large distributed storage systems metadata management is a critical aspect of overall system performance. Data throughput to and from storage can be increased using the concept of storage virtualization and parallel data paths. Along with data parallelism, avoiding performance robbing metadata access in the data path can greatly accelerate the performance which can be achieved through out-of-band storage virtualization system [2].

The literature survey revealed that in the current architecture of out of band the interaction between agent and virtual storage manager (VSM) brings too much communication latency, thus reducing the throughput of the disk [3].

The overall solution proposed by various authors was based on various hardware enhancement. In the research article proposed by authors new out-of-band virtualization system-MagicStore were presented which is not hardware specific and supports three widely used host platforms: Solaris, Linux, Windows [10].

The literature survey also revealed that, the researchers who have worked on bringing the compute part closer to data termed it as Active Storage which was realized through the concept of Active Disk which is discussed in[5,6]. The key idea used in Active disk architecture is offloading the bulk of computation to the disk resident processor where data resides and to use the host processor primarily for coordination, scheduling, and the combination of results from individual disks.

A. Current Storage Issues

Big data environment judged on characteristic - storage performance. It is highly noted when the client is waiting for input and output from the storage system. The increase in storage performance is necessary to accelerate workflows for greater discovery, productivity, profitability, and innovation. The storage architecture must be designed such that it overcame the limitations of standard network protocols and serialized data paths to achieve greater performance.

Enterprises use Network-attached storage (NAS) instead of direct-attached storage(DAS) to store such a massive data because it scales better than DAS. The challenges for network attached storage increases with scalability. To achieve linear scalability, high performance, and high reliability most NAS products scale to some degree but it creates the problem in data management [2].

This issue can be solved by implementing virtualization using out-of-band architecture. Processing of metadata outside the data path and enabling client to access file data directly in object storage devices (OSDs) can expel most of the metadata related performance issues, and its effects[10].

In an out-of-band virtualization system, the virtual storage manager (VSM) which is responsible for address mapping is located outside of the data path, so as to improve performance and scalability, but different Agents are required on different platforms. Since the agents are on host/server machine, they indirectly increase the load on the server as the server has to perform various functions like data presentation, data preservation, data access, path control, etc.

B. Lustre File System

Lustre file system is a large-scale parallel file system which plays a key role in High-Performance Computing (HPC) system. Lustre file system can serve tens of thousands of client systems can have petabytes of storage with an I/O throughput of hundreds of gigabytes per second. Components of Lustre file system are - Metadata server (MDS), metadata target (MDT), Object storage server (OSS), object storage target (OST) and lustre clients. The MDS makes metadata available to Lustre clients which is stored in one or more MDTs. The MDT stores metadata such as a name of a file, directories, permission on file. For one or more local OSTs the OSS provides file I/O service and network request handling. File data is stored in one or more objects, and every object is stored on a separate OST in a Lustre file system. Lustre clients are computational nodes running Lustre client software[12].

In Lustre file system high performance and scalability can be achieved by providing abstraction at different levels. Lustre treats files as objects at file system level that are located through metadata servers (MDSs).

When a client wants to write some data, it communicates with the MDS with a write request. The MDS checks for user authentication and the intended location of the file. According to file system settings, the MDS sends back a list of OSTs to the client. After receiving a reply from MDS, the client can now exclusively interacts with the assigned OSTs without having to communicate with the MDS.

C. Active Storage

Active Storage is a computer system architecture that executes application code by utilizing processing power in disk drives. The basic idea behind Active Storage (also referred as Active Disk or Intelligent Disk) is to offload computation task and data traffic from host computers to the disk drives. Active Storage can be used to achieve high performance for data-intensive applications such as multimedia applications and decision support systems[14].

D. Lustre with Active Storage

Lustre file system embedded with Active Storage can tackle the problem of data management in the context of high-performance computing systems. Lustre file system use main stream server computers as storage nodes which are very similar or identical to the compute nodes present in a cluster. The underutilized CPU time in the storage nodes can be utilized to perform computation when it is extended with Active Storage[13].

Active storage concept can enhance the I/O throughput of Lustre file system as it provides an opportunity for reducing the amount of data flow between the storage node and compute node. Using Active Storage, it is possible to considerably reduce the data movement across the network thereby reducing the overall network traffic. In Lustre file system Active Storage concept can be implemented either in a kernel space or a storage space.

3. PROPOSED SCHEME

The proposed system is a framework which can be implemented using Active Storage concept in Lustre file system.

A. Proposing Active Metadata Storage

1. Virtualization technique will be applied to Metadata server/appliance in out-of-band storage infrastructure.
2. File system consistencies and metadata updates need to be tracked, recorded by this metadata appliance without the intervention of the server.
3. This metadata appliance should able to carry out operations like : Volume management, Dynamic volume migration, Point-in-time copies, LUN virtualization, Data mobility, Copy services.

B. Integrating the Active Storage Concept with Out-of-band Storage Metadata Appliance

1. Metadata Appliance will understand the I/O Pattern between the various server and storage. A thermal model can be developed here for further decision making.
2. Based on the thermal Model, the offloading decisions can be taken to perform the near data processing.

C. System Architecture

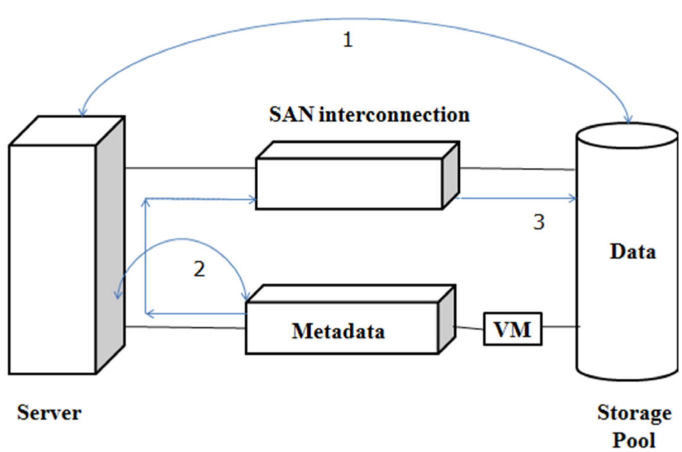

Figure 1: Proposed System Architecture

In proposed system architecture path 1 is Data path, path 2 is Metadata access, and path 3 is Control path. Active storage concept will be integrated with Out-of-Band storage metadata appliance.

Active storage system takes the advantage of underutilized CPU time on file system servers to process the data instead of utilizing the time just to store the data.

The key idea used in Active disk architecture is offloading the bulk of computation to the disk resident processor where data resides and to use the host processor primarily for coordination, scheduling, and the combination of results from individual disks.

File system consistencies and metadata updates need to be tracked, recorded by the metadata appliance without the intervention of the server. Metadata Appliance will understand the I/O Pattern between the various server and storage. A thermal model can be developed here for further decision making. Based on the thermal model, the offloading decisions can be taken to perform the near data processing.

In out-of-band storage virtualization system, data transmission and address mapping are both done by the virtual storage manager (VSM). VSM is present outside the data path, to improve performance and scalability and different software Agents are present on each Host/server machine.

The VSM knows the states of physical devices and manages virtualization metadata. The software Agents which are connected to the manager over an Ethernet has to query VSM to create logical volume devices and to perform address mapping from logical address space to the physical address space.

4. CONCLUSION

Conventional out-of-band storage virtualization system can be used to achieve high performance by processing metadata out of data path, but the interaction between Agent and Virtual Storage Manager increase communication latency and reduce the throughput of the disk. By using metadata appliance and Active Disk in out-of-band architecture, a load of metadata management on the server can be reduced to some extent thereupon achieving high performance and throughput of the disk.

Acknowledgment

The authors would like to thank all the Bharati Vidyapeeth College of Engineering staff members for their valuable inputs and support.

REFERENCES

[1] L. Management, "Metadata Management : An Essential Ingredient for Information Lifecycle Management," *Access*, No. October, 2005.

[2] "Copyright © 2015 Panasas, Inc. www.panasas.com November 2015," No. November, 2015.

[3] A. W. L. a N. Chipset and T. H. Meng, "Design and Implementation of an," Vol. 50, No. August, pp. 160–168, 2003.

[4] EMC, "Networked Storage Virtualization," pp. 1–10.

[5] E. Felix, K. Fox, K. Regimbal, and J. Nieplocha, "Active storage processing in a parallel file system," *Proc. 6th LCI …*, 2006.

[6] A. Acharya, M. Uysal, and J. Saltz, "Active disks: programming model, algorithms and evaluation," *SIGPLAN Not.*, Vol. 33, No. 11, pp. 81–91, 1998.

[7] Y. F. Y. Fu, N. X. N. Xiao, and E. Z. E. Zhou, "A Novel Dynamic Metadata Management Scheme for Large Distributed Storage Systems," *2008 10th IEEE Int. Conf. High Perform. Comput. Commun.*, No. September, pp. 987–992, 2008.

[8] R. Salunkhe, A. D. Kadam, N. Jayakumar, and S. Joshi, "Luster A Scalable Architecture File System: A Research Implementation on Active Storage Array Framework with Luster file System.," *Iceeot*, 2015.

[9] S. N. Technologies and E. M. C. P. Professional, "Storage Virtualization," 2009.

[10] G. Zhang, J. Shu, W. Xue, and W. Zheng, "Design and implementation of an out-of-band virtualization system for large SANs," *IEEE Trans. Comput.*, Vol. 56, No. 12, pp. 1654–1665, 2007.

[11] J. Zhang, X. Jiang, Y. Tian, X. Qin, M. I. Alghamdi, M. al., Assaf, and M. Qiu, "ORCA: An offloading framework for I/O-intensive applications on clusters," *2012 IEEE 31st Int. Perform. Comput. Commun. Conf. IPCCC 2012*, pp. 81–90, 2012.

[12] T. K. Petersen, "Inside The Lustre File System."

[13] J. Piernas and E. J. Felix, "Evaluation of Active Storage Strategies for the Lustre Parallel File System," No. 1.

[14] https://link.springer.com/referenceworkentry/10.1007%2F978-0-387-39940-9_1326

[15] "Evaluation of a Performance Model of Lustre File System," No. July, 2010.