Chapter 5
Metadata Management in PetaShare Distributed Storage Network

Ismail Akturk
Bilkent University, Turkey

Xinqi Wang
Louisiana State University, USA

Tevfik Kosar
State University of New York at Buffalo (SUNY), USA

ABSTRACT

The unbounded increase in the size of data generated by scientific applications necessitates collaboration and sharing among the nation's education and research institutions. Simply purchasing high-capacity, high-performance storage systems and adding them to the existing infrastructure of the collaborating institutions does not solve the underlying and highly challenging data handling problem. Scientists are compelled to spend a great deal of time and energy on solving basic data-handling issues, such as the physical location of data, how to access it, and/or how to move it to visualization and/or compute resources for further analysis. This chapter presents the design and implementation of a reliable and efficient distributed data storage system, PetaShare, which spans multiple institutions across the state of Louisiana. At the back-end, PetaShare provides a unified name space and efficient data movement across geographically distributed storage sites. At the front-end, it provides light-weight clients the enable easy, transparent, and scalable access. In PetaShare, the authors have designed and implemented an asynchronously replicated multi-master metadata system for enhanced reliability and availability. The authors also present a high level cross-domain metadata schema to provide a structured systematic view of multiple science domains supported by PetaShare.

DOI: 10.4018/978-1-61520-971-2.ch005

INTRODUCTION

The unbounded increase in the size of data generated by scientific applications, such as high energy physics (Carena et al., 2008) (Newman, 2003), computational biology (Ponomarev, Bishop, & Putkaradze, 2009) (Yang & Yang, 2008), coastal modeling (Chen, Zhao, Hu, & Douglass, 2005) (Stamey, Wang, & Koterba, 2007), computational fluid-dynamics (Gaither, 2007) numerical relativity (Allen, Goodale, Masso, & Seidel, 1999), and astrophysics (Tyson, 2002) necessitates collaboration and sharing data among the nation's education and research institutions. Having stringent performance requirements, large volume of data sets, and geographically distributed human, computational and storage resources makes existing data management infrastructures insufficient (Chervenak, Foster, Kesselman, Salisbury, & Tuecke, 1999). Simply purchasing high-capacity, high-performance storage systems and adding them to the existing computing infrastructure of the collaborating institutions do not solve the underlying and highly challenging data handling problems. Scientists are compelled to spend a great amount of time and energy on solving basic data-handling issues, such as how to find physical location of data, how to access it, and/or how to move it to visualization and/or compute resources for further analysis.

There is a wide variety of distributed file systems developed to alleviate data management challenges in cluster environment, such as AFS (Howard et al., 1988), NFS (Shepler et al., 2003), Lustre (Lustre, 2010), PVFS (Ligon, & Ross, 1996), GPFS (Schmuck, & Haskin, 2002) and Panasas (Nagle, Serenyi, & Matthews, 2004). These file systems are sufficient and widely used in LANs as a cluster file system. However, when the volume of generated data sets increases and data sets are distributed over the clusters through WANs, it becomes very expensive to maintain a unified shared file system running across distributed clusters. This is due to the constraints of WAN, heterogeneity of distributed resources and environments, and authorization/authentication policies of different administration domains. To address the challenges of data handling issues in geographically distributed and heterogeneous environments, distributed data storage systems have been proposed and implemented.

Distributed data storage systems provide flexible mechanisms for controlling, organizing, sharing, accessing and manipulating data sets over widely distributed resources that are under the control of different administration domains. One of the important features of distributed storage systems is providing global unified name space across distributed resources, which enables easy data sharing and accessing without the knowledge of actual physical location of data. This is known as 'location transparency'. The location transparency of distributed data sets is provided efficiently by distributed data storage systems. Distributed storage systems enable scalable, efficient and transparent access to the distributed resources including replicated data sets in different resources to enable fast access while ensuring data coherency. They can issue fine-grained authentication and authorization policies over shared and distributed resources as well as data sets. In general, the data sets stored in distributed resources are accessed through data servers, while metadata is managed separately by metadata servers in distributed data storage systems for efficiency.

The NSF funded PetaShare project aims to enable transparent handling of underlying data sharing, archival, and retrieval mechanisms, and make data available to scientists for analysis and visualization on demand. The goal is to enable scientists to focus on their primary research problems, assured that the underlying infrastructure will manage the low-level data handling issues. In the design and implementation of PetaShare, a novel approach has been employed to solve the distributed data sharing and management problems. Unlike existing approaches, PetaShare treats data storage resources and the tasks related to data

Related Content