panzura

Panzura White Paper

Metadata

Panzura's game-changing Distribute Cloud File System finally brings the full power and benefits of cloud storage to enterprise customers, helping to break the unending on-site storage expansion cycle while eliminating islands of storage that inhibit cross-site user interaction and productivity and real-time data protection. Panzura makes deploying cloud storage and a distributed cloud file system easy and transparent to users.

In this white paper, we explore how shared Metadata provides the foundation for enterprise collaboration.

Introduction

In 1969 the term “metadata” was coined by Jack E. Myers to represent products that implemented his idea of the MetaModel. Today this term is more generally described as “data about data”; information that helps us describe and locate objects. From the card catalog in the public library that we use to locate books on library bookshelves, to the yellow pages that we use to find business addresses, metadata provides us with information about objects that we visit, touch, and use.

This white paper reviews a particular type of metadata that we use every day, the metadata used in a file system. In this case, the metadata provides an index to the files we store on our smartphones, laptop computers and enterprise workstations. When we use these devices to locate a photograph, video or document, we are actually walking through file systems using the file system metadata.

Consider then what happens when we make multiple copies of the file system metadata, and share them with sites around the world. Every user who accesses a local copy of the shared metadata would see an identical version of the file system. In a nutshell, shared metadata is the foundation of the Panzura Distributed Cloud File System (Panzura DCFS). Panzura stores a copy of the file system metadata at each site in order to provide users with a single view of an entire global file system, no matter where they are located.

The purpose of this white paper is to describe Panzura’s patented technology for sharing metadata, and how it forms the foundation of the Panzura DCFS.

What is Metadata?

Metadata can be defined as “data that describes and gives information about other data”. In the case of a file system, the metadata is used to describe directories, files, and file attributes. Particular file attributes include file size, creation date, last modified date, file owner, and file access permissions.

When users interact with a file system, the bulk of their actions consist of navigating through directory structures, opening folders, looking through lists of files, and sorting files based on name, size, type, and date. These actions make use of the filesystem metadata but do not require access to the actual contents of the files.

For example, when Windows Explorer is used to browse the file system, the file system metadata is displayed in the Explorer window. The actual data in a file is not accessed until the file is opened, which is usually triggered by double-clicking the filename with the mouse.

SSDs Provide High Performance Access to File System Metadata

The speed with which a file system responds to individual file commands is known as the response time of the file system. In a shared file server, one that is storing files for hundreds or thousands of users, the needs of thousands of users must be considered in the design of the file system. If you observe the way that individual users interact with a file system, they will navigate through multiple file directories and view the file attributes before they determine which file to open and edit. When navigating directories, the response time of the file system is determined by the speed with which the system responds to metadata requests. When thousands of users simultaneous navigate the directories of a shared file system, the system must serve thousands of small random read requests per second, which are measured as IOPS (Input-output Operations Per Second).

Once the user identifies and opens a file, a high performing file system will serve the data quickly and without pausing. The throughput with which a file system reads sequential blocks of data and serves them to the user is measured in MBps (Mega Byte per second).

A high performance shared file server must respond equally well to both directory navigation commands (metadata IOPS) and file commands (throughput MBps), so these types of data are usually separated, stored and served from the physical storage technology that is best suited to serve them.

Solid State Drives (SSD) based on flash memory storage technology provide superior response times for random IOPS-based workloads. A single SSD provides levels of IOPS performance that can be only be accomplished by using hundreds of traditional hard disk drives. When SSD technology is used to serve metadata IOPS, users experience a level of responsiveness that is the same, or better, then they achieve when browsing their local file systems. SSD technology costs more than Hard Disk Drive (HDD) technology, but systems do not need a lot of it to provide the performance that is desired.

On the other hand, HDD technology is designed to provide the throughput (MBps) required for sequential file data access. When the right amount of fast storage technology (SSD) is combined with low cost storage (HDD) it is possible to produce a shared storage system that achieves both fast response time and high throughput in one solution.

Cloud Storage Transformation

The explosion of utility cloud computing provides enterprise customers with access to highly scalable and durable file storage solutions for an incredibly low cost. Cloud storage allows any amount of data to be stored and retrieved, at any time, from anywhere in the world. The storage is accessed over a wire, but in this case the wire happens to be the Internet.

Recognizing that customers want to use the cloud to store backup and archive data, cloud storage gateways are being deployed to connect the corporate LAN with low cost cloud storage. These gateways are being used to move backup and archive data to cloud storage. While this decreases the costs of storing archive and backup data, it fails to provide the performance that customers require for enterprise storage.

A high performance file system that stores files in cloud storage is made possible by combining three elements: a file system that separates metadata from file data, SSD-based controllers, and shared cloud storage. Panzura is an example of a high performance file system that leverages cloud storage to lower costs. The Panzura DCFS connects these three elements together to create storage that provides both fast response times and offers the low cost of cloud storage.

But there is another inherent feature of cloud storage that can be leveraged to create an even more revolutionary file system. Since cloud storage can be accessed at any time, from anywhere around the world, then it can be shared by any number of globally distributed controllers and thus form the basis for a globally shared file system. A globally shared file system is a single file system that can be accessed by any number of users, in any number of locations, around the world. With a single file system there is no need to duplicate files or to mirror data between sites, as there is only one file system and only one copy of the files.

The Panzura DCFS is a high performance global file system that connects sites around the world so users can work together as if they were located in the same office.

Panzura and the Power of Shared Metadata

Panzura's revolutionary ability to provide a single, globally consistent file system is made possible through technology that shares file system metadata and atomic metadata updates between controllers at globally distributed sites in real-time. With shared file system metadata on each controller, clients who connect to their local Panzura Freedom Flash Cache can view and access the entire distributed cloud file system. With one golden copy of the files in the cloud, the users can access individual files from any site around the world.

Panzura file system metadata contains a complete view of the file system directory structure and the attributes of individual files. If 100 copies of this metadata were made, each of the 100 metadata copies would show the entire file system. When this metadata is placed on Panzura controllers at sites around the world, each site has a complete view of the file system.

The file system metadata is served from the SSD on each system. As users navigate the Panzura file system they appear to be viewing files and folders, but they are actually browsing the file system metadata. Their experience is identical to working with a local file system because the user is, in fact, browsing a local copy of the file system metadata. Only after a user requests to open a file will the golden copy of the file be fetched from the cloud storage and presented to the user.

When a user updates a file, the file system metadata is also updated and synchronized with other Panzura Flash Cache systems. These small changes are easily shared between Flash Cache systems. Synchronizing small atomic updates to the file system metadata ensures that all users see one consistent, up-to-date view of the entire distributed cloud file system.

Distributed Cloud File Systems Require Distributed File Locking

It may seem like an obvious statement, but file systems must not corrupt the files they store. When multiple clients share access to a common file repository, provisions must be made to ensure that two or more clients do not make conflicting changes to the same files. A shared file system uses a lock manager to control and grant locks on individual files. When a file is locked, other clients may read the file, but they must wait until the lock is relinquished before they can write to the same file.

A global file system also employs a lock manager to control read and write access to individual files, but in this case it is called a global lock manager. There are two types of lock managers that can be used with a distributed cloud file system, central and distributed.

As their name implies, central lock managers use a single lock manager to manage file locking. This lock manager is usually located in the cloud so that all sites have equal access to the lock manager. Each controller contacts the central lock manager to view and request locks for a file. The performance of this solution can suffer if there is network latency between the controllers and the cloud. It can also suffer performance issues in larger deployments as the single lock manager attempts to work through large numbers of locking requests received from all the individual controllers. These design issues effectively cap the performance of the lock manager and possibly affect the performance of the global file system.

Distributed lock manager technology solves these performance issues by deploying individual file lock managers to each system. The individual lock managers provide fast performance for the local users, and lock privileges are shared between controllers as needed. When a user on a remote Freedom system requests write access to a file that is currently being locked on the local system, the ownership of that file, and the permission to lock the file, is passed to the remote user. At that point the locking for that file will be local to the users on the remote system.

Panzura has chosen to employ distributed file locking technology to provide high performance file locking and guarantee consistency for the Panzura Distributed Cloud File System. The Panzura distributed file locking technology ensures that only one person at a time is allowed to edit a file while maintaining the speed and efficiency of a localized storage system.

Summary

Metadata provides an index into a file system. This index gives a path to every file and directory in the file system. If we separate metadata from file data, then we can accomplish several goals:

Goal #1 - Performance

A measurement of file system performance is the speed with which it responds to commands to walk directories and view file attributes, which are translated as requests to view file system metadata. Within the file system metadata access patterns are seen as high amounts of small and random IOPS which is a perfect workload for solid-state storage (SSD) technology. A file system design that separates metadata from file data, and uses SSD technology to store metadata provides users with excellent response times.

Goal #2 – Distributed Cloud File System

When copies of file system metadata are synchronized with sites around the globe, users at those sites have the same view of the file system. With the addition of distributed file locking and atomic metadata updates between sites, those same users can now edit files and collaborate on projects in the same file system but from great distances. By sharing the file system metadata the file system is transformed into a distributed cloud file system.

Goal #3 – Low Cost File Storage

With a file system architecture that separates file data from file system metadata, it is possible to employ a storage solution that balances throughput with low cost. The emergence of cloud storage provides the optimal place to store file data as it is the lowest cost solution for storing large amounts of file data.

The Panzura Distributed Cloud File System achieves all three of these goals. Panzura serves file system metadata from an SSD for high performance, shares and distributes atomic metadata updates between systems to provide users with one view of a globally shared file system, and leverages cloud storage to provide low cost file storage.

Panzura uses these technologies so that users located at sites around the globe, have the same high-performance file access as they would when working side-by-side in the same office.

Panzura, Inc. | 695 Campbell Technology Pkwy #225, Campbell, CA, USA | 855-PANZURA | www.panzura.com