TANEJA GROUP
TECHNOLOGY ANALYSTS

T E C H N O L O G Y B R I E F

Archiving Beyond File Systems: Object Storage

EMC Centera And Disk Archival

January 2009

THE NET NET

Disk-based archiving answers many challenges in an organization, but this strong trend also creates questions for end users. Specifically, because of their distinct architectural approaches. Taneja Group sees some confusion regarding the question of whether to deploy a traditional file system or an object storage (e.g. Content Addressed Storage, or CAS) approach in support of an enterprise archival initiative. While it may not appear critical at initial deployment, we believe that the wrong choice in the "file system vs. object storage" question will lead to far-ranging challenges that compound over the course of an archive's lifetime.

Taneja Group has spent significant time researching object storage archiving and we firmly believe that Content Addressed Storage provides differentiated business value and a lower total cost of ownership over traditional file system based approaches for long-term online disk archival requirements. In this brief we will examine the world of file system based archiving, then provide a comparative look at the advantages of a CAS solution such as EMC Centera.

Changed Game: Disk Archival

Taneja Group has spent many hours speaking with both prospective and existing disk archival end users. Across all of these interactions, one commonality comes through clearly: disk archival has changed the game with fundamentally unique requirements that distinguish it from the tape and optical world. We find that some end users come to this realization early in their selection process while others discover after their initial deployment that they have a new kind of "beast" on their hands.

Some of the key characteristics that we see defining the unique and emerging requirements of disk archival can be summarized as follows

- **Hyper-scalability.** As disk-based archiving becomes the preferred method for long-term content preservation, we have seen the need for unprecedented scalability reaching into the tens, hundreds and thousands of terabytes. We observe that these scalability growth rates are further compounded as some administrators are retaining as much content as possible in the readily accessible disk medium as opposed to sending data to an offline "static" archival on tape media. The speed and ease of use of disk-based archives has in fact made it practical for administrators to create what Taneja Group defines as "Active Archives," disk-based archives where an organization's information is likely to be retained for long periods of time

is moved to the archive much earlier than ever before and it is used from there rather than from primary storage. By doing so, the "Active Archive" becomes a very cost effective extension of primary storage allowing an organization to better manage primary storage capacity utilization and reduce the overall cost of storage and its management.

- **Centralized archives.** A properly architected disk-based archive changes stored data into a readily available, highly usable information asset. Because of this fact, we have seen organizations increasingly approaching their disk archives from an infrastructure-wide perspective. Specifically, we observe the trend that organizations want to deploy a *centralized* archiving platform in support of all relevant business operations. This trend towards centralized archives is driven by a number of factors, including total cost of ownership, internal governance, regulatory compliance, and storage consolidation projects across an organization. We have examined that in a high-growth disk archive, the alternative approach of supporting individual archive "silos" on a per-application basis has proven itself to be fundamentally unmanageable as these repositories grow in capacity over time.

- **Dynamic application support.** Because disk-based archiving often touches many applications (e.g. content management, email, file data, proprietary applications) disk-based solutions must be able to provide an abstracted view into all of the supported applications in a seamless fashion. This manner of dynamic application support has been historically absent in disk-based archiving solutions that instead were structured as application "silos", each with their own archival content associations.

  Going further, we have observed that disk-archiving solutions are increasingly required to support multiple "views" across all of these applications, providing the end users with the ability to perform complex, simultaneous queries for data based on a range of programmable, business-relevant characteristics (e.g. various content attributes, usage history, and application associations.)

- **Long-term online.** One of the interesting but little noted qualities of disk-based archiving is its tendency to become an attractor for more and more archival content. Regularly, we speak with end users who share that their growth rates in disk archives have exceeded their best projections prior to deployment. Upon examination, the reasons become clear: disk-based archives, because of their "online and always available" status, transform an organization's traditional relationship with archived content. Specifically, disk archives have enabled users to access and

retrieve stored content within the context of their normal usage patterns. The historical "retrieval gap" that prevented offline and nearline archive content from playing an active role in real-time business has been removed.

As a result, today, archived information is playing a more strategic role in workflows and business processes. With this increased access to information, the data repositories are growing at a accelerated rate with an ever-increasing requirement for immediate access. Our client engagements show that this general "always on" quality of disk-based archiving will persist over the lifetime of the archive, creating the challenging requirement that solutions be both supportable over many decades and still always available to users, on demand.

The File System Challenge

Given the unique characteristics of disk archiving outlined above, it is no wonder that we see increasing numbers of end users asking serious questions regarding the ability of their traditional file systems to deploy, scale, and manage disk archives effectively.

The various questions regarding file systems result from one core technical issue: traditional file systems access and manage data in a hierarchical fashion, with significant dependencies on both the application and operating systems with which they are associated. As a result of that decades-old design principle, traditional file systems face undeniable challenges when it comes to supporting an enterprise disk archive with the profile provided above. Taneja Group has grouped these challenges into three general categories that we encourage end users to consider in their disk archival evaluation process.

Challenge: File System Lock-In

Because file systems straddle the kernel and user levels of a computing system, they create necessary dependencies on both the operating systems (OS) and applications of their hosts. Over the years, these OS and application dependencies have fostered sophisticated software innovations that have *abstracted* file systems in appropriate and useful ways (e.g. cluster file systems, virtual machines, application clustering.)

However, when placed in the context of today's disk-based archiving demands, these sophisticated augmentations to file systems are of little to no assistance in freeing the archive from "lock-in" to a specific application and OS.

Specifically, the challenge resides in how file systems store and retrieve data. File systems store data in a hierarchical fashion, always relying on the data's placement within a file and directory structures for its storage and retrieval. As a result of this approach, traditional file systems cannot create an abstraction layer for archival data that treats stored data as an independent data object. In other words, all data stored via a file system is tightly associated with both its application and the OS that supports it.

In the context of long-term disk archiving, this tight coupling of application and OS creates "lock-in" challenges on two fronts: first, it represents a management challenge for archiving content across multiple applications (and operating systems) in a centralized manner. Second, file systems pose a viability risk to the archive over time as they obsolesce along with applications and operating systems, thereby forcing obsolescence onto the captive archived data.

Challenge: File System Growth

As a file system grows in relation to its operating system and application, it eventually encroaches on the outer bounds of its available address space for storing data. The practical implication of hitting this boundary is a noticeably negative impact on performance. This is a very common IT concern, and it is especially well known to anyone who has ever faced a growing departmental file server. With today's dominant enterprise file systems (e.g. NTFS for Windows environments and the various Linux-based file systems), the maximum accessible limit hovers effectively around 2 terabytes per file system. Before reaching that capacity boundary, users will proactively extend their production environment into a new file system that provides a new address space onto which data can be stored.

The requirement to migrate a production environment to a new file system is typically a time-consuming and manually intensive task. In the context of disk-based archiving, this manner of file system growth management quickly becomes untenable. With archives that regularly range into the multiple terabytes in size and continue along that growth trajectory, the need to continually manage the scaling and migration of multiple file systems and their associated applications constitutes a massive challenge.

Challenge: File System Access

When a user establishes a given file system as an interface into an archival pool, they have made a commitment to begin layering data into increasingly complex hierarchies. Even when that single archiving file system is presented to multiple applications through a network mount (e.g. a NFS or CIFS interface), it still represents a unified, deep hierarchy of directory and file data. As the archive grows, the file system will have to expend increasingly more time performing deep queries into its directories to extract data. More critically, the data being stored is frozen in its relation to both its application and the other data stored around it.

This tight coupling prevents the file system from being able to easily support dynamic data views into the environment across multiple applications and operating systems. Based on our client work, Taneja Group has seen that the true business value of disk-based archiving is derived from the ability of multiple archiving applications (e.g. content management, email, voice & video recordings, medical images, proprietary applications, file data, etc…) to communicate with each other in a seamless fashion. For this reason, we are confident that the restricted access flexibility of a traditional file system approach is increasingly unacceptable to an organization's end users.

Challenge: File System Backup
File systems in an archive solution have all the management challenges already discussed and no built in mechanism for assuring content integrity and authenticity. As such, file systems can be easily corrupted. Knowing this, a common best practice is to conduct frequent backups, which further adds cost and complexity to the management burden of using file systems for archiving. With object storage approaches increasingly common, the advantages of this end-to-end data integrity and authenticity have become obvious to end users.

Object Approaches to Archiving

Looking beyond traditional file system based approaches to disk archiving, what else is available? Taneja Group knows that viable alternatives are in the market. In particular, a distributed object storage approach to disk archiving has been in use by many organizations for over half a decade. Because of its strikingly different architecture and additional use cases, the implications of object storage archiving are now clearly comprehended by the enterprise community.

We have seen that the difference in approach is exemplified by the market-defining EMC Centera archival appliance. Centera utilizes a distributed object software model known as Content Addressed Storage (CAS). CAS-based archiving differs from traditional file system-based approaches in several key respects that have had a profound impact for all deployments. Most notably for this discussion, CAS does not utilize traditional file systems, nor does it need to utilize specified storage media, nor does it require kernel level integration with host applications. Clearly, the compounding effect of these differences add up to a fundamentally different kind of archive architecture and a lower total cost of ownership. However, the most salient, driving difference resides in how CAS stores and retrieves data. In other words, what CAS does *instead* of using a hierarchical file system.

To assist with educating an organization with cutting through the complexity in evaluating potential CAS-based solutions versus traditional file systems, we have summarized the following points of differentiation brought to the table by CAS:

CAS: Flat address space
Unlike traditional file systems, CAS does not rely on a hierarchical scheme of directories and files to organize data. Rather, such solutions rely on unique hash-code identifiers (a digital fingerprint) specific to each unique content element. This content-based addressing schema that encapsulates entire files or sets of data independently from any file system enables CAS to create what Taneja Group calls “archival objects”. We define archival objects as *digital assets that have been processed by an object-based addressing technology and enhanced with metadata attributes that enable the asset to be utilized as an independent resource*. With CAS, a unit of data and its metadata are inextricably linked, and captured as a unique object stored within a flat address space. The most important results of storing archival objects in this flat address space are (1) the content authenticity of archived objects is assured and (2) the archived objects are now

abstracted and independent of their application and operating system associations. This translates into high flexibility with regard to the number and type of applications and operating systems with which CAS can be deployed.

By freeing archival storage from the constraints of hierarchical, limited capacity file systems, CAS reduces administrative complexity. Moreover, since CAS removes file system complexity and fragility, and increases the integrity of stored data objects, organizations can rely solely on replication for disaster recovery, and negate on-site archive backup. As a result of this one-two punch against management overhead, Taneja Group has observed cases where organizations can easily manage magnitudes more archived information using a CAS solution vs. tape, optical or traditional file system based storage. In one observed case it was greater than 100 times more information.

CAS: A Single Instance Store

CAS Metadata is specific to each user's use of the content, yet points to the same piece of unique content. The result can be dramatic reductions the quantity of storage required for an archive.

CAS: Metadata

By storing metadata about content use, applications can often complete given information requests by searching the storage-based metadata and never open the content objects. The result is increased application performance. More profound is the ability to do cross-application information queries without using application cycles. This is possible because (1) content and metadata stored within CAS is application, file and operating system independent, (2) metadata is searchable and (3) specific to EMC Centera CAS there is a search engine available in the repository. Easy cross-application querying provides immense benefits for day-to-day business, governance and compliance.

CAS: Application level access

Because of the unique content-based addressing approach of CAS solutions, they are able to integrate directly with application environments via APIs. Unlike file systems that have kernel level dependencies on the operating system, CAS solutions extend their archival support cleanly within the user space of a given application. There are several significant impacts of this design approach: first, it means that multiple applications can simultaneously leverage the same centralized CAS archival storage infrastructure. Second, it means that very specific archiving management attributes (e.g. aging of data, protection of data, and access to data) can be executed on a *per-application* basis. These capabilities create a "complete chain of information custody", allowing data to be completely controlled, managed and authenticated after leaving the primary application. These are capabilities not native to traditional file system archival approaches.

CAS: Media Independence

File systems and the operating systems on which they depend are designed and certified for deployment with specific disk types (e.g. SCSI, ATA,) and protocols (e.g. Fibre-channel, iSCSI). By contrast, CAS based

archiving solutions are truly media independent. Because CAS leverages an object-based model for its indexing, it remains neutral to any storage media on which it resides. The implications for a long-term online disk archival are therefore very significant: When a CAS archival solution is deployed, it can migrate to new storage media over time without disturbing the integrity of the archived objects. For long-term disk-based archiving, this represents significant risk mitigation and investment protection that is not readily achievable with traditional file system archiving solutions.

CAS: High Scalability

With traditional archive solutions, scaling into higher storage capacities over time requires a constant awareness of the status of the file system versus remaining available address space. As the file system reaches its maximum capacity, administrators must expand the entire file system "silo" (operating system, file system, application) in order to scale the archive. By contrast, CAS-based archival solutions can expand in an open fashion into extremely high capacities (multiple petabytes) due to their flat address space. In addition, because CAS solutions can abstract themselves across multiple applications and storage media, they enable very granular and dynamic online scaling to take place for both application hosts and storage capacities, each according to their immediate demands.

CAS: Self-managing

Management of the archive infrastructure constitutes a major point of differentiation between the CAS object-model approach and traditional file systems. With file system-based archives, the administrator faces a familiar range of tasks in deployment, recovery, migration, and change management of the "silo". By contrast, CAS-based approaches leverage their non-hierarchical architecture to distribute management controls across the entire archive infrastructure. For example, if a Centera disk or node fails, the archive cluster knows how to self heal without manual intervention. This distributed management structure extends to cover the deployment, scaling, recovery and protection of all the archival objects being stored by Centera. As a result of this approach, Centera removes a significant number of mundane "touches" from the disk-based archive that still exist with traditional file system based approaches. As an archive scales to higher capacities with more application associations, these self-managing qualities of CAS add up to a meaningful increase in overall environment efficiency.

Considered together, these qualities of CAS demonstrate that there are distinct advantages to creating disk-based archives outside of traditional file systems.

Taneja Group Opinion

We know very well the challenges that end users face in the deployment of disk archives. End users need to ask whether or not they desire a disk-based archive that provides high levels of scale, is readily available, can survive for long durations, and possesses minimal management requirements. For end users that satisfy those criteria, they will find traditional file system-based approaches to disk archiving inadequate.

As indicated above, the Taneja Group has observed there are many critical advantages to be gained by leveraging object-based storage in the form of CAS disk-based archiving solutions, such as EMC Centera. By stepping outside of the silo-effect created via hierarchical file systems, CAS opens up a wide new range of functionality that allows a complete reconsideration of the role archival information plays in an organization.

Since we first wrote on this subject more than 6 years ago, we have observed several things. First, we have seen these distinctions become self-evident, as more users adopt and scale CAS solutions to capacities that clearly demonstrate the unique capabilities of object storage. Second, because of CAS and EMC Centera in particular we have seen organizations change how they use archiving. When first introduced, disk-based archives replaced tape and optical solutions which had been relegated to deep archives because of their lack of information retrieval speed. These were archives an organization would use to store information that they hoped they would seldom need. However, today we see a new storage dilemma for organizations where archiving is helping. Specifically, for organizations that are being asked to store 30%. 50% and sometimes 100+% more information with flat or reduced IT budgets, orgnaizations are moving information that can be archived much more quickly to the archive. They are creating what we have already discussed as "Active Archives". These Active Archives further lower an organization's cost per megabyte to store information at the same time they are being leveraged to take large quantities of information out of the organization's backup streams. These Active Archives reduce backup costs and simplify the organization's IT infrastructure because the information no longer lives on primary storage and no longer needs to be backed up. However our observation is that these organizations only create Active Archives when they are confident in the robustness, scalability, performance and cost effectiveness of their archive platform. With thousands of customers and hundreds of PBs of product shipped since its inception, EMC Centera is the shining example of how organizations are using object-based storage to create deep archives and this new generation of archives, Active Archives.