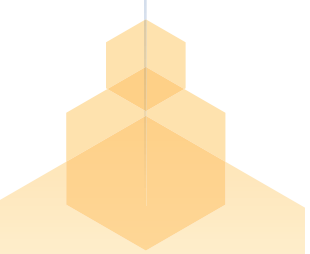


Considerations for Secondary Storage Consolidation

Dan Oister, Digi-Data



Considerations for Secondary Storage Consolidation

Every file or email your company creates starts aging the moment its completed. This helps to explain the 80/20 rule of static data to active data. This proliferation of data tends to amass at a faster than desired rate causing IT professionals worldwide to ask the same questions:

- *How do I protect such a large amount of data?*
- *How do I move the data to higher capacity systems as they become available?*
- *What provisions do I need now, and in the future to assure I can “mine” this data if required?*
- *How do I minimize the cost of capacity and power consumption over time?*

All secondary storage applications are not created equal...

Consider three of the more common secondary storage applications:

- *Disk-to-disk backup*
- *Image Storage*
- *Fixed Content / Archiving*

Disk-to-disk backup has gained rapid market momentum by improving the percentage of completed nightly backups primarily through accelerating data transfers and improving overall system reliability. In looking at the “storage signature” of this application, it requires very rapid ingestion (upload) as data is copied into the system from various backup servers. It may also require rapid data transfers out of the system for restores, (although not quite as fast, as restoring file systems takes about three times longer than the backup operation).

Image storage can mandate fast bandwidth requirements both for the storage, and for the recall of images. In some cases, it requires multiple high-speed streams, depending on how many servers or systems require video or image files.

Archival Storage usually involves a lower speed of ingestion, more like a trickle, but in other environments, it may require significant, parallel processing efforts if a critical discovery is required for legal reasons, etc. Unfortunately, it is difficult to predict the actions of various legislative bodies in terms of information retention and discovery requirements. Hypothetically, in five years companies may be required to archive video emails, which could tax the capacity and bandwidth of most systems today.

Since it is impractical to buy an expensive fibre channel disk system for each of the potential secondary storage applications you may need, it makes sense to consolidate these tier-2 applications into a single storage farm. Companies are often caught in the trap of buying multiple secondary storage systems, which amplifies maintenance and management costs.

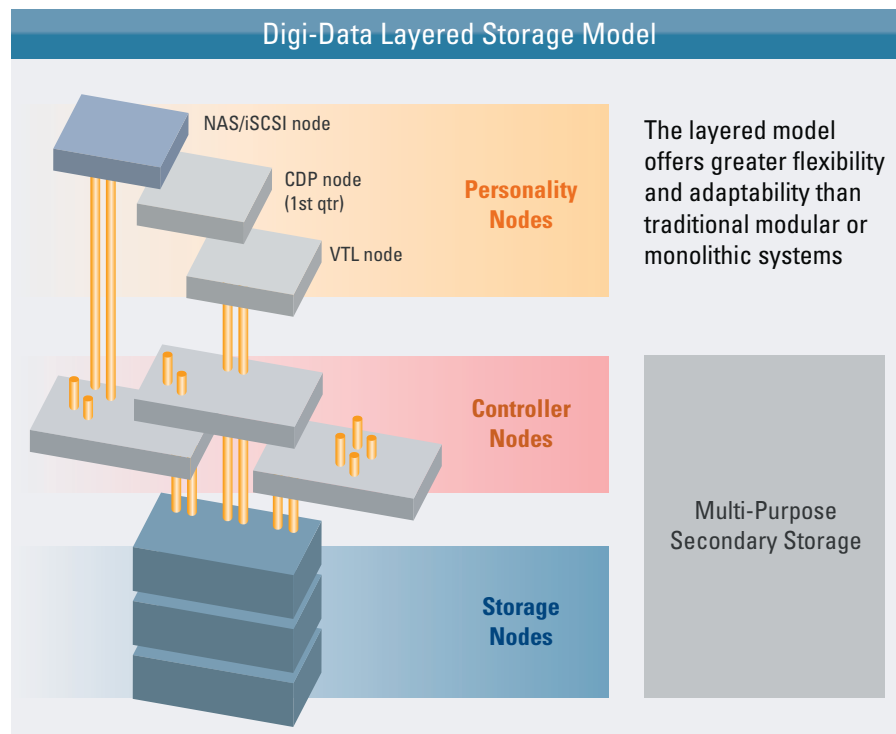
To Consolidate Secondary Storage Applications – Consider a Clustered Storage Architecture

This is where a clustered storage system makes sense to provide the proper infrastructure for current and future secondary storage requirements. The key benefit to a clustered architecture is flexibility.

Each of the components in each of the layers can be re-deployed in the event of a change in ownership, direction, or data center location. In a clustered storage model, there are three main layers that are ideally suited to cover the varied storage signatures of secondary storage apps.

At the storage layer, different types of JBOD - Just a Bunch Of Disks enclosures are utilized to maximize density, or match the I/O requirements of the application. As different profiles of applications may use the system over time, multiple types of banks of disk may be employed. The independent growth of backup, image storage, or archive could potentially drive the need for a very high disk-count. Therefore the system must be able to scale above the typical 220 - 240 disk drive limits common in today's modular storage systems. Mixed drive types, rotational speeds, and sizes must also be supported, easing the migration challenge and saving you money.

The next layer "up" involves disk controllers, which stripe the data across groups of disk to eliminate the possibility of data loss due to a disk failure. They also amplify performance by employing many



disks in each read or write operation. To get the most out of large banks of disk drives, clustered systems use multiple controllers to tune performance, access, and redundancy. If more bandwidth is required to protect data, adding a controller into the cluster can be done in a few minutes. To offer the greatest degree of granularity, these systems generally package single controllers in thin 1u enclosures.

Another key challenge in secondary storage systems is migrating up to higher capacities. Considering the disk drive of the early 90's was a 200 MB drive, and Terabyte drives are starting to be evaluated today, there needs to be an easy way to migrate within the subsystem. Make sure the storage architecture under consideration supports a migration capability that can provide full-volume background copies between volumes or systems without interruption to system operations.

Other key attributes of a good secondary storage farm include virtualization and storage pooling. To simplify growth, and provide easy connectivity from multiple servers, these systems quickly build storage pools and provide a simple mechanism to attach servers into the pool. To simplify the attachment of all types of servers, these architectures support automatic failover, without host based agents involved.

Power management features become more and more critical over time as well. Infrequently accessed information should not continue to consume energy at the same rate as critical corporate information.

Other attachments and personalities

To leverage the investment in a consolidated secondary storage system, it's beneficial to be able to present different “personalities” to various backup, NAS, or application servers. Virtual Tape, mentioned earlier, continues to be a popular way to accelerate backups and improve overall backup completion rates. Continuous Data Protection is also gaining popularity for customers with critical databases and/or e-mail servers that need a higher frequency of backups. And providing a file system interface (Network Attached Storage) is also a key gateway for large storage farms, easing the access to high availability storage.

With today's regulatory climate, most customers aren't sure what to keep, and what to throw away. Most companies are not willing to take the chance, so they keep inactive data for extended periods of time, incurring the storage and management overhead and expense.

If your firm faces similar challenges, consider doing things a little differently. Clustered Storage Systems offer a higher degree of versatility and simplified management than independent islands of storage, whilst improving performance and availability for each application connected. As regulations and rich media have their impact on managing the IT infrastructure, take a look down road before you create another island of secondary storage.