The benefits of file virtualization services (FVS)

FVS, or file/NAS virtualization, simplifies the management of proliferating, heterogeneous NAS systems.

By Ray Lucchesi
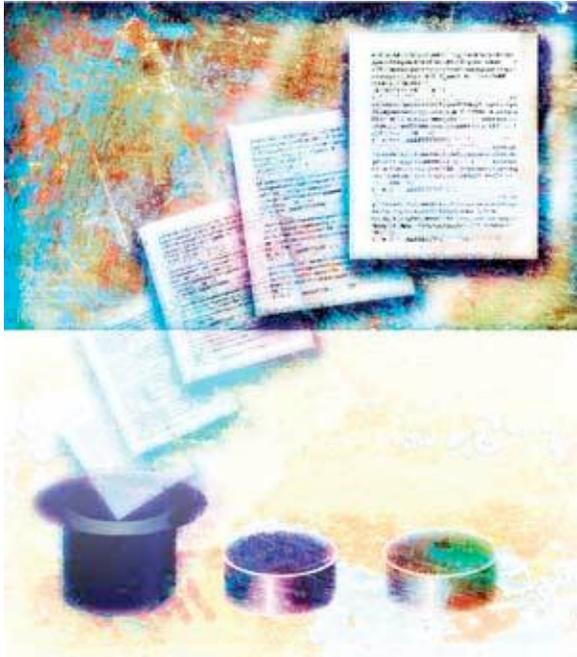
File virtualization services (FVS) vastly simplify managing heterogeneous NAS environments. Managing multiple NAS servers is a significant challenge that does not scale well. In fact, at some point (typically at about 20 to 30 NAS boxes), the additional administrator time per deployed NAS device is exponential rather than incremental. However, NAS appliances continue to be added because companies are keeping information longer and NAS is easy to deploy.

FVS products can potentially make managing 30 NAS filers as easy as managing one filer. It's not quite that simple, but by aggregating multiple NAS boxes into one management environment, FVS can make NAS management much easier.

FVS equates to transparent (no user impact) migration between heterogeneous NAS boxes. An administrator can move a directory or even a single file from one NAS file system to another without any user interruption.

Silverton Consulting
Strategy, Storage & Systems
P.O. Box 492|Broomfield, CO 80038-0492|
720-221-7270|info@silvertonconsulting.com

Moreover, administrators can refresh NAS technology painlessly, consolidate many small filers quickly, and balance capacity or performance across heterogeneous NAS devices effortlessly-all without interrupting the user base.

To provide transparent migration, FVS also supplies advanced namespace capabilities by aggregating NAS namespaces. The two most prominent advanced namespace capabilities are

- A federated namespace, which concatenates all file system namespaces and directories together into one contiguous namespace; and
- A shared namespace, which places all file system namespaces and directories under one mount/share point.

Using either namespace technology, albeit via different routes, FVS can seamlessly find and migrate all files within an FVS domain.

Centralized management, capacity, and performance reporting for NAS filers under FVS control is another key benefit. For example, such reporting can list NAS boxes that are reaching capacity or are suffering performance degradation. It can also show which filers are empty or idle.

Yet another important feature of FVS is heterogeneous remote replication. This capability enables one vendor's NAS file data to be replicated via a WAN to another vendor's NAS product.

Finally, using the policy automation feature of FVS together with all the features described above, FVS can provide rudimentary information lifecycle management (ILM) functionality for file data. Consequently, FVS can automatically scan file metadata, determine which files have not been accessed recently, and transparently migrate them to higher capacity, lower-cost storage. Conversely, frequently accessed files can be automatically migrated to higher-performance (and more-costly) storage systems.

How is it done?
File virtualization technology is implemented in one of three ways: on the host, in a stand-alone appliance, or in the network.
Host implementations depend on operating system specific services, such as Microsoft's Windows DFS, to provide federated name-spaces and user redirection after migration. Host-based FVS products have some limitations, depending on the operating systems in use, when migrating open files. When host FVS products fail, new file migrations are halted but previously migrated data remains unaffected by the outage. However, host-based FVS products can be installed quickly and integrated easily with the operating system.

Stand-alone appliance implementations depend upon being directly and

constantly in the file I/O data path to provide transparent migration through a shared namespace. The issues with being constantly in the data path include 1) a larger FVS appliance may be required to add subsequent NAS filers; 2) any failure of the appliance could be catastrophic, so high-availability appliances are required; and 3) some advanced features of the underlying NAS filers are superseded and thus need to be duplicated by the appliance. However, in an appliance-based FVS approach administrators can perform almost any desired level of virtualization because all file I/O flows first through the appliance then to the NAS filers. The FVS appliance is uppermost in the hierarchy and has no dependence on the host operating systems, NAS boxes, or network switches.

FVS remote replication
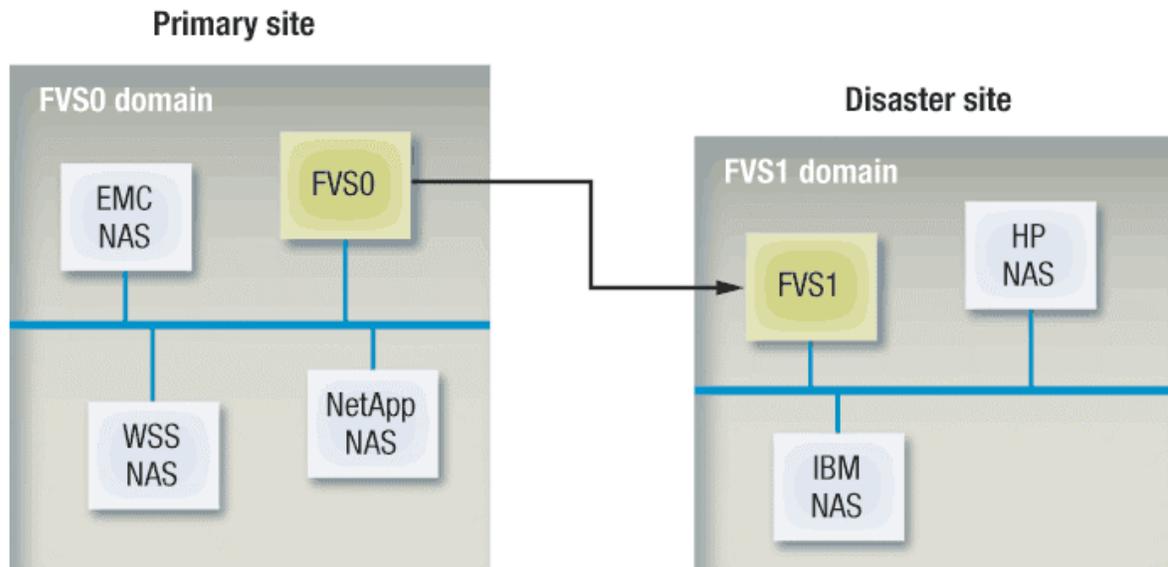
FVS remote replication provides cross-vendor duplication of NAS data under an FVS domain from a primary site to a disaster-recovery site over a WAN.

Network-based FVS implementations depend upon close interaction and interfacing with network switches and NAS products to provide federated namespaces and user redirection after migration. Network-based products must move in and out of the data path to provide file virtualization on an as-needed basis. When network-based FVS products fail, all active file migrations and other advanced file management capabilities are halted, but previously migrated data remains unaffected by the outage. While the network-based approach can be particularly useful, it relies on switch and NAS vendors' APIs. However, given API availability and compatibility, network-based FVS implementations can optimize file virtualization much like the appliance

implementations yet not suffer the downsides of the appliance approach.

Why FVS?

Transparent migration is the key feature to simplifying NAS management. To understand the advantages of FVS in this context, consider what an overburdened administrator must do to move a directory without FVS:

1. Identify a directory to move.
2. Find space on another filer to hold the directory to be moved.
3. Identify all the users that access the directory.
4. Schedule a convenient time across all users to take the directory offline.
5. Take all user access to that directory offline.
6. Move the directory to the new filer.
7. Update all the users' mount/share points to the new file system.
8. Re-enable user access to the directory.

Most of these steps are short, but the scheduling involved in Step 4 may be impossible for some directories and extremely difficult for others. Duration of the whole process is dependent on the number of users involved, but it's almost guaranteed to take days to complete.

In contrast, directory migration with FVS only requires an administrator to:

1. Identify a directory to move.
2. Find space on another filer to hold the directory to be moved.
3. Move the directory to the new filer.

Note that this does not involve user scheduling, user downtime, or updating users' mount/share points. Depending on the amount of file data being moved, this process could only take minutes. Thus, once FVS is in place administrators have a powerful tool in managing many NAS filers.

What else can FVS do?

Once effortless migration is possible, it becomes much easier to accomplish NAS capacity balancing. For example, when one NAS filer runs out of capacity,

- Finding space consuming, inactive directories on this filer only requires viewing a report screen;
- Finding available space on another NAS box only requires viewing another report screen; and
- Moving the directory is just a couple of mouse clicks.

In a matter of minutes, not days, administrators can free up valuable capacity on a critical resource and deploy it on another NAS box with available space. Furthermore, this same approach applies to performance balancing (i.e., find a highly active directory on an over-burdened NAS filer, find a relatively idle NAS box, and move the data). Once again, the entire process can be accomplished

in minutes, as opposed to days.

With FVS, NAS technology refreshes can also be done much easier. FVS allows an administrator to easily install new NAS products and quickly deploy active data to them without user impact. The same holds true for NAS consolidation. Painless capacity and performance balancing, technology refreshes, and NAS consolidation are a direct result of the ease and transparency of file migration under FVS.

FVS products help with many of the balancing, refreshing, and consolidating activities by providing policy automation. For example, policies can be created that cause stale file data to be moved to lower-cost, higher-capacity NAS filers, thus eliminating much of the NAS capacity-balancing problem. Administrators can do the same for file extensions such as .MP3, .JPG, or other media files. For data consolidation or technology refreshes, an automated policy could move all file data on one NAS box to another box. Similarly, policy automation can be used to move and secure file data into special-purpose archives for permanent storage based on directory path name combined with last-access timestamps. All of this automated policy-driven activity is based on standard NAS file system metadata, such as file name, extension, or last- access timestamps.

Equally important, FVS provides heterogeneous remote replication, enabling use of multi-vendor NAS environments for highly reliable disaster-recovery solutions. Furthermore, this FVS feature eliminates vendor "lock-in" because the FVS remote replication capability is not dependent on vendor-specific features or matching NAS boxes. An FVS product may be required in both the primary and remote site, but redeployment of diverse vendors' NAS boxes is certainly made easier.

In contrast to the above advantages, the advanced namespace capabilities of FVS have received less universal support. While most IT administrators appreciate the advantages of a single mount/share point, users are less enamored. All mount/share points must be changed, which is often perceived by users as a loss of departmental control. As for federated namespaces, users can retain original user mount/share points, but must rely on automated services to correctly redirect access when data is migrated. Both approaches work well and can provide advanced namespace services necessary to help manage a NAS environment.

Where to deploy FVS
Most data centers with more than 20 NAS boxes would benefit significantly from FVS, but even an environment with a few heterogeneous NAS appliances may benefit from FVS. Some NAS vendors provide transparent migration among multiple versions of their products, but seldom offer these features across

other NAS vendors' products, which is where FVS enters the picture.

Representative FVS vendors

Attune Systems	EMC (Rainfinity)
Brocade (StorageX)	F5 Networks (Acopia)
Cisco (NeoPath)	

Most often the first customer reaction after he or she implements FVS is disbelief when he or she sees how quickly inactive data accumulates on a network. It's not unusual to have 80% of file data un-accessed for more than 90 days. With inexpensive, abundant file storage, data is often kept long after its useful life. FVS can facilitate the management of this additional data.

Some users are surprised that they can implement FVS on a trial basis for one file system or even a few directories. It's easy to set up these FVS products to manage only a few directories or file systems and expand later. Once the technology has proven itself, it can be easily expanded to incorporate more of the NAS environment.

IT managers are typically surprised at the ease and transparency of the directory migration process. Too often, a workable, yet unchanging and oftentimes cumbersome process is used to migrate directories, which affects the operating efficiency of the entire data center. Once having seen the effectiveness of FVS, most managers are eager to adopt FVS products to improve overall file data management.

Finally, when data centers first start to improve file management they often buy less NAS storage-temporarily at least. In reality, this is only a minor delay in their NAS acquisitions. Once capacity utilization is balanced, any data growth ultimately requires additional NAS devices. By making file data easier to manage, eventually companies buy bigger, more-capable NAS systems because they can now easily see a way to painlessly implement consolidation.

What's next?
Recently, storage vendors and standards organizations have been working to define the file area network (FAN), and FVS can be a critical component of a FAN.

Market research firms do not formally track the FVS market, but one clear indicator of its size and growth potential is the spurt of acquisition activity in this space. Almost every start-up in this space has been acquired over the last 24 months (see vendor box on p. 26)- an indication that rapid market growth is imminent.

Moreover, it's clear that the need for FVS products is acute and growing. FVS products will continue to supply increased functionality and automation to make managing NAS systems and file data more effective. In addition, in the near future we may see a merging of data-classification services and file virtualization services into one solution. Because data-classification services address file management by examining file content, the coupling of the two technologies could make the promise of ILM a reality-at least for file data. For today, however, FVS provides a clear advantage for data centers managing multiple NAS products, and may provide the ILM services needed for current and future file data requirements. ❏

© *InfoStor* October, 2007

About the author:
Ray Lucchesi *is president of Silverton Consulting, a Storage, Strategy & Systems consulting services company, based in the USA offering products and services to the data storage community*.
mailto:info@silvertonconsulting.com
http://www.silvertonconsulting.com

Silverton Consulting
Strategy, Storage & Systems
P.O. Box 492|Broomfield, CO 80038-0492|
720-221-7270|info@silvertonconsulting.com

© 2008 InfoStor
All Rights Reserved

Page 7 of 7