

Commentary

June 14, 2006

Find It! Move It! The Power of EMC's VisualSRM and DiskXtender Integration

The problems already facing overloaded and still growing file systems — such as overworked backup/restore processes, increased risk to application availability and more-frequent-than-desired purchases of expensive primary storage — are only going to get worse. Much of the overloaded file data is read-only data that can be offloaded to an active file archive to help alleviate those burdens. The difficulty has been how to identify the files to be moved and then move them to the active archive. Now EMC makes those tasks much easier by integrating its VisualSRM and DiskXtender for NAS products to perform the file targeting and file migrations in concert.

Trouble Right Here in River City With Overburdened File Systems

The integration of EMC's VisualSRM and DiskXtender for NAS products helps organizations attack the ever-growing problem of overburdened file systems. The never-ending growth of storage as well as enterprises' need to meet more stringent legal and regulatory requirements continues to put increased administrative and cost burdens on IT and the firm as a whole. These pressures result in four challenges related to the efficient and effective management of data and the storage upon which that data resides:

- *Improving the backup/restore process* — both full backups and restores take longer with more data; the result is threats to availability (e.g., backups encroaching upon online availability and re-

stores overrunning allowed unplanned-downtime service levels) as well as increased cost for resources, such as backup media

- *Ensuring better application availability* — data protection processes need to be improved, and the risk of availability-impacting out-of-space conditions needs to be dramatically reduced
- *Making better use of storage assets* — using one or more tiers of less expensive secondary storage reduces the strain on the IT capital budget
- *Setting the stage to make it easier to manage legal and regulatory demands* — IT needs to put the data that has to be managed for compliance in a retention-managed pool of storage

The single answer to all these challenges is archiving — and, more specifically, file archiving. The greatest growth in storage comes from files: not only from tradi-

tional business documents, such as word processing documents and presentations, but also from digital media assets, medical tests (such as X-rays and MRIs), check images, and CAD/CAM drawings; and on and on.

Fast Relief through Active File Archiving

Read-only data files (especially as read access requirements may diminish over time) are candidates to be migrated to an active archive. An active archive is a pool of read-only data that is still online, accessible transparently by end users and applications. Thus, the files are still production data (as they can still serve useful business purposes even though they are not being modified in the “live” production pool).

The “live” production pool of active data becomes much smaller — and perhaps dramatically so. One large EMC multi-national customer attested to the fact that “70% is always fixed” and a telco customer said that “less than 1% changes after one week.” Simply reducing the size of the active changeable pool while creating an active archive pool solves each of the four problems facing overburdened file systems:

- Backups and restores take less time, are more reliable, and require fewer resources
- Freeing up disk space on primary storage lessens the risk of unavailability-causing out-of-space conditions
- Using ATA disk for secondary storage in a tiered storage approach simplified by an ILM architecture reduces the continuous need to purchase additional high

priced Fibre Channel primary storage

- Compliance constraints can only be imposed on a read-only content pool of storage (which an active archive is) since, by definition, the files are not allowed to be changed.

If active archiving of files brings such fast relief, why isn’t everyone doing it?

EMC Combo Enables Effective Active File Archiving

Once IT organizations understand that they need to have a systematic and comprehensive approach for archiving files, they typically run into the problem of what they should archive, and/or the problem of how to get a file migration tool to migrate only the targeted files. EMC solves both problems through a combination of VisualSRM and DiskXtender for NAS (Figure 1).

VisualSRM is a robust open host-based storage resource management (SRM) software tool that provides wide operating systems and applications support (e.g. Windows, Linux, etc.) With VisualSRM IT administrators can reach out and discover what files are controlled by what NAS systems and select those files that meet criteria based on file metadata (such as date last accessed and file size) as candidates for migration.

An IT administrator at a VisualSRM workstation can go through the VisualSRM Console process to interface with DiskXtender for NAS to create a policy “job” that specifies the primary storage source and secondary storage destination for the files that are to be migrated. VisualSRM, either on a scheduled basis or interactively, executes a policy called an Intelligent Action that enables DiskXtender for NAS to migrate the

candidate files to the secondary storage active archive.

In an “all in the EMC family” approach, these files are moved from a Celerra-managed primary storage NAS array to a Celerra-managed secondary storage array (although the destination can be tape, optical, Centera, or a plain-vanilla disk array).

Looking Under the Hood

Both products — VisualSRM and DiskXtender for NAS — have to be installed, but the integration is managed through the VisualSRM Console. The DiskXtender user interface maintains its character within the context of the VisualSRM Console, which means that the administrator has to know and understand how to use both products. Typically, one of the two products may already be in use, so users would need only to learn how the other works. As both of these products are straightforward to use, the learning curve should not be burdensome.

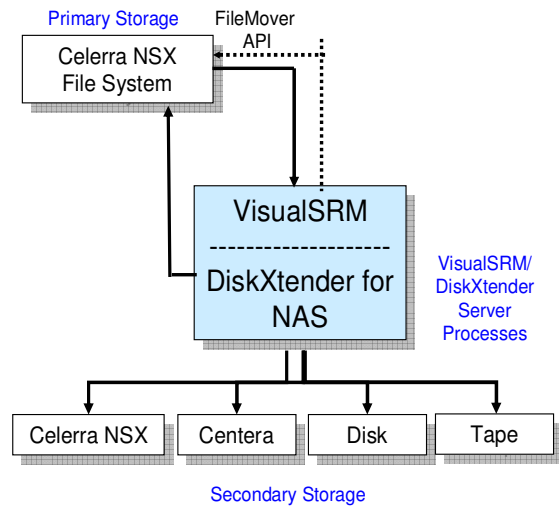
VisualSRM establishes the mechanisms for collecting file data, for analyzing that data, and then for setting policies that, when invoked, take the action specified in the policy. Those mechanisms include the use of agents, metadata, and what are called Intelligent Actions.

Agents

VisualSRM deploys agents to systems. Agents are "pushed" to systems. The Push Install process is non-intrusive, i.e. the IT administrator can install, upgrade, and de-install agents without rebooting the remote system. Once an agent is installed on a host, that system is considered to be a "managed

resource." An administrator can then use the VisualSRM Console to manage the agents.

Figure 1: Architectural Overview of VisualSRM and DiskXtender



Source: Mesabi Group, EMC June 2006

An IT administrator will set up scanning policies that will receive data from the agents on how the storage controlled by the systems is being used; these policies can be run interactively or automatically on a pre-determined schedule.

Although agents collect data for reporting on storage consumption, agents can also be directed to take policy actions, such as moving and deleting data.

Because of the ability of agents to take actions, the VisualSRM integration with DiskXtender for NAS is a proactive management tool that is able to carry out policy-directed actions anticipating and forestalling problems in addition to being a planning tool.

Metadata

Once agents are installed, policies to scan the files on the remote system are set up. A file scan can be established at the server, volume, mount-point, folder, or path level. A scan will gather all the necessary file metadata (e.g. data about data). That file metadata includes filename, path, file owner, create date, modified date, last backup date, last accessed date, size allocated, and size used plus system attributes.

The file metadata is stored in a repository. An IT administrator can assess and analyze the storage environment using VisualSRM's extensive reporting capabilities to interrogate the repository. For example, files that are not considered business-related (for example, MP3 files and vacation pictures), files that are at risk because they have not been backed up recently, and files that have not been accessed for some period of time can all be identified. Identified means that they can be classified not only individually, but as part of a larger group (say, all files that have not been accessed for the last sixty days). Files that have been identified are candidates for migration.

Intelligent Actions

Intelligent Actions are policies that can move, delete, copy, stage, and compress files, or run a script. Files are selected as the target of Intelligent Actions through the use of filters on the metadata — such as file type, last access date, or create date — that the agents have collected in a scan. Intelligent Actions can be run interactively directly from the VisualSRM console, or can be scheduled to run at

one or more designated times automatically.

Putting It All Together to Make the Right Move

As noted above, through the VisualSRM Console, an IT administrator identifies (i.e., finds) the files that need to be migrated. Then, an Intelligent Action has to be prepared. To do this, script information has to be entered and filters have to be applied. Multiple filters can be used to get the proper level of granularity — for example, all the files with a specified file extension that have not been accessed in the last 30 days.

An IT administrator then uses the DiskXtender for NAS policy wizard to create a policy “job” that describes the migration process —the source, the destination, and the readback method.

A source is the NAS location of the files that should be migrated. The destination is a directory on the secondary storage where the migrated files will be moved. The readback method gives the access method for getting migrated data back. When the migration takes place, a small file stub is created on the primary storage Celerra in place of the migrated file. When an end user or application wants to access a file, the stub knows where the file now is. The typical readback access method is pass through, which means that the file is read from secondary storage, but the file remains on secondary storage.

Through VisualSRM, an administrator can “kick off” the DiskXtender for NAS policy job that actually carries out the migration.

When a Threshold Is Exceeded

VisualSRM can set thresholds for acceptable storage consumption in order to prevent out-of-control situations from arising. Those thresholds can be set on percentage limits or on the amount of space consumed. When a threshold has been exceeded, a threshold policy can trigger other policies, such as automatically starting a DiskXtender policy via an intelligent action and a task list. Among the tasks in the task list could be the job of migrating files from the production file systems to secondary storage to relieve the pressure on primary storage. That migration could be based upon criteria, such as file size, file type, last accessed, etc. Multiple criteria can be applied through filters.

What Does the Future Hold

The current integration of VisualSRM and DiskXtender is NAS-only and tightly coupled with EMC products such as Celerra EMC seems likely to extend the functionality to DiskXtender for Windows as well as other EMC information management products in the future, but it had to start somewhere, and integration within the family is always easier as a starting point.

Conclusion

The concepts of ILM and tiered storage have always sounded good,

but putting them in practice has always been the challenge. “ILM made simple” is what the integration of VisualSRM and DiskXtender for NAS delivers for NAS environments.

Now, read-only files can be easily migrated from a tier 1 primary storage pool of active changeable data to a tier 2 secondary storage active archive pool. Freeing up space on primary storage speeds up both backups and restores while at the same time improving application availability. Using cost-effective disks for the active archives puts less strain upon the organization’s wallet. And the data is now in a position to be retention-managed effectively — notably for compliance requirements.

This integration solution meets the twin challenges of classification and migration — challenges that must be met if users are to build an effective active archive. VisualSRM enables the identification of the migration-candidate files and DiskXtender then does the “heavy lifting” of actually migrating the files. If an organization is thinking about building an active archive in a NAS environment, strong consideration should be given to VisualSRM and DiskXtender for NAS integration.

David Hill

Analyst Name: David Hill
Topic Area: Data Management

Mesabi Group LLC
26 Country Lane
Westwood, MA 02090
www.mesabigroup.com

Mesabi Group LLC is an affiliate of Valley View Ventures that aims to provide thought leadership and sound advice to both vendors and users of information technology. .

Phone: (781) 326-0038
email the author: davidhill@mesabigroup.com

The information contained in this publication has been obtained from sources Mesabi Group LLC believes to be reliable, but is not warranted by Mesabi Group LLC. Commentary opinions reflect the analyst’s judgment at the time and are subject to change without notice. Unless otherwise noted, the entire contents of this publication are copyrighted by Mesabi Group LLC, and may not be reproduced, stored in a retrieval system, or transmitted in any form or by any means without prior written consent by Mesabi Group LLC