Why file virtualization? It's easy

Up-and-coming technology melds multiple storage devices into one, boosting efficiency and data availability

By John Moore
Published on October 29, 2007

Some co-location data protection strategies defy the reach of small and midsize agencies, but New York City's consumer watchdog organization found a way to get the job done using a new technique called file virtualization.

File virtualization provides a single, unchanging point through which users access their files, though in reality those files may be stored on multiple storage systems in different locations and moved around by administrators as business and infrastructure needs change.

"What I was interested in doing was to try to figure out some way to provide network access and access to users' files in the event that anything happened to the [department's] main location," said Matthew Miller, local-area network administrator at New York's Department of Consumer Affairs.

The department's solution involves placing single network-attached storage devices at two locations, replicating files between the storage boxes and using file virtualization to give some 360 employees in five offices seamless access to files as if they were housed on just one system.

Rick Villars, vice president of storage systems at IDC, said file virtualization is an increasingly popular option for managing multiple NAS devices.

"The challenge of managing all the systems and moving information between the systems — you can't do that on a device-by-device level," Villars said. "File virtualization is one technology that can make lots of individual file systems look like — and be manageable as — one aggregate environment."

Transparent access

Last year, the Department of Consumer Affairs installed an EMC Celerra NS500 in its main office in lower Manhattan as part of a Windows Server 2003 network upgrade. The department then deployed a second NAS device, a Network Appliance FAS 270, for its Queens location as a hedge against data loss.

The Manhattan location at 42 Broadway has experienced two disasters within the past decade: the World Trade Center attacks a few blocks away in 2001 and the Northeast blackout, which shut down the city in 2003. The latter occurrence "pretty much took us off-line," Miller said.

The department's disaster response approach, the last elements of which went into place this past summer, involves both replication and virtualization. The organization's user, division and general shared directories all save to the Manhattan NAS unit and are replicated to the device in Queens. The two facilities are linked via fiber-optic cable.

With replication in place, the department deployed Microsoft's Distributed

File System and NetApp's Virtual File Manager to handle file virtualization duties. Virtual File Manager, which is NetApp's rebranding of Brocade Communications Systems' StorageX software, aggregates the distributed file data across the mixed-vendor environment and pools the multiple file systems into one logical file system.

Like many file virtualization solutions, NetApp's Virtual File Manager resides on a separate server rather than the storage device.

With virtualization, "the location of files is transparent to the users," Miller said. "If there's a problem with 42 Broadway, users are automatically redirected to Queens, and everything proceeds as if nothing had happened."

Miller said the system has been tested, adding that he has been pleased with its functioning thus far. The EMC NAS machine has 6T of capacity, and the NetApp filer is capable of holding a comparable amount of data.

The roster of data types protected by the virtualization setup includes Microsoft Word and Excel files along with a couple of small databases, Miller said. In addition, the department licenses businesses in 55 categories — from bingo hall operators to tour guides — and maintains images of licensees' photo identifiction cards in the replicated NAS environments.

The cost of the solution came in much lower than a storage-area network approach that was briefly considered. "The SAN…was just way too expensive for what we had to do," Miller said. The department's EMC NAS device was priced at around $100,000, and the NetApp device and associated software cost about half that amount, Miller said.

In contrast, high-end SAN replication can cost $100,000 or more for only the software.

Miller said the department's file virtualization won't protect against every disaster scenario. But he described the solution as a relatively low-cost way of assuring information access in the event of many types of disruption.