

# Backup strategies for networked storage

A quick review of NAS, SAN, iSCSI, serverless backup, NDMP, snapshots, and DAFS in the context of backup.

■ BY SOUBIR ACHARYA  
AND SUSAN G. FRIEDMAN

Recently there has been a plethora of articles on SANs, NAS, and NAS vs. SAN for network storage—almost an overload of opinions and controversy. The prevailing wisdom is that eventually they will not be seen as separate and dichotomous solutions. Instead, they will be viewed as complementary aspects to an overall network storage strategy.

However, this middle group is flanked by extremes: On one side are those who view storage area networks (SANs) as too costly, complex, and time-intensive to maintain, and on the other side are those who see network-attached storage (NAS) as a low-end product that does not effectively meet the needs of many storage-intensive applications.

However, a middle ground is coming, and there are many new technologies and tools in the pipeline that will aid in meeting the considerable challenge of backing up data on heterogeneous networks, whether SAN, NAS, or a hybrid of both. Two major areas of advancements are in interconnect technologies and evolving protocols from the Storage Networking Industry Association (SNIA).

## Network-attached storage

NAS optimizes the Ethernet, or Gigabit Ethernet, IP networking technology of a LAN.

A NAS device (or server) is a box with a minimal kernel operating system, designed to facilitate data movement. It is platform-agnostic, and any user using virtually any operating system

can send data to or receive data from a NAS device using file-oriented protocols such as NFS, CIFS, or HTTP over TCP/IP. Clients access storage on a NAS server on a file-by-file basis, as opposed to accessing storage block-by-block over Fibre Channel or iSCSI. Attached to the LAN, a NAS server has an any-to-any connection, enabling shared and remote storage over a heterogeneous network.

The Network Data Management Protocol (NDMP) is used to send backup instructions over TCP/IP to a NAS device. Backup becomes a much easier task—and more standardized—with the NDMP protocol. NDMP will be discussed in more detail later in this article.

## Storage area network

A SAN is a network built specifically to handle storage and backup traffic. It is separate and distinct from the LAN and takes storage traffic off the regular network. For all practical purposes, at this point all SANs use Fibre Channel interconnects, although technologies such as iSCSI will become more viable next year.

In a Fibre Channel SAN, every server and storage device is linked directly to the SAN, so it is a many-to-many connection on a dedicated storage network. Any of the servers on the SAN can access any of the storage devices, allowing sharing of expensive tape libraries and other storage hardware between multiple servers.

Because it uses Fibre Channel, hosts and applications see storage devices attached to the SAN as locally attached storage. The Fibre Channel architecture uses block I/O protocols, and so it is higher speed than TCP/IP-based protocols. Because of its network characteristics, a SAN can support multiple protocols and operating systems and can be managed much the same way as a heterogeneous network is managed. It



is also relatively scalable, after overcoming the first large-scale building step.

## iSCSI

iSCSI is an emerging storage networking technology that is currently in the standards approval process in the Internet Engineering Task Force (IETF). Among other potential advantages, iSCSI is tailor-made for heterogeneous networks.

iSCSI is a TCP/IP-based storage networking specification that enables any storage connected to an IP network to be backed up from any point on that network. By carrying SCSI commands over an IP network, iSCSI facilitates long-distance storage deployment, management, and data transfer over intranets. This capability makes deployment flexible. Storage and servers can be added at a convenient location, not in locations dictated by proximity, as with Fibre Channel. Since it's sent over TCP/IP, iSCSI can sustain multiple simultaneous channels, allowing many-to-many conversations between hosts and storage.

Because IP (Ethernet and Gigabit Ethernet) networks are pervasive in all areas of networking—the Internet, WANs, and LANs—they can make use of iSCSI. Storage traffic can be carried over the existing intranet/Internet without any need to upgrade. There is already considerable knowledge and investment in TCP/IP networking, making it very likely that end users will adopt a protocol that makes use of the existing infrastructure. In this sense, iSCSI is a threat to Fibre Channel SANs, which require investing in new and separate plumbing.

iSCSI maps a serial SCSI protocol on top of TCP to enable a host computer to see and move data in and out of remote IP-attached storage devices as if they were local devices. Disk or tape storage plugged directly into a LAN is then visible to all computers in the LAN, much like any regular computer or server would be. This is ideal for backup considerations because it means that, subject to security restrictions, any client can back up to any device, regardless of its location.

## iSCSI or NAS?

Both iSCSI and NAS provide the flexibility of multiple, shared storage devices over existing TCP/IP networks, but each has advantages and disadvantages. A NAS device is essentially an appliance with a power cord and network cable that comes plug-and-play ready. Extensive hardware configuration is not required.

iSCSI will similarly come with plug-and-play ease. However, creating an iSCSI disk device requires a lower level of investment for vendors, compared to creating a NAS appliance, which requires an operating system, file system, and other components. All that is needed for an iSCSI disk device to work is a network card, and associated logic, that can act as an iSCSI target, extracting SCSI commands from TCP packets.

An iSCSI disk subsystem is a better solution than directly (locally) attached storage because there is less disruption when adding additional storage. Storage-on-demand becomes easy when there is no downtime: Just add another device to the network. Conversely, if storage is directly attached, downtime is almost guaranteed when additional units are connected.

iSCSI disk subsystems will deliver data to hosts faster than a NAS arrangement, because the protocol is block I/O-oriented and does not have the overhead of a separate network protocol to talk to a client. In the early stages, we will also see iSCSI gateways that are plugged into a gigabit network at one end, with SCSI or Fibre Channel storage behind it. This could provide an easy migration path, leading to early adoption.

iSCSI is still in its infancy, however, and the existing network infrastructure and TCP/IP protocol are not well-suited for storage, which requires high bandwidth and relatively error-free transmission. Traditional SCSI implementations have assumed a dedicated link to storage with very low error rates and low-latency burst transfers. TCP/IP has to be adapted to meet these requirements. iSCSI implementations must employ a gigabit network as well as modified

TCP/IP stack software for optimal performance. Some kind of storage virtualization software will also be needed as an enabling technology to orchestrate all the pieces involved in this storage pooling.

In addition, iSCSI's newness as a protocol introduces interoperability issues. Competing vendors are developing different implementations that may not work together. Much the same problem is found with Fibre Channel SANs. Emerging industry standards generated from the IETF will hopefully address these issues, rendering them moot in the near future.

Quality of Service requirements may need to be enforced to achieve a workable storage network. iSCSI has requirements such as the ability for Point A and Point B to negotiate dedicated bandwidth for a period of time, which networks do not normally provide. Network security can also be an issue. However, the iSCSI protocol negates that concern because it incorporates peer-to-peer authentication as well as encryption, if necessary. An iSCSI login phase, during which session characteristics are negotiated, precedes actual data transfer.

## Serverless backup

Serverless backup is an advance that will further facilitate data management on heterogeneous networks. More accurately called “non-intrusive image backup,” it works by backing up volume-by-volume rather than file-by-file, yet allowing file-level restore. Equally important, while it does this it creates only a small footprint on the application server. Therefore, it frees up network and server resources for other uses. There are several ways of accomplishing so-called serverless backup, including third-party copy or SCSI extended-copy command (X-copy) and using an alternate node or host for backing up the application server.

Serverless backup is dependent on three things:

- A single, consistent point-in-time image (e.g., a snapshot);
- Intelligent SAN device hardware (e.g.,

a storage router) that can move data independently of the host server. The intelligent hardware takes over the function previously performed by the server, sending instructions to the data; and

- Ability to restore a single file from an image backup. This basically involves mapping file-allocation information down to the physical block level.

Serverless backup with iSCSI is easier than with other technologies, because if one client can see the storage, another client can see the same storage and act as the backup management station. In contrast, with NAS the client can only talk directly to the NAS server, which impairs performance. It also ties the backup method to the specific NAS software, limiting options.

iSCSI targets may interact to support X-copy, or intelligent mediation may be required. It is also possible that a complementary protocol to iSCSI will evolve. This would enable hosts to transfer bulk data directly from disks to archival media without needing explicit X-copy support. Another possibility is that an NDMP server engine would be embedded in hardware that would enable backup-and-restore processes to be initiated and controlled by remote nodes on the network.

Backup vendors will have to discover new ways of moving bulk data from IP storage devices to IP-enabled tape or regular SCSI tape in a manner that doesn't affect server or network bandwidth significantly. The challenges are in achieving optimal performance and a high degree of reliability, since iSCSI would introduce its own issues of error-handling and time-outs on heterogeneous networks.

## NAS backup with NDMP

Currently, there is a task force working on version 4+ of NDMP, which will be the next-generation protocol for backup of storage on heterogeneous networks. Some of the anticipated developments include the creation of extensions to communicate to vendor-specific functions, the ability to

provide integrated application-specific solutions, support for snapshot management, and support for specific library vendors' tape servers implemented in hardware.

NDMP is becoming an IETF standard, which will lead to more acceptance in the marketplace. Vendors are also creating value-added features with vendor-specific extensions. The extensions are outside the core protocol and, because the core remains standard, interoperability is maintained.

## Snapshots

One important extension is snapshot support. With the appropriate extension, the backup application can coordinate and integrate the snapshot process on a NAS appliance. The backup vendor can control when to suspend and when to resume the application. To create a snapshot, it is necessary to freeze the application by bringing it to a consistent state, so that the files are also in a frozen, consistent state for their location to be recorded. (Snapshots are basically the ability to take a "picture" and record where the data is located. The data can be restored from this information in the case of loss or corruption.)

Snapshots apply to both NAS and more-general image backups. One technique used to create snapshots is Copy On Write (COW), which involves tracking changes to the snapshot source file system. The original data for blocks with changes are copied off to a cache area before the data is overwritten, enabling the snapshot to remain frozen in time.

Snapshots take much less time to complete than backups, so the downtime entailed is less than what an actual backup would cause. For example, a 500GB Exchange database can take up to three hours to back up and is not available for transactions during this time. In comparison, a snapshot would cause only *minutes* of downtime. Some applications already provide a "QUIESCE" command, which freezes all I/O while the snapshot is taken. Therefore, the application does not need to be shut down at all, and there is zero downtime.

Snapshot management also implements near-line storage; since a snapshot takes up much less space than a backup, the data can go to disk, not only to tape. For disaster recovery, the snapshots generated over time (t1, t2, t3) can be archived and saved to tape.

This is a much more application-integrated approach because the vendor-created interface will enable users to create, delete, and manage snapshots on the server as well as schedule when the snapshots will occur. The interface supports browse and restore, so data on drive D (a snapshot of drive C at a moment in time) can be restored to drive C on a file-by-file basis.

NDMP tape libraries are integrated as part of the hardware. Though still connected by SCSI or Fibre Channel, they are virtualized as software entities. (A tape library now looks like a server software endpoint whereas, before, it was only a hardware endpoint.) Tape libraries are also available with network attachments so they can be attached to a LAN or directly to a network.

This configuration allows any node with NDMP to back up to any tape or tape library without having to use SCSI, so it can be shared. This lowers the total cost of ownership dramatically and also allows for a remote library—which is difficult with Fibre Channel SANs without complex and expensive configurations.

## DAFS

The Direct Access File System (DAFS) grew out of NFS V4 and enables NAS clients faster access to NAS disks. With DAFS, a NAS server exposes a local file system to a remote client. Backup vendors may use the highly optimized DAFS client APIs under certain circumstances to archive for a fast, less resource-intensive backup path.

DAFS turns the existing models upside down, since the backup burden is transferred to the client from the server. This is a more efficient process, so at the same time it subverts the current methods, it opens up more options for backup.

## Backup with NDMP and SANs

More and more NDMP servers are being integrated with SANs, employing serverless backup techniques via SCSI extended copy. In many cases, the client is not aware that SCSI X-copy is being employed behind the scenes. This may need to be purchased as a separate—and perhaps expensive—option.

Serverless backup implemented by SCSI X-copy could also be used by NAS and backup software vendors as a means of moving data from disk to tape, as a part of their value-added NDMP offering. Some vendors are already moving in this direction.

A SAN and NAS-aware backup solution needs to automatically discover and configure SCSI devices on the network and be able to manage them without the need for operator intervention. The challenge is to select, control, and arbitrate these devices as part of a shared-backup solution. The NDMP discovery interface aids in this automatic discovery, with backup clients doing the management piece, resulting in a more “hands-off” solution.

## Conclusion

All of the developments discussed and others too numerous to cover here will

further the ease and flexibility of storage management and backup on heterogeneous networks. Most likely, no single technology or method will dominate. Increasingly, the solutions will be hybrids of several different technologies. The result is that end users will have more options to choose from for managing data storage and backup. □

---

**Soubir Acharya** is a senior software architect, and **Susan G. Friedman** is a technical analyst at Syncsort ([www.syncsort.com](http://www.syncsort.com)) in Woodcliff Lake, NJ.



50 Tice Boulevard  
Woodcliff Lake, NJ 07677  
(201) 930-8200  
<http://www.syncsort.com>