

How to build SCSI and SCSI+FC SANs

It's possible to build a SCSI-centric SAN, but a Fibre Channel/SCSI combination may offer advantages.

■ BY JERRY NAMERY

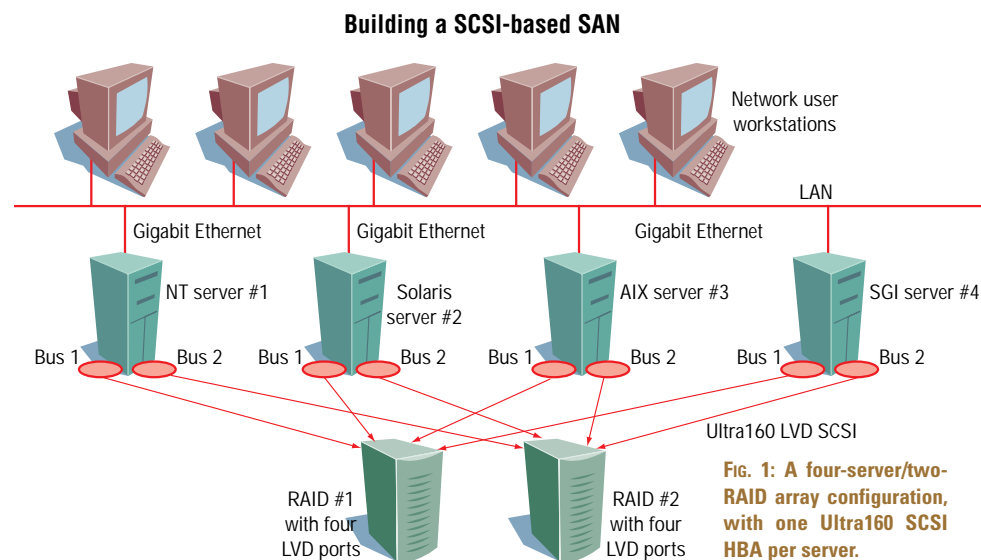
Can you build a high-performance storage area network (SAN) with tried-and-true SCSI? Yes. Can the same SCSI-based SAN take advantage of Fibre Channel, while reducing some of the SAN's limitations? Yes.

The key to the SCSI part of the SAN is the latest implementation of the venerable interface—Ultra160 SCSI. Backward compatible with previous versions

of SCSI, Ultra160 provides a 68-pin bus with 160MBps per port between systems and storage devices. Single PCI cards typically come with two of these ports, for an aggregate throughput of 320MBps. To achieve this speed, you need a PC with a 64-bit PCI bus running at 66MHz or better.

To build a SAN with Ultra160 SCSI, use a star topology with storage in the middle.

Ultra160 SCSI provides very reliable, point-to-point connecti-



ty. However, it does have a cable distance limitation of 25 meters from host to storage system. (For more information on Ultra160 SCSI, see "Ultra160 SCSI boosts performance, reliability," *InfoStor*, March, p. 44.)

To build a SAN with Ultra160 SCSI, use a star topology with the

storage in the middle of the star. A cable goes from the star to the storage to each one of the servers. For example, a typical RAID subsystem might support four servers sharing the same storage, vs. 126 nodes supported by Fibre Channel. However, four servers can provide a powerful network backbone. Four dual-port PCI adapter cards with eight Ultra160 SCSI ports provide more than 1Gbps of aggregate bandwidth. And this type of SAN

SAN solves traffic problems at pre-press plant

By Elizabeth Ferrarini

Some pre-press applications require the transfer of multi-megabyte files, which can take several hours over a LAN. A SAN offers a more effective way to transfer large files. A SAN comprises a central pool of storage with multiple host servers attached to RAID storage devices over a network of high-speed interconnects, such as SCSI and/or Fibre Channel. Unlike a LAN, a SAN enables applications with heavy file-transfer demands to have direct access to a shared-storage repository at considerably faster transfer rates.

A SAN helped R.R. Donnelley & Sons' Glasgow Division, in Glasgow, Kentucky, increase the speed of a color-swap application. This division does pre-press work and printing for magazines and catalogs, such as *PC Magazine*, *Yahoo Magazine*, and *Harvard Business Review*. For the color-swap application, the plant prints low-resolution versions of the original high-resolutions files. Users make changes to the low-resolution image files, which don't take up a lot of memory. When all of the files are ready for final printing, Open Prepress Interface (OPI) software swaps the modified low-resolution image files for the matching high-resolution files.

Tony Wallace, the systems analyst who set up the SAN, says "swapping 50MB files to 100MB files over a Gigabit Ethernet network was taking too long and was consuming too much bandwidth." To speed things up, Wallace created a SAN by moving two Intergraph InterServ 8400 servers, which run the same application, off the Gigabit Ethernet network, attaching them to a SCSI-based Winchester Systems FlashDisk RAID array, and adding Tivoli's SANergy middleware software.

The RAID array provides a terabyte-level central storage repository with 36GB disk drives. A single array can connect directly to 4 to 36 servers running different operating systems. SANergy, which runs on Windows NT, Mac, Sun, or SGI platforms, extends the file system so that multiple servers can share the same files directly on the RAID array. SANergy allows the files to be swapped via a point-to-point SCSI connection from the subsystem to the servers and from the servers to workstations. The software, which runs transparently on the servers, takes the data transfer off the LAN and moves it across the SAN.

This technique bypasses the network, reducing network congestion. Wallace says the configuration provides 10 times the performance of the previous Gigabit Ethernet network swap-out approach.

Elizabeth Ferrarini is a Boston-based freelance writer. She can be contacted at iswive@aol.com.

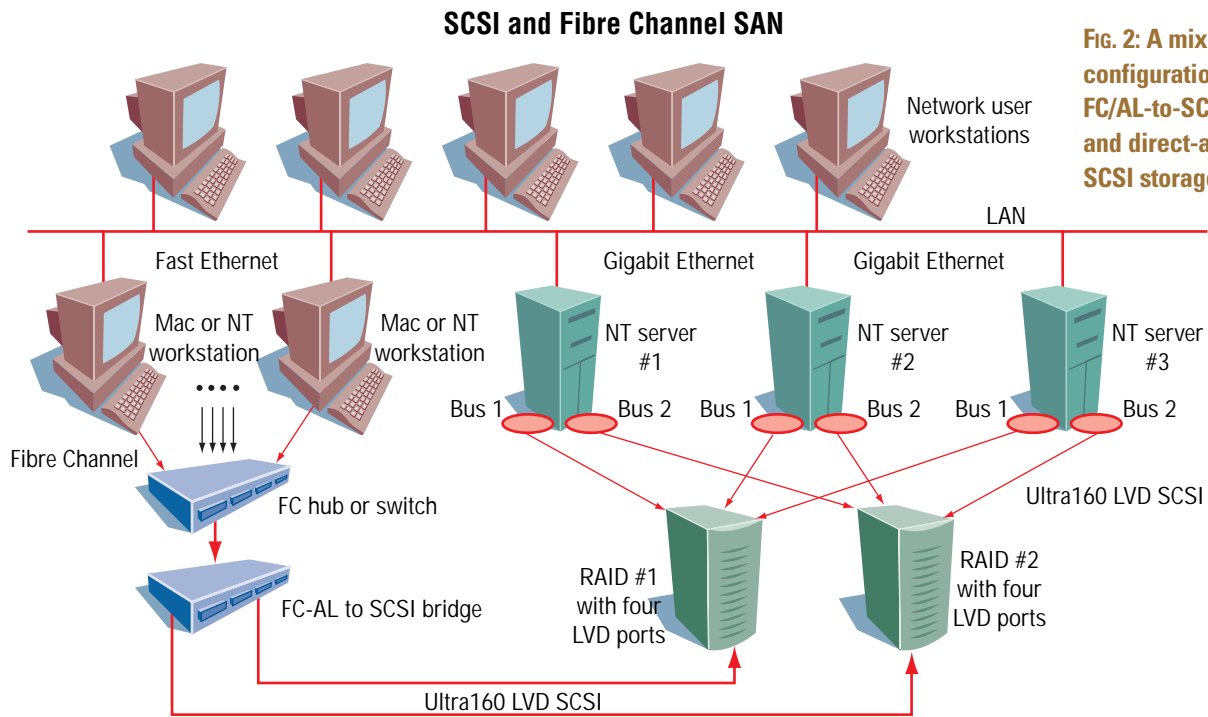


Fig. 2: A mixed configuration, with FC/AL-to-SCSI bridge and direct-attached SCSI storage.

offers quick and easy installation and testing.

Figure 1 shows four servers (with different operating systems) and two

RAID arrays. You can have multiple RAID systems for multi-terabyte requirements. This configuration also shows a single Ultra160 SCSI host bus

adapter per server, with two ports per adapter—bus 1 and bus 2. (The adapter is backward compatible with 80MBps low-voltage differential SCSI).

With middleware software, such as Tivoli's SANergy, any server in the system can fail and the entire configuration will keep running. The software gives the other servers in the network direct access to the storage. It also handles file locking and file management. The file handshake transaction comes over TCP/IP Ethernet. The data can be transmitted via Ultra160 SCSI at speeds 3X to 20X the bandwidth of TCP/IP Ethernet. This speed comes as a result of the servers' direct access to the file; no data is transferred over the network.

Mixed SCSI-FC SANs

For some environments, adding Fibre Channel to a SCSI-centric SAN has advantages. Fibre Channel runs the SCSI protocol at 100MBps per port over optical cables and runs a unique storage protocol at 1.06Gps in packets. (Fibre Channel does not currently run IP.) It uses a network topology with hubs or switches as concentrators. Fibre Channel typically supports up to 500-meter cable distances, which is suitable for most applications, although you can spend more money and purchase special cables and drivers for distances of up to 10 kilometers.

Fibre Channel-Arbitrated Loop (FC-AL) currently has one downside:

it runs Class-3 service. There are three classes for quality of service of transmission, and Class 3 does not guarantee transmission, nor does it acknowledge it. If Fibre Channel drops a packet and the software fails to catch it, the result is a hang-up (or a timeout), which causes the system to freeze for a second. A loop initialization process (LIP) then starts resetting the entire bus.

Because it's a point-to-point connection, SCSI doesn't have the benefits of running in a loop. However, you pay more for the benefits of Fibre Channel over Ultra160 SCSI.

Although 126 nodes per arbitrated loop are possible, you may have difficulty managing and debugging that many, and high node counts translate into decreased performance. As a result, you may want to limit the node count to 12 or 14.

In Figure 2, primary servers are attached directly to two RAID subsystems via Ultra160 SCSI connections. In addition, this configuration includes a low-cost Fibre Channel-to-SCSI bridge to convert SCSI to Fibre Channel.

You can attach a hub or switch to some workstations and servers. Why do this? Because you many want to dedicate these primary servers to mission-critical functions. Mac and NT power users or remote users, for example, may

need fast data access without clogging the network. They can connect via Fibre Channel to a Fibre Channel hub or switch, which in turn connects via the bridge directly to the RAID arrays.

Again, SANergy software provides the servers with simultaneous access to the storage on the shared RAID systems. Installing this software on the Macs and NT workstations gives them transparent access to the data.

If anything happens to the optical transmission or to the switch, The software automatically reroutes the traffic over the LAN. So if anything hangs up or if a connection fails, the critical SCSI-attached servers continue to run.

SCSI provides a simple, reliable, and fast connection. Fibre Channel provides connectivity for a large number of users, and for much longer distances. In the case of a failure, SANergy allows users to continue running (although not necessarily at optimal speed), until you have a chance to fix or replace the failed component. The end result includes the best of both worlds. □

Jerry Namery is chief technology officer at Winchester Systems Inc. (www.winsys.com) in Woburn, MA.

WINCHESTERSYSTEMS®

400 West Cummings Park, Woburn, MA 01801 USA
800-325-3700 • Int'l: 781-933-8500 • FAX: 781-933-6174

www.winsys.com

