New Storage Interfaces

A White Paper Prepared by The SCSI Trade Association

April 1997

SCSI Parallel Interface Technology -
a path to the future!

Table of Contents

Preface

This white paper is based on the article "New Storage Interfaces" authored by Greg P. Schulz of MTI that appears in the April 1997 edition of *Sys Admin Journal.* This paper contains additional material not contained in the above mentioned article as well as general formatting changes.

Introduction

Changes are occurring with storage technology including lower cost per megabyte, faster drives, higher bandwidths, and applications requiring larger amounts of storage and throughput. This has resulted in a storage and I/O growth explosion closing the gap between I/O, Storage, and CPU performance. Our discussion centers around storage interfaces that enable data to be stored and retrieved from storage devices (Disk Drives, RAID Arrays, Tape Drives, and Solid State Disk). Network interfaces such as Ethernet®, FDDI, Token Ring, Fast Ethernet, and ATM can also be used as storage transports for distributed or network file systems such as NFS. However, this discussion centers on technology used <u>primarily</u> for storage.

Parallel SCSI

<u>SCSI-1 and SCSI-2</u>

Today's most popular storage interface for high performance servers and workstations is the Small Computer System Interface (SCSI, pronounced "scuzzy"). Parallel SCSI (SCSI-1) was standardized in 1986 supporting seven 8-bit devices operating up to 5MB/second[1]. Parallel SCSI was further enhanced in 1992 with SCSI-2 (the current standard) implementing Fast (10MB/second) and Wide (20MB/second) data transfers (up to 32 bits however only 16 bits are currently used), Synchronous transfers, Differential signaling, Command Tag Queuing, LUN support, Parity Checking and connectors smaller than the SCSI-1 Centronics® 50-pin connectors.

Key components in an parallel SCSI implementation include a host adapter or SCSI interface, cabling (flat ribbon or round cable) to daisy chain devices (Disk Drives, RAID Arrays, Tape, etc.) together and terminators. Parallel SCSI busses must be terminated on each end of the bus with the host adapter typically providing termination. If a SCSI device does not provide termination at the end of the bus, then a physical terminator or plug must be used as illustrated in Figure 1.

Each SCSI node or device including the host adapter has a SCSI address or ID which must be unique and in the range of 0-7. Normally the host adapter will use SCSI ID 0 however some systems will use a different SCSI ID for the host adapter. It important to insure proper termination and assignment of SCSI IDs and avoid ID mismatch to prevent problems. SCSI-2 also provides LUN support which enables multiple logical or virtual devices to share the same physical ID, however each logical unit has a unique sub-address as is common with RAID Arrays. As shown in Figure-1, SCSI implementations are normally single host configurations however multi-host and multi-initiator configurations can also be setup. A multi-host or multi-initiator includes more than one host having access to the same SCSI devices over the same SCSI bus as in a Windows NT, Unix, or OpenVMS SCSI Cluster.

Parallel SCSI (named for its parallel architecture where data is sent in parallel over a series of wires to maximize performance) is what people mean when they refer to "SCSI". For example, in narrow mode, eight data bits are sent in parallel over a series of parallel wires while wide transfers are performed by sending 16 bits in parallel. With wide (16-bit) SCSI, there are twice as many data bits and data wires (paths) to double the maximum bandwidth. Wide SCSI can improve performance during the data transfer phase for applications performing large I/O such as imaging, sequential processing, and backups. The maximum bandwidth numbers include data transmission, command execution, timing and other overhead. For example, a wide interface (maximum bandwidth of 20MB/second) can perform around 15-18MB/second data transfer rates factoring in overhead.

1 Throughput, transfer rates and bandwidth speeds are sometimes quoted using mega<u>bit</u> per second (often abbreviated as M<u>b</u>/second) as compared to mega<u>byte</u> per second (abbreviated as M<u>B</u>/second). With eight bits per byte and parity overhead, a rough translation is 10Mb = 1MB.

Parallel SCSI bus distance varies from 3-25 meters depending on attributes such as Single-ended or Differential. Differential (High Voltage Differential or HVD) parallel SCSI operates at distances up to 25 meters using a wider cable with extra wires to compare differences in signals and make adjustments to the signal due to noise and other factors as needed.

There are shortcomings with parallel SCSI; however, improvements such Ultra-SCSI and new serial storage interface technology (SSA and Fibre Channel) are overcoming some of the limitations. Parallel SCSI should remain viable for some time (past the year 2000) as more devices and systems are added on a daily basis. New storage interfaces discussed here provide backward compatibility with existing SCSI devices using bridges. This enables customers to maintain their investment in existing devices while migrating to the newer technologies.

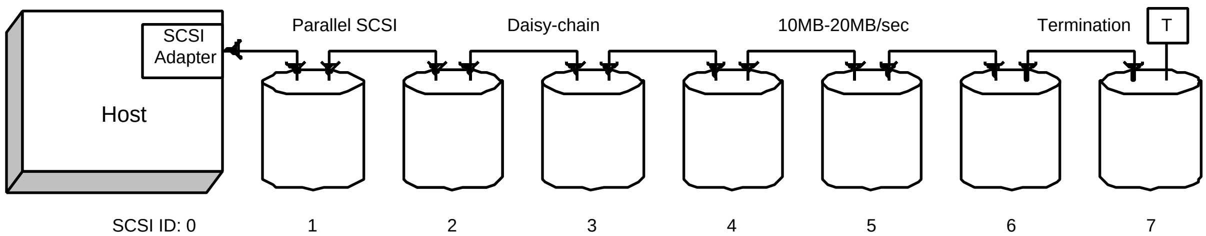


Figure 1. Parallel SCSI-1 and SCSI-2

SCSI-3

SCSI-3 improves on SCSI-2 by increasing the number of devices, support for Ultra-SCSI and Fibre-Channel, SCAM (automated configuration with no need for device jumpers), and new device connectors including SCA (Single Connect Assembly), SCA-2 (SCA with hot plug capability) and VHDCI (Very High Density Cable Interface). SCA eliminates the need for cumbersome cables required to interface drives and canisters to backplanes. VHDCI implements a reduced size cable interface than enables up to four wide SCSI interfaces on a single PC board or attach to a PCMIA type II card.

Ultra-SCSI (FAST-20)

Ultra-SCSI implements the SCSI-3 protocol doubling the maximum bandwidth to 40MB/second. Ultra-SCSI is parallel SCSI with faster performance using a new signaling and clocking technique. Ultra-SCSI can coexist with SCSI-2 devices; however, the maximum speed or bandwidth is limited to the slowest device on the bus. For example, assume an server with a Ultra-SCSI adapter has a SCSI RAID (FWD) device and an Ultra-SCSI RAID device on the same bus. In this case, the bus will run at a maximum throughput of 20MB/second. These limitations apply when SCSI-2 devices are mixed on the same bus using the same controller. When devices are connected to separate unique adapters, these limitations do not exist.

Some Ultra-SCSI items to be aware of include:

- Ultra-SCSI adapters and drivers for HP, IBM, Sun, SGI, Digital, and Windows NT platforms should be available by year-end.
- Verify that operating system revision level supports Ultra-SCSI adapters and the devices that will be used. Patches or drivers may be needed and there may be device configuration restrictions.
- Ultra-SCSI is less forgiving than existing parallel SCSI due to faster clock cycles. This will require closer attention to detail with cabling, types of cable used, number of devices per bus, and termination.
- Adhere to Ultra-SCSI guidelines placing older, slower SCSI devices and non-hard disk devices (CD-ROM, scanners, etc.) on separate host adapters.

Ultra-SCSI provides a “mid-life kicker” for parallel SCSI by doubling the bandwidth and increasing the number of devices. Future enhancements of Ultra-SCSI call for doubling of bandwidth to 80MB/second and implementing Low Voltage Differential SCSI (LVD). This new version has been referred to as Ultra2-SCSI with some technology, such as LVD, now under development. LVD as its name implies, uses low voltage differential signaling compared to high voltage differential (HVD) with existing parallel SCSI increasing performance and distance.

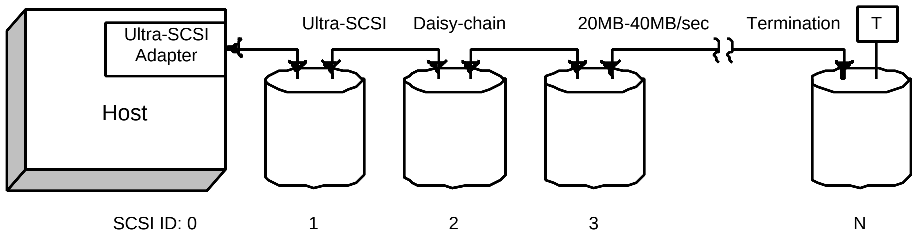


Figure 2. Ultra-SCSI

New Serial Storage Interfaces

New serial storage interfaces include Serial Storage Architecture (SSA) and Fibre Channel that address some of the deficiencies of existing parallel SCSI and other storage interfaces. Features found in serial interfaces include fiber optic cables, as well as copper, dual-porting or dual-Pathing for fault-tolerance, higher performance, and support for more devices over longer distances. Unlike parallel interfaces that send data bits over individual parallel wires, serial technology sends data bits in series as packets over a single wire simplifying data transmission and reducing overhead. Though this may seem slower, by simplifying the underlying signaling and timing mechanism, serial technology can significantly outperform existing parallel technology.

Serial Storage Architecture (SSA)

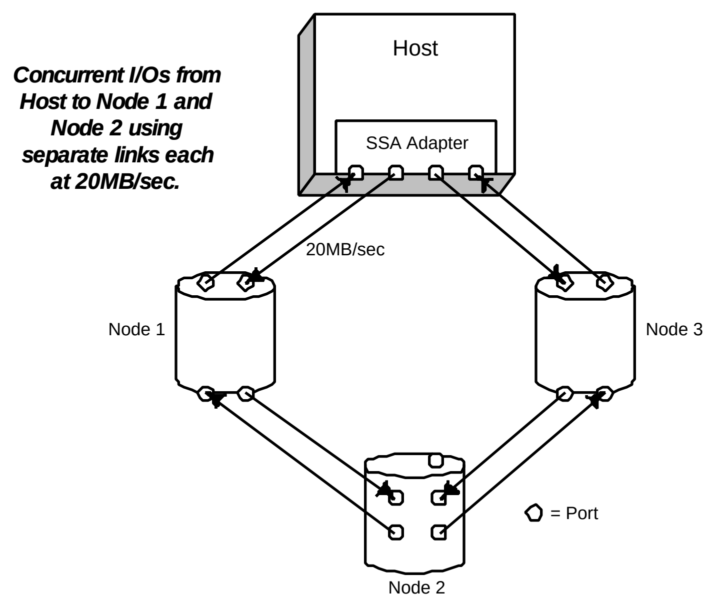
Serial Storage Architecture (SSA) has been designed by IBM® in partnership with other vendors as a new open storage interface standard (ANSI X3T10.1) to replace parallel SCSI as well as other older and slower interfaces. One of the key attributes of SSA is “Spatial Reuse” that enables concurrent traffic to exist on a bus increasing the aggregate bandwidth. Unlike parallel SCSI storage which is configured as a bus or chain, SSA devices are configured in loops with daisy chained connections between devices. Each node or SSA device can support two loops each with two links (in and out). Spatial Reuse allows concurrent data transmission on both loops which increases bandwidth without the need for arbitration.. However, the maximum bandwidth of any given link is only 20MB per second (the same speed as Fast Wide parallel SCSI or one-half the speed of Ultra-SCSI (40MB/Second)).

Improvements in packaging have enabled SSA devices to be hot-swappable and auto configurable. SSA-to-parallel SCSI bridges such as that from Vicom, enables existing parallel SCSI devices to be accessed from an SSA system and vice versa. When using a bridge device, “Spatial Reuse” does not apply to the parallel SCSI devices. A single SSA loop can support 127 devices compared to 15 for parallel SCSI with distances up to 15 meters between the devices. A four-conductor copper wire cable with a four-pin connector replaces the 50 or 68 pin connectors and cables used with traditional parallel SCSI. This reduces cost, as well as simplifies installation and configuration. Some host adapters such as those for IBM RS/6000® systems provide connection for up to four loops.

A future goal for SSA is to double the speed from 20MB/second per link to 40MB/second. Using Spatial Reuse, total bus bandwidth would increase from 80MB/second to 160MB/second. A practical limit of 96 devices per adapter is suggested for SSA implementation by some vendors. Benchmarks by IBM have shown about 35MB/second at the processor using realistic workloads which compares with about 12MB/second using a Fast Wide SCSI (20MB/second) adapter on an RS/6000. It should be noted that several drives must be concurrently accessed to achieve these types of data rates. Although the 35MB/second is considerably lower than the 80MB/second theoretical transfer rate, IBM’s MicroChannel® internal bus on the RS/6000 appears to be a bottleneck. Additional IBM laboratory testing has shown that adapters based on the newer PCI bus may be able to process 70MB/second at the adapter level.

Some SSA issues to be aware of include:

- SSA is well suited for multi-host or clustered systems where connectivity and aggregate storage bandwidth are important. SSA is not well suited for single stream, bandwidth-intensive applications.
- Although not being widely accepted by the industry in whole, some users of IBM systems are implementing SSA storage such as the IBM 7133.
- SSA uses an addressing scheme that requires knowledge of the path to the device. Should the path change, such as when a drive is removed, the address must be recalculated. An analogy would be telling a cab driver where to go. With SSA, the driver might be told to go two blocks, turn right, go one block, turn left, then go two more blocks and stop at the corner house on the right. With this approach, the cab driver is not able to optimize the route (which may or may not be a good thing to do). This scheme works well as long as there are no detours that would require setting up a new route.



<u>Fibre Channel (FC)</u>

Work began in 1988 to develop a replacement for existing storage interfaces including parallel SCSI, HiPPI and ESCON. Fibre Channel is an open standard (ANSI X3T11) that can be used as a storage interface as well as a network interface for protocols such as TCP/IP. Similar to parallel SCSI, Fibre Channel is evolving with several standard speeds and topologies including point-to-point, fabric (a type of network) and Arbitrated Loop. Fibre Channel is capable of speeds up to 1Gbit per second (100MB/second) in each direction simultaneously (400MB/second Full Duplex). In addition, Fibre Channel can support longer distances of up to 10Km compared to Parallel SCSI 25m.

The spelling of “Fibre” compared to “Fiber” for Fibre Channel is intentional. Fibre Channel can be implemented using copper wire as well as fiber optic cables. Fibre Channel devices utilize a Global Link Module (GLM) which facilitates easy conversion from copper wire to optical cable. GLMs also make it easy to increase speed or bandwidth such as when migrating from _ speed to full speed (100MB/seccond).

Fibre Channel establishes a point-to-point connection between devices providing up to 100MB/second bandwidth to each device. Although this bandwidth is faster than SSA, the downside is only one device can transmit or receive over a link at a time (similar to Parallel SCSI). This means that multiple devices that want to transmit or receive at the same time may have to wait for the bus. Each Fibre Channel connection has a send and receive port and dual loops can be implemented for higher aggregate bandwidth (Full Duplex). Fibre Channel ports include N_PORT, F_PORT, NL_PORT, and FL_PORT that are used to establish links for point-to-point, arbitrated loop, and fabric (i.e., network) topologies. N_PORT refers to a node or device port that connects to an F_PORT that is part of the fabric or perhaps to another N_PORT in a point-to-point topology. The "N" indicates a node or device while the "F" indicates the port is part of the fabric. Arbitrated loop ports are referred to as FL_PORT and NL_PORT with the letter "L" indicating that it supports the Arbitrated Loop protocol. Fibre Channel in simple configurations can be implemented with out switches or hubs.

Point-to-point, Arbitrated loop and fabrics may be combined such as a workgroup being on an Arbitrated loop with nodes connected to a fabric via a switched network to the rest of the company. A particular node may have local disk devices attached via an Arbitrated loop with a direct connection over a N_PORT to a fabric. Likewise, a disk array may be attached to a workstation or server via a point-to-point connection. The port limit for point-to-point connections is two ports connected via a single link.

Fabric devices are referred to as elements that can be thought of as a switch or router. Fabrics can consist of a single fabric element (a node or device) or several fabric elements. This enables new technology to be introduced in the middle of the fabric without disrupting nodes on the outer edge of the fabric. This is similar to how a phone system works. The phone company can change their internal networks transparently (such as when they implemented digital switching) while ensuring that the signals will be the same for your phone. In a fabric, any node can talk to any node with the fabric performing the required switching and routing, thus providing a peer-to-peer service. This enables multiple protocols to exist at the same time on the fabric or loop. The individual nodes are responsible for speaking the proper protocol or command set (similar to a telephone conversation where it is the people on the line that must understand what each other are saying).
Fibre Channel issues to be aware of include:

- Early Fibre Channel products, such as storage devices from Sun and HP, use slower point-to-point fiber connections and are not compatible with the emerging FC-AL standard.
- Fibre Channel support of various topologies presents configuration issues that require careful planning and implementation. Several host adapters are available for different host buses such as S-Bus®, HP-PB and PCI.
- Verify operating system compatibility with the devices and adapters that will be used. Determine if any special patches or drivers are needed.
- Fibre Channel uses a flexible addressing scheme that relies on knowledge of where the destination is and leaves the routing up to the network. Using the previous cab driver example, Fibre Channel packets contain a destination that is used by Fibre Channel components to ensure delivery. Knowledge of the actual route is not needed by the sender; only the destination address is required. Fibre Channel works on the premise the cab driver is given the address and location and he/she is free to determine the best route to get there.

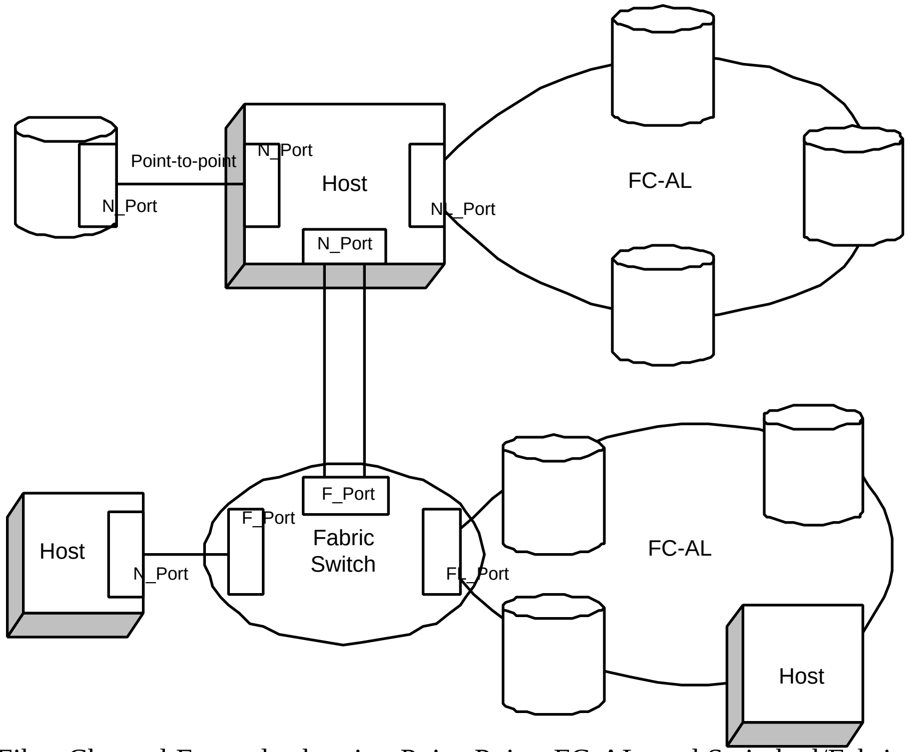


Figure 3. Fibre Channel Example showing Point-Point, FC-AL, and Switched/Fabric topologies

Storage Interface Comparison

By now you should have a better understanding of existing as well as new storage interface technologies. Let's look at some questions and issues to consider when evaluating and selecting a storage interface.

1. Does the host platform support Ultra-SCSI or will it?
2. Does the host platform support Fibre Channel and, if so, which variant?
3. What operating system version is needed to utilize a new bus technology?
4. How will applications benefit from a new I/O bus (e.g., performance, distance, more device, etc.)?
5. Do the devices that will be used support any of these new interfaces?
6. Are there current problems that will be addressed by new technology?
7. Are there cabling or distance issues?
8. Are all infrastructure components ready such as hubs, switches, etc.?
9. Are device ports or adapter connections available?

With these and other questions in mind, you should now be able to better understand the following information and apply it to your requirements.

Interface[2]	Maximum Transfer Rate (MB/sec)	Bus Width (bits)	Maximum Bus Length[3] (meters)			Maximum Devices[4] (Loads)
SERIAL						
			Single-ended	**Differential(HVD)**	**LVD**	
SCSI-1	5	8	6	25	NA	8
Fast Narrow SCSI[5]	10	8	3	25	NA	8
Fast Wide SCSI[5]	20	16	3	25	NA	16
Ultra-SCSI[6]	20	8	1.5	25	NA	8
	20	8	3	NA	NA	4
Wide Ultra-SCSI[6]	40	16	NA	25	NA	16
	40	16	1.5	NA	NA	8
	40	16	3	NA	NA	4
Ultra2 SCSI[7]	40	8	NA	25	25	2
	40	8	NA	12	12	8
Wide Ultra2 SCSI[7]	80	16	NA	25	25	2
	80	16	NA	12	12	16
PARALLEL						
			Copper	**Fibre**		
SSA	80[8]	16	10m	2,450m		127
FC-AL	400[8]	16	100m	10,000m		126

Table 1. Storage Interface Technology Comparison

Beyond Fibre Channel and SSA

A new storage interface combining the best features of SSA and Fibre Channel has been proposed resulting in a "cease fire" in the storage interface wars (some viewing IBM as surrendering, although dictating terms of the surrender!). This initiative is being led by Adaptec®, Seagate and IBM with several other system vendors participating in the review process. If approved, this should result in a hybrid serial storage interface that combines the best of Fibre Channel and SSA while maintaining backward compatibility. Although details are still being worked out, Seagate has indicated that future Fibre Channel drives will be compatible with the new interface. The new interface does not yet have a formal name, however, many have nicknamed it Fibre Channel-Enhanced Loop or FC-EL (the use of which is being discouraged by the combined groups). This new interface has been submitted to the ANSI review process, which if approved should result in a formal name sometime in the future with actual products not available until late in the 1990's.

Summary

Although there are shortcomings with parallel SCSI, it should continue as a popular storage interface for some time. Ultra-SCSI is becoming popular as a high performance storage interface to access existing and new parallel SCSI devices. New serial storage interfaces (SSA and Fibre Channel) should gain popularity in late 1997 and early 1998 with support for existing parallel SCSI devices via

2 The Parallel SCSI names (e.g., SCSI-1, Fast SCSI, etc.) are the official names used by the SCSI Trade Association and American National Standards Institute.
3 With Ultra-SCSI, the maximum bus distance is affected by the number of nodes (devices or loads) on the bus. This is due to tighter signaling requirements to support the higher Ultra-SCSI speed.
4 Includes host adapter. Actual number of attached devices is one fewer.
5 Fast and Fast Wide also known as SCSI-2.
6 Ultra-SCSI and Wide Ultra-SCSI are also known as SCSI-3 and/or Fast 20.
7 Ultra2 SCSI and Wide Ultra2 SCSI are also known as Fast 40.
8 Full duplex.

bridges. It will probably take several years before these (or the proposed "best of both worlds") technologies become as available and common in the marketplace as SCSI is today.

SSA and Fibre Channel have benefits and weaknesses that should be considered with respect to your environment. SSA has a higher aggregate bandwidth for multiple systems (Spatial Reuse) while Fibre Channel has higher individual device bandwidth. Look for a new hybrid serial storage interface with the best of SSA and Fibre Channel. This new proposed interface should provide backward compatibility for Fibre channel devices and compatibility with SSA devices via a bridge.

Applications requiring backward compatibility with existing parallel SCSI devices including disk, RAID and tape will find Ultra-SCSI a good fit. Similarly, for systems that do not yet support Ultra-SCSI, SSA, or Fibre Channel adapters, Ultra-SCSI devices can provide improvement in performance over parallel SCSI. Applications requiring high data transfer (throughput) rates from individual storage devices, such as RAID arrays, will find Ultra-SCSI and Fibre channel better suited than SSA. Databases and high bandwidth applications such as imaging, batch processing and video are good examples of high throughput applications. SSA, on the hand, is better suited for applications requiring simultaneous access to multiple devices from one or more hosts. Examples include host-based RAID, databases distributed over several drives, and server clustering.

Finally, keep in mind there is not a right or wrong adapter or interface as long as it meets the needs of your environment for <u>your</u> applications.

Acknowledgements

The STA would like to thank the author of this white paper, Greg P. Schulz of MTI, 4905 E. La Palma Ave., Anaheim, CA, 92807. Please visit the MTI home page at <www.mti.com>.

Additional Information

The SCSI Trade Association has a wide variety of documents and information on SCSI Parallel Interface Technology, including presentations, articles in periodicals, seminar material and white papers. Please contact Association offices at:

The SCSI Trade Association
404 Balboa Street • San Francisco, CA 94118
Tel: 415-750-8351 • Fax: 415-751-4829
Info@scsita.org • http://www.scsita.org/