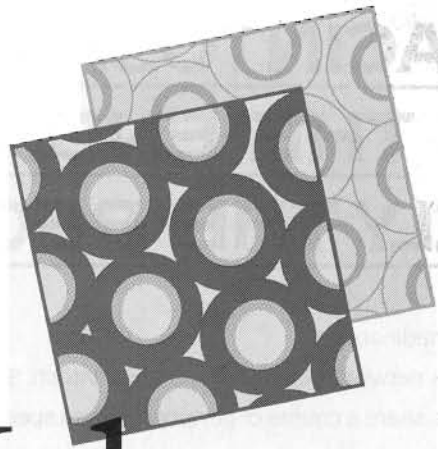


VAXcluster Tape Subsystems

SCSI-based alternatives to tape drives provide needed capabilities.

BY BRADFORD T. HARRISON

AS STANDARDS GROWTH brings to market an ever-greater range of compatible products from third-party suppliers and manufacturers, system managers are under increasing pressure to select between equipment from their system vendor and third-party sources. Purchasing products from the system vendor may seem a safer choice. But with the technical skill, support capabilities and marketing strength that standards growth has fostered among third parties, the capabilities of third-party products prove attractive.

Nowhere is this seen more clearly than in tape drives for VAXcluster systems. DEC's limited choices in the TA79 and TA81 products fall far short of more innovative drives from such companies as Fujitsu, Storage Technology Corporation and Exabyte Corporation. Drives from these companies are finding widespread market acceptance because of their technical sophistication, performance, reliability and support of industry-standard interfaces, such as SCSI. With these drives, a much wider range of capabilities becomes available at a lower cost than the DEC drives — capabilities such as large capacity, unattended backup, transfer rates as high as 3 MB per second and DEC-to-IBM interchange.

An Open Invitation

Tape products in general have proved a thorn in DEC's side, culminating in the problems the ill-fated TK50 experienced in the field shortly after shipping. DEC's tape offerings serve more to round out its product line than to offer innovative storage solutions.

This lack of technical sophistication in tape products is an exception to the rule that DEC strives to operate on the leading edge of commercial technology. But it points out a major discrepancy in the company's approach to handling customers' data storage and access needs. DEC's Digital Storage Architecture (DSA) encourages replacement of older hardware with up-to-date products. But DEC often fails to offer those products in a timely manner, or with the enhancements that are the hallmark of third-party companies, such as System Industries and others.

DSA offloads all device-specific functions to the peripheral subsystem itself, handling disk and tape as classes of intelligent devices rather than as specific devices with specific geometries. Peripherals thus can be added or replaced in the system with no changes to operating system software or applications. It would follow that DEC would equip all of its systems with industry-standard intelligent interfaces, such as SCSI, so that customers immediately could take advantage of new storage technology.

But this isn't the case. In fact, in recent years DEC has moved in the opposite direction, attempting through specialized interface hardware to close off its systems to the industry. System managers, therefore, need to pay close attention to the development of standards and the products that standards are helping bring to market. Currently, SCSI-based products provide a good hardware example of the trend, and a wealth of products exists. Third parties are making SCSI-based tape drives available to large DEC systems even though DEC provides no native SCSI interface on VAXclusters. These drives are added to the system to increase performance and reliability over similar products from DEC, and to provide the customer with important capabilities not obtainable from DEC's VAXcluster tape products.

SCSI And VAXclusters

SCSI host adapters have been available for Q-bus and UNIBUS systems for more than three years, and many DEC customers have outfitted their PDPs, MicroVAXs, VAXstations and UNIBUS VAXs with a wide variety of SCSI-based products.

Currently, the ANSI-standard synchronous version of SCSI that follows the Common Command Set is the specification to which most third-party companies are designing product. This version of SCSI is yielding transfer rates as high as 4 MB per second, a more than two-fold increase in performance over earlier asynchronous SCSI. This throughput is acceptable for support of high-speed fixed Winchester drives, but the vast majority of manufacturers continue to implement the Storage Module Drive (SMD) interface for high-end systems, including VAXclusters.

THIRD PARTIES ARE making SCSI-based tape drives available to large DEC systems even though DEC provides no native SCSI interface on VAXclusters.

The situation is different for tape. Tape drives are slower devices, accepting data at about 1 MB per second at most. Therefore, SCSI for high-end systems got its start with these drives and will continue to offer the required performance as tape-drive throughput increases. Equipping VAXclusters with SCSI-based tape products occurs with absolutely no interface performance penalty for any available tape drive. In fact, the interface opens these systems to the highest tape drive transfer rates the industry has to offer. This is exemplified in the recent System Industries' introduction of an IBM 3480-compatible product that features a throughput approaching 3 MB per second across a SCSI interface.

Transparent Integration

Integration of SCSI tape drives into the operating system environment of a VAXcluster is a simple matter. This is because DSA allows subsystem-based intelligence to offload the CPU of all I/O tasks specific to the peripheral device. As far as the operating system is concerned, peripheral storage consists of some number of logical units, and each logical unit consists of some number of 512-byte logical blocks. The operating system polls at boot time to discover the number and sizes of logical units on the system. The standard peripheral drivers under VAX/VMS are the Mass Storage Control Protocol (MSCP) disk driver (DUDRIVER) and Tape Mass Storage Control Protocol (TMSCP) tape driver (TUDRIVER). To the MSCP driver, disk storage consists of some number of randomly accessible logical units filled with some number of logical blocks. The same is true of the TMSCP driver, but it views its storage blocks as sequentially accessible only.

The MSCP and TMSCP drivers communicate with their intelligent peripheral controllers via MSCP and TMSCP command and message packets. Typical packets send and receive data, cause formatting and request diagnostics. Again, the drivers know nothing of the specific characteristics and geometries of the devices with which they're doing business.

The advantages to this architecture are many, but two of the most important are:

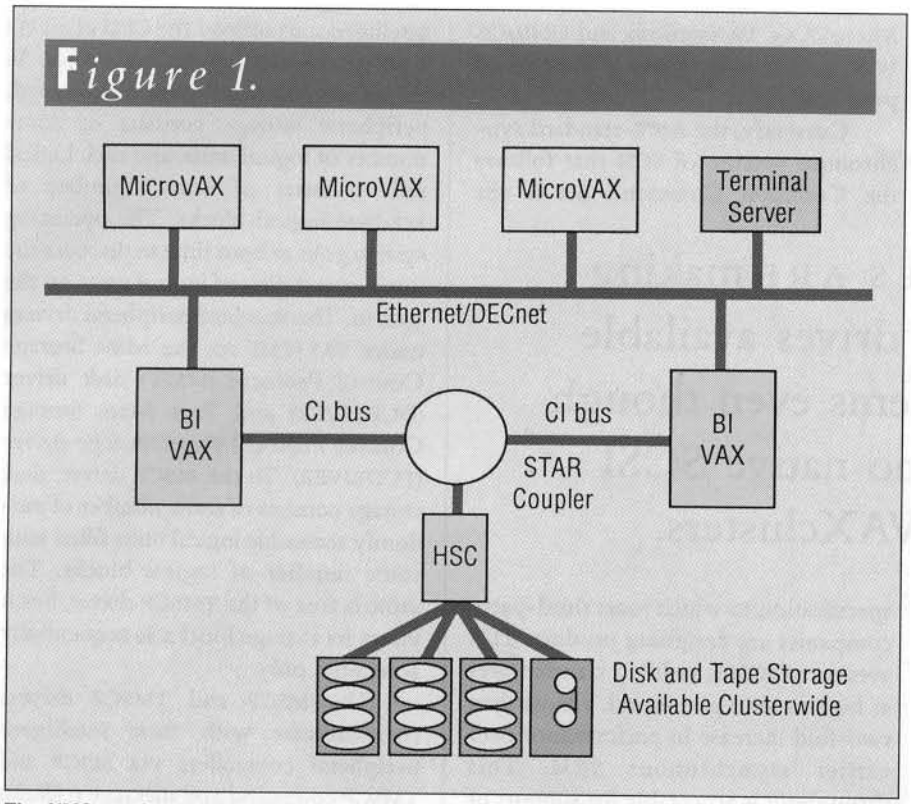
1. Peripherals can be added or replaced with no change to system software.
2. Peripherals can be shared by multiple processors with full file sharing, including record-locking capabilities.

These are two of the most important features of VMS and allow for construction of clusters.

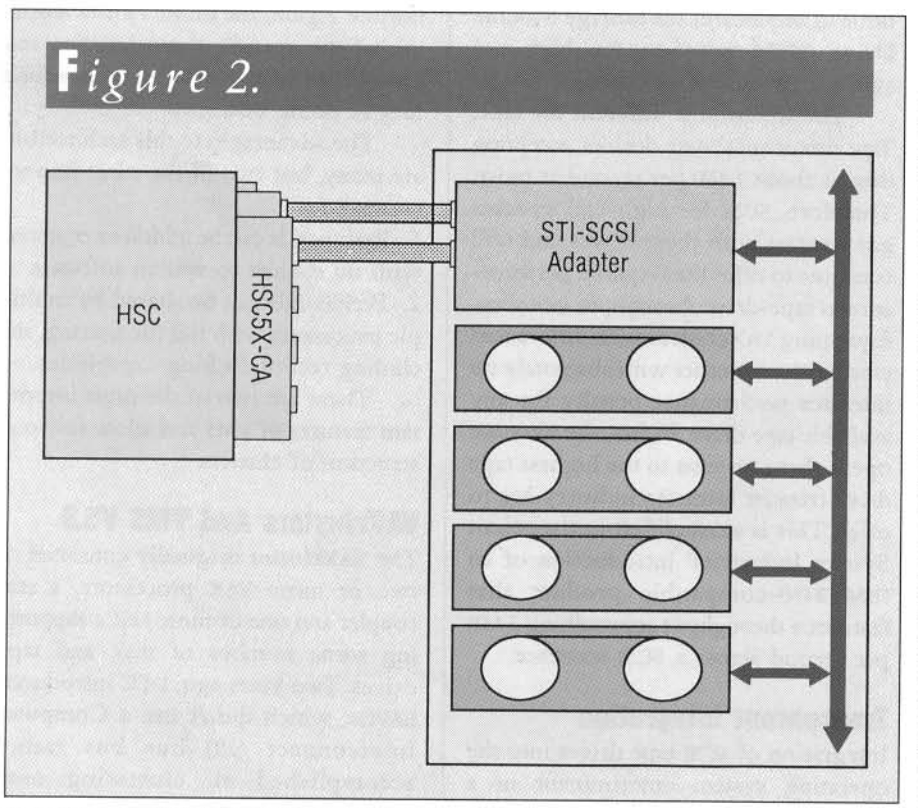
VAXclusters And VMS V5.0

The VAXcluster originally consisted of two or more VAX processors, a star coupler and one or more HSCs supporting some number of disk and tape drives. Two years ago, DEC introduced LAVCs, which didn't use a Computer Interconnect (CI) bus but rather accomplished all clustering over Ethernet.

Now, with VMS version 5.0, the



The VMS version 5.0 VAXcluster.



Connection of SCSI tape drives to HSC and SCSI bus.

two types of clusters have been combined, serving one, large base of mass storage to as many as 42 processors running VMS (see Figure 1). A single copy of VMS can be kept on the system disk and served to all processors in the cluster. Similarly, all files are served down to the record level clusterwide from the single bank of storage peripherals. This kind of centralized storage solution for the entire enterprise is finding greater utility than configurations using widely distributed storage units.

In VAXclusters, tape plays an important role as a backup device, although tape often is used for archiving, batch processing, software distribution, data interchange between systems and journaling.

In VAXclusters, the HSCs serve as the intelligent controllers that receive the MSCP and TMSCP command and message packets. HSCs use the DEC F-11 (HSC50) and J-11 (HSC40 and HSC70) 16-bit CPUs. HSCs handle such functions as command queuing, data buffering and reading and writing of data, and they contain on-board utilities for formatting, copying and archiving without host-CPU intervention.

HSCs are configured with the HSC5X-BA (disk) and HSC5X-CA (tape) channel cards to provide the Standard Disk Interface (SDI) and Standard Tape Interface (STI) connections to the supported disk and tape drives.

The HSC40 and HSC70 share the same performance, but the HSC70 comes with 32 ports, whereas the HSC40 comes with 12 ports expandable to 32. The HSC50 supports fewer I/O requests per second and supports 24 devices.

Installing SCSI Tape Drives

SCSI tape drives are installed on a VAXcluster by cabling a single- or multiple-drive tape subsystem containing an STI-to-SCSI adapter directly to STI connectors on one or two HSC5X-CAs (see Figure 2). This must be done when the cluster is powered-down, and the cluster learns of the new subsystem during polling at the next system boot. If the

subsystem is cabled to more than one HSC5X-CA, the channel cards may be in either the same or different HSCs.

Because SCSI is a bus-oriented interface, a multiple-drive subsystem requires only one STI-SCSI adapter. Additionally, as is possible with the DEC TA79, some reel-to-reel drives may be configured in a master-slave relationship. Under SCSI, just the master drive resides directly on the SCSI bus. The slave drives are daisy-chained from it. This configuration is used to reduce the cost of the subsystem: Only the master drive contains the formatter. SCSI allows each drive to be viewed separately by the operating system in either the master-slave configuration or in the configuration where each drive has a separate address on the SCSI bus. No changes in operating software are required for operation of all supported drives in either configuration.

UNDER SCSI, just the master drive resides directly on the SCSI bus.

The SCSI drives can be used for all functions the TA79 and TA81 are used for and run under all VMS software that uses the TMSCP device driver. CPU-controlled or HSC-controlled backup and restore operations work in the same way, and you can manage your tape systems with any VMS tape management software, such as DEC's Storage Library System (SLS) package.

STI-SCSI Adapter Operation

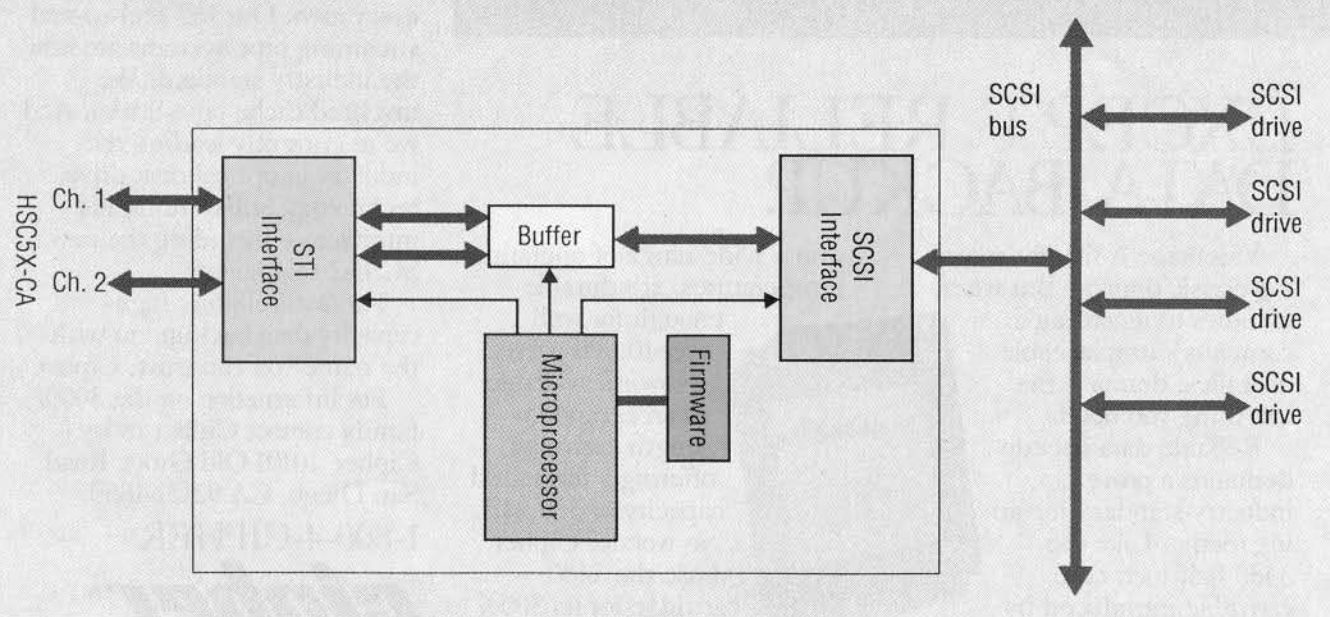
At the heart of a VAXcluster SCSI tape subsystem is the STI-SCSI adapter (see Figure 3). This adapter consists of five primary components: the STI interface, the SCSI interface, a microprocessor, firmware (EPROM) and a RAM buffer.

The STI interface supports two STI channels for high-performance dual-ported applications and for redundancy in the event of HSC or HSC5X-CA hardware failure (at which point autofailover occurs). The STI connectors on an HSC5X-CA provide signals meant to control DEC tape drives with the vendor-unique STI interface, but the adapter translates these signals into the appropriate SCSI commands for transmission over the SCSI bus. Operation of the STI-SCSI adapter is in fact similar to the operation of a DEC tape drive STI adapter board, which converts STI signals for operation with the native Pertec interface of the DEC drives.

STI-to-SCSI conversion operations are handled by the microprocessor under control of the firmware. During data transfer operations, the RAM is used as a data buffer in either direction and via either STI channel. Data is buffered for each drive supported. In this way, data transfers to and from the drives continue uninterrupted and at maximum drive transfer rates.

An STI-SCSI adapter can handle as many as four drives directly on its SCSI bus. SCSI bus devices are daisy-chained

Figure 3.



The STI-SCSI adapter.

together using a common cable, and each is assigned an address. All signals are common between all devices. SCSI-bus signals are either control signals or data signals (see Figure 4). The SCSI bus is terminated at both ends.

When the adapter is to connect with a drive, as directed by an I/O request from the HSC, it puts that drive's address out onto the SCSI bus. It's acting as initiator on the SCSI bus, and the drive then establishes itself as a target when it responds to the request for service. The adapter is assigned an address on the SCSI bus — the highest address. Distributed arbitration awards control of the bus to the device with the highest address that's contending for use of the bus. Under distributed arbitration, both the adapter and the drives compete for bus control.

Distributed arbitration radically speeds the SCSI bus. It allows for disconnect/reconnect operation by which a drive disconnects from the adapter after receiving a SCSI command and reconnects (by arbitrating for the bus) after it's finished executing the command. In this way, overlapped operations are performed on all four drives so that the delay caused by one drive, especially during data transfer operations, won't hold up the others.

When installing the SCSI tape subsystem, it's important that the configuration of the cluster be analyzed to ensure that no hardware bottlenecks currently exist or are created. Because of the architecture of DSA, cluster bottlenecks easily are removed by simply adjusting the hardware configuration as opposed to manipulating VMS tunable parameters. When the subsystem is installed it's also important that it be cabled into the cluster in such a way that it operates at full capacity.

Configuring Cluster Storage

As intelligent controllers, HSCs can command queue I/O requests, overlap seeks and perform simultaneous data transfers to and from supported devices. If VAXcluster performance is unacceptable and the cause doesn't lie with disk or tape drive performance, the most dramatic improvement is to upgrade to HSC70s, add additional HSCs or dual port the drives.

Mechanical aspects of computer systems ultimately prove to be the limiting factor in performance, so if all other system bottlenecks have been removed and the cluster is still sluggish, an upgrade to faster drives is necessary. In tape, the higher transfer rates offered by the new SCSI-based devices from Fujitsu and Storage Technology will reduce backup and archiving time drastically, and access time during batch processing and journaling will be much improved. Even in a fully configured subsystem consisting of four of these devices, maximum throughput can be obtained by dual porting to two HSC5X-CAs or, if necessary, two HSC5X-CAs in either the same or different HSCs.

The flexibility of the DSA architecture allows the integrator to implement the most efficient configuration.

Cluster performance improvements

Figure 4.

Busy

Select

Control/Data

I/O

Message

Request

Acknowledge

Attention

Reset

Data bits plus parity

SCSI-bus signals.

Figure 5.

- Test unit ready
- Rewind
- Request sense
- Read block limits
- Read
- Write
- Write filemarks
- Inquiry
- Space
- Recover buffered data
- Mode select
- Reserve unit
- Release unit
- Copy
- Erase
- Mode sense
- Load/unload
- Receive diagnostic results
- Send diagnostics
- Prevent/allow medium removal

SCSI tape drive commands.

also can be obtained through the use of on-board HSC utilities. For example, the backup/restore utility on-board these controllers dramatically decreases backup time, because the data doesn't pass through VAX main memory. Multiple disk and tape drives can be supported simultaneously. However, with this utility, data compression during restore doesn't occur. Only physical backup can be performed — not image or incremental. For data compression to occur, the data will have to pass through a VAX processor and memory.

SCSI Tape Drives On VAXclusters

Figure 5 lists the SCSI commands used for tape storage. All commands have counterparts in the TMSCP command set, enabling compatibility. Standardized use of the commands allows any SCSI device to be replaced by or operate alongside any other. Devices can be mixed and matched on the same SCSI bus according to the requirements of applications. And no matter what the configuration, all devices will operate at maximum performance if properly integrated into the system.

In the tape industry, another kind of compatibility exists: media interchange compatibility. SCSI tape devices support this capability insofar as it's available. In the VAXcluster environment, this is done via support of three popular tape formats: nine-track reel to reel, IBM 3480-compatible cartridges and 8mm cassettes.

Tape standards for recording and file formats historically have been set by volume sales of IBM products. The rest of the industry then embraces and improves upon the standard. This is the case with nine-track tape and, more recently, 3480 cartridge tape products.

Nine-track tape has been the staple medium of data processing for many years, and nine-track tape drives continue to provide some of the best, fastest and most reliable technology. Recording density has stabilized at 6,250 bpi with the group-coded recording (GCR) format, and recording speeds reach as fast as 200 ips. Additionally, file storage formats and blocking factors have been established for nine-track by ANSI, so tapes are interchangeable industrywide,

Tape Storage Manufacturers

American Data Systems Marketing Inc.
53 Elderwood Dr.
Stoughton, MA 02072
(617) 341-0171
CIRCLE 473 ON READER CARD

American Digital Systems Inc.
490 Boston Post Rd.
Sudbury, MA 01776
(508) 443-7711
CIRCLE 474 ON READER CARD

Applied Data Communications
14272 Chambers rd.
Tustin, CA 92680
(714) 731-9000
CIRCLE 475 ON READER CARD

Aviv Corp.
26 Cummings Park
Woburn, MA 01801
(617) 933-1165
CIRCLE 477 ON READER CARD

Cipher Data Products Inc.
9715 Business Park Ave.
San Diego, CA 92131
(619) 693-7084
CIRCLE 478 ON READER CARD

CMD Technology Inc.
3851 S. Main St.
Santa Ana, CA 92707
(714) 549-4422
CIRCLE 479 ON READER CARD

CMS Enhancements Inc.
1372 Valencia Ave.
Tustin, CA 92680
(714) 259-9555
CIRCLE 480 ON READER CARD

Codar Technology Inc.
1500 Kansas Ave.
Longmont, CO 80501
(303) 776-0472
CIRCLE 481 ON READER CARD

Contemporary Cybernetics Group Inc.
11830 Canon Blvd.
Newport News, VA 23606
(804) 873-0900
CIRCLE 482 ON READER CARD

Digi-Data Corp.
8580 Dorsey Run Rd.
Jessup, MD 20794
(301) 498-0200
CIRCLE 483 ON READER CARD

Digital Equipment Corp.
146 Main St.
Maynard, MA 01754
(508) 897-5111
CIRCLE 403 ON READER CARD

EMC Corp.
171 South St.
Hopkinton, MA 10748
(508) 435-2541
CIRCLE 404 ON READER CARD

Emulex Corp.
3545 Harbor Blvd.
Costa Mesa, CA 92626
(714) 662-5600
CIRCLE 405 ON READER CARD

Exabyte Corp.
1748 38th St.
Boulder, CO 80301
(303) 442-4333
CIRCLE 440 ON READER CARD

Exsys Storage Systems
1340 Tully Rd.
San Jose, CA 95122
(408) 292-0343
CIRCLE 484 ON READER CARD

First Computer Corp.
8230 S. Madison St.
Burr Ridge, IL 60521
(312) 920-1050
CIRCLE 485 ON READER CARD

Fujitsu America Inc.
3055 Orchard Dr.
San Jose, CA 95134
(408) 432-1300
CIRCLE 441 ON READER CARD

Gigatrend Inc.
2234 Rutherford Rd.
Carlsbad, CA 92008
(619) 931-9122
CIRCLE 486 ON READER CARD

Group Three Electronics Inc.
4715 Viewbridge Ave., Suite 150
San Diego, CA 92123
(619) 292-0525
CIRCLE 555 ON READER CARD

Honeywell Inc.
P.O. Box 5227
Denver, CO 80217
(303) 773-4491
CIRCLE 487 ON READER CARD

Continued.

facilitating program distribution and data sharing between systems.

The IBM 3480 18-track standard is challenging reel to reel with a much greater density (37,000 characters per inch) at a read/write speed of 79 ips. File storage formats again have been standardized by IBM, and third-party products are designed to read and write in these formats, making interchange between systems as standardized and transparent as for nine-track. Installation of both nine-track and 3480 products allows for complete transportability of media between systems without the need for special driver utilities to use the tapes.

Another kind of tape compatibility between systems results when a tape drive that hasn't been standardized suddenly finds widespread application. Such is the case with the 8mm EXB-8200 helical-scan product from Exabyte, a spin-off of Storage Technology. The extremely high-capacity cassettes (2.3 GB) used by this product can be transferred between systems implementing the drive. Helical-scan technology, based on the same hardware used in videocassette recorders, has met with strong success for backup applications since its introduction in 1987.

A proposal stating the electronic recording specifications and characteristics for the 8mm helical-scan media is currently before ANSI, and a standard is expected soon. At that time, more 8mm helical-scan drives will become available from a variety of manufacturers, but the issue of file format compatibility still will need to be resolved.

WITH WIDELY SUPPORTED STANDARDS such as SCSI, there are few risks in choosing third-party VAXcluster storage peripherals instead of the DEC offerings. These peripherals are easy to cable and integrate into the operating environment, and offer "big-company" performance and support because of collective industry backing of standards. *—Bradford T. Harrison is a free-lance writer specializing in DEC systems.*

ARTICLE INTEREST QUOTIENT
Circle On Reader Card
High 455 Medium 456 Low 457

Tape Storage Manufacturers (continued)

IBM Corp.
Old Orchard Rd.
Armonk, NY 10504
(914) 765-1900
CIRCLE 407 ON READER CARD

Innovative Data Technology
5340 Eastgate Mall
San Diego, CA 92121
(619) 587-0555
CIRCLE 488 ON READER CARD

Iomega Corp.
1821 W. 4000 S.
Roy, UT 84067
(801) 778-3000
CIRCLE 489 ON READER CARD

IPS Technology
11201 Richmond Ave.,
Ste. A102
Houston, TX 77082
(713) 870-0880
CIRCLE 536 ON READER CARD

Kennedy Co.
9292 Jeronimo Rd.
Irvine, CA 92718
(714) 770-1100
CIRCLE 537 ON READER CARD

Laser Magnetic Storage Int'l.
4425 ArrowsWest Dr.
Colorado Springs, CO 80907
(719) 593-7900
CIRCLE 538 ON READER CARD

Megatape Corp.
1041 Hamilton Rd.
Duarte, CA 91010
(818) 357-9921
CIRCLE 442 ON READER CARD

Micro Technology Inc.
1620 Miraloma Ave.
Placentia, CA 92670
(714) 632-7580
CIRCLE 443 ON READER CARD

Miltope Corp.
1770 Walt Whitman Rd.
Melville, NY 11747
(516) 420-0200
CIRCLE 539 ON READER CARD

Scientific Micro Systems Inc.
777 E. Middlefield Rd.
Mountain View, CA 94043
(415) 964-5700
CIRCLE 553 ON READER CARD

Storage Technology Corp.
2270 S. 88th St.
Louisville, CO 80028
(303) 673-5151
CIRCLE 445 ON READER CARD

Summus Computer Systems
17171 Park Row, Ste. 300
Houston, TX 77084
(713) 492-6611
CIRCLE 541 ON READER CARD

System Industries
560 Cottonwood Dr.
Milpitas, CA 95035
(408) 432-1212
CIRCLE 414 ON READER CARD

Telebyte Technology Inc.
270 E. Pulaski Rd.
Greenlawn, NY 11740
(516) 423-3232
CIRCLE 542 ON READER CARD

3M Data Products
8200 Highwood Dr.
Minneapolis, MN 55438
(800) 888-1889
CIRCLE 543 ON READER CARD

Total Tec Systems Inc.
2 Gourmet Ln.
Edison, NJ 08837
(201) 906-6500
CIRCLE 544 ON READER CARD

Transitional Technology Inc.
1411 N. Batavia Ste. 203
Orange, CA 92667
(714) 744-1030
CIRCLE 545 ON READER CARD

Trimarchi Inc.
P.O. Box 560
State College, PA 16804
(814) 234-5659
CIRCLE 546 ON READER CARD

Tristar Technology
10 Reuten Dr.
Closter, NJ 07624
(201) 784-1557
CIRCLE 547 ON READER CARD

TSA/Advet
P.O. Box 44145
4722 Campbells Run Rd.
Pittsburgh, PA 15205
(412) 787-0980
CIRCLE 554 ON READER CARD

U.S. Design Corp.
4311 Forbes Blvd.
Lanham, MD 20706
(301) 577-2880
CIRCLE 548 ON READER CARD