TECHNICAL BRIEF
March 2005

A comparison of StorageTek's Virtual Storage Manager (VSM) system's clustered VTSS and IBM's Virtual Tape Server Peer-to-Peer option

ABSTRACT

This paper explores two different approaches to virtual tape architecture. It provides a general overview of the Virtual Tape Server Peer-to-Peer (VTS PtP) capability from IBM and the clustered virtual tape storage subsystem (VTSS) functionality from StorageTek. In addition, this paper compares IBM's Peer-to-Peer (PtP) implementation with StorageTek's Virtual Storage Manager® (VSM®) system's clustered VTSS implementation. This document assumes the reader has a basic understanding of each vendor's virtual tape architecture.

1 Executive summary

There are significant differences between IBM's VTS PtP capability and StorageTek's clustered VTSS functionality.

There are significant differences between IBM's Virtual Tape Server Peer-to-Peer (VTS PtP) capability and StorageTek's clustered virtual tape storage subsystem (clustered VTSS) functionality.

StorageTek provides an integrated solution within the VSM system architecture. This is a continuation of enhanced recovery capabilities that the VSM system has delivered since its market introduction in 1998. Because IBM's VTS architecture does not allow for these types of capabilities, it requires substantially more hardware components in order to provide a similar level of functionality.

With either approach, the functionality is intended to provide greater data protection against unscheduled outages, such as component failures, as well as scheduled, or planned, outages. Planned outages allow for such tasks as maintenance of failed components, hardware upgrades and microcode upgrades, yet still provide users with access to data.

2 Overview of IBM 3494B20 Peer-to-Peer Remote Copy

The IBM 3494B20 Peer-to-Peer Remote Copy enhancement provides the ability to copy virtual volumes into two 3494 Model B20 VTS systems.

The IBM 3494B20 Peer-to-Peer Remote Copy enhancement provides users with the ability to copy virtual volumes into two 3494 Model B20 VTS systems. The capability can place a copy of virtual volumes into these two VTS systems while each resides in different 3494 tape libraries. The virtual tape servers are paired, with one being the master VTS and the other the I/O VTS. The I/O VTS may also be called a user interface (UI) library, each being a distributed library. Collectively, they are called a composite library.

It is important to understand that peer-to-peer VTS presents the image of a single logical VTS to the attached OS/390 host. In other words, the host is unaware that there are two physical VTSs connected. At the current time, a peer-to-peer VTS can emulate up to 128 virtual tape drives (VTDs), at an additional charge. In addition, there is only one set of logical volumes, for a maximum capacity of 250,000 logical volumes.

Two Magstar 3494 tape libraries, two 3494 virtual tape servers, and a minimum of four 3590 drives per 3494 tape library come together to form a peer-to-peer VTS environment.

The hardware components that form a peer-to-peer VTS environment are two Magstar 3494 tape libraries, two 3494 virtual tape servers and a minimum of four 3590 drives per 3494 tape library (six are recommended). Initially, IBM introduced two new VTS components for this support, AX0 Virtual Tape Controller and the CX0 Auxiliary Frame. The AX0 Virtual Tape Controller (VTC) is a RISC-based controller whose role is to create the peer-to-peer dual copy onto the other VTS, maintain the two VTSs in synchronization and preserve the logical volume integrity across the peer-to-peer complex. The configuration requires four VTC Virtual Tape Controllers because they provide the virtual drive emulation function in the peer-to-peer environment and each one of them is able to emulate one 3490 control unit with 16 drives. The four AX0 Virtual Tape Controllers emulate 64 total virtual tape drives (16 x 4). Users will be able to add an additional four AX0s in order to attain a total of 128 virtual tape drives.

The AX0 Virtual Tape Controllers are located between the host and the VTSs. They are connected to the host and to the VTSs with two ESCON channels from each side. This provides a total of eight ESCON channels (two from each AX0 Virtual Tape Controller) connected to the host.

The CXO Auxiliary Frame can contain two or four AXO Virtual Tape Controllers. In a local configuration, all four Virtual Tape Controllers are contained in the CXO frame. In a remote configuration, you can have two CXO frames, each one containing two AXO Virtual Tape Controllers. Each AXO frame has two power control compartments, which allow for dual power cords for redundancy.

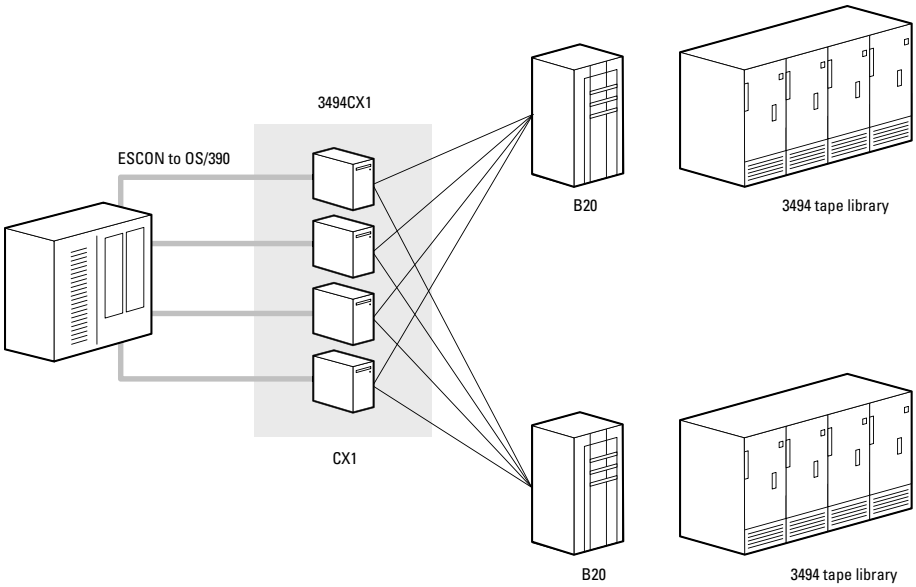
In July 2003, IBM announced a change to this feature by introducing a new product known as the 3494CX1, which replaced the AXO/CXO configuration. This simply is a repackage of the AXO/CXO with the same amount of hardware. The CX1 is used to connect two VTSs to form a PtP VTS. The original PtP used AXOs housed in CXO cabinets for this purpose. In addition to being a cabinet, the CX1 contains the virtual tape controller (VTC) functions that the old AXO provided. The major enhancement that the CX1 provides is support for FICON connections between the host and the CX1 and between the CX1 and the VTSs. The announcement states that FICON PtP performs at twice the speed of an ESCON PtP (significantly less than one would expect when using FICON connections), although a FICON VTC has a bandwidth of six times that of an ESCON VTC. Customers with existing AXOs/CXOs can upgrade the CXOs to CX1s and use two of the existing AXOs in each CX1 (the old CXO could house four AXOs). IBM also announced the ability to have two connections between a VTS and its 3494, eliminating the connection as a single point of failure. IBM calls this the dual path concentrator.

Supporting configurations are limited. Only eight configurations are supported, depending upon the VTS model.

Supported configurations are limited — only eight configurations are supported, depending upon the VTS model. A B20 VTS requires eight VTCs and only supports three configurations — all ESCON, all FICON or a 50/50 ESCON/FICON split. Those eight VTCs can be in two, three or four CX1s as each CX1 only supports two or four VTCs. A FICON VTC supports 12 combinations of long wave/short wave to the host VTS. However, all of the FICON VTCs within a single CX1 must have the same combination. In addition, for the 50/50 split, all of the VTCs within a CX1 must be either ESCON or FICON. Also note that a VTC with ESCON to the host supports only ESCON to the VTSs. There is no provision for using ESCON between the host and VTC and FICON between the VTC and the VTSs. It also appears that starting out with an ESCON PtP and then adding FICON to the PtP will be a cumbersome and expensive process, as the system administrator will have to upgrade a minimum of four VTCs at the same time, as well as adding the FICON feature and FICON channels to both VTSs.

The host software components are an OS/390 operating system with DFSMS/MVS. DFSMS/MVS needs a Small Program Enhancement in order to recognize the composite library. A new IBM Copy Feature (No. 4010) is required for the B20 Virtual Tape Server in a peer-to-peer configuration. As a prerequisite for this Copy Feature, two VTSs in the peer-to-peer configuration must be installed in two different 3494 libraries, Extended High Performance Option and Performance Accelerator Feature must be installed in both VTSs, and four Extended Performance ESCON/FICON channels are required.

Figure 1. Overview of IBM PtP architecture



As previously mentioned, one VTS is the master while the other is the I/O VTS. Host tape mounts are issued through the VTC Virtual Tape Controller, and the Virtual Tape Controller selects the logical volume and logical tape drive. It then selects which distributed library will satisfy the mount. Either of the two distributed libraries may be assigned to handle any scratch mount. The distributed library that satisfies the scratch mount will be responsible for all subsequent references to that logical volume. An additional 3494CX1 can be installed at a remote site where the VTSs are physically separated to provide disaster backup.

When data has been received into the VTC, it is routed to the master VTS. The master VTS compresses the volume and then one of two copy options is available. One is Immediate Copy mode, which will make a copy of the logical volume on the alternate VTS. The second mode is Deferred Copy. As the name implies, the logical volume is added to a copy queue and some time later the VTC copies the volume to the second VTS.

3 Overview of clustered VTSS

Clustered VTSS from StorageTek provides increased data availability to users of the VSM system, beyond what dual or quad ACS delivers today.

Clustered VTSS from StorageTek provides increased data availability to users of the VSM system, beyond what a dual or quad automated cartridge system (ACS) delivers today. This enhancement enables customers to replicate virtual tape volumes (VTVs) from one VTSS to another VTSS. The two VTSSs operate together to provide improved data availability without host involvement. The two VTSSs can either be attached to one ACS at a local location or two ACSs, one being local and the other remote. The VTV replication is started automatically at VTV dismount time. Clustered VTSS can be utilized to provide automatic remote copying of a VTV without incurring the overhead of managing physical media (mounting and positioning of multiple volume cartridges), which is currently the case with dual ACS.

The system administrator designates a pair of VTSSs that form a cluster. One VTSS is the primary system and the other is configured as the secondary one.

The two VTSSs are connected via ESCON channels, which are uni-directional. (FICON channel connectivity is scheduled to be available 2005.) Data can flow from the primary VTSS to the secondary. Under normal conditions, the primary VTSS will be responsible for supporting the tape workload. The primary VTSS will then replicate VTVs, using the ESCON channel connection between the primary and secondary VTSSs.

The system administrator can define which data should be replicated through an advanced management policy class known as replicate. As VTVs are dismounted on the primary, with the replicate policy enabled, VTCS will send a copy of the VTV to the secondary VTSS. Both copies of the VTV are resident in the VTSS buffers. When a VTV is removed from the primary VTSS (migration/delete-on-scratch), then the copy is also removed from the secondary. Migration can occur from both the primary and secondary VTSS, but is typically done by the secondary to maximize primary replication efficiency. As soon as VTVs are replicated to the secondary VTSS, they will be queued for immediate migration. The advantage occurs when the primary VTSS must auto-migrate, resulting in trivial overhead.

If the primary VTSS becomes unavailable, the user can vary the primary VTSS offline, allowing the workload to continue on the secondary VTSS. Once the primary is available, the user can vary the primary VTSS online and return the workload to the primary VTSS. When the primary is restored back online, VTCS will automatically re-synchronize the contents of the two buffers.

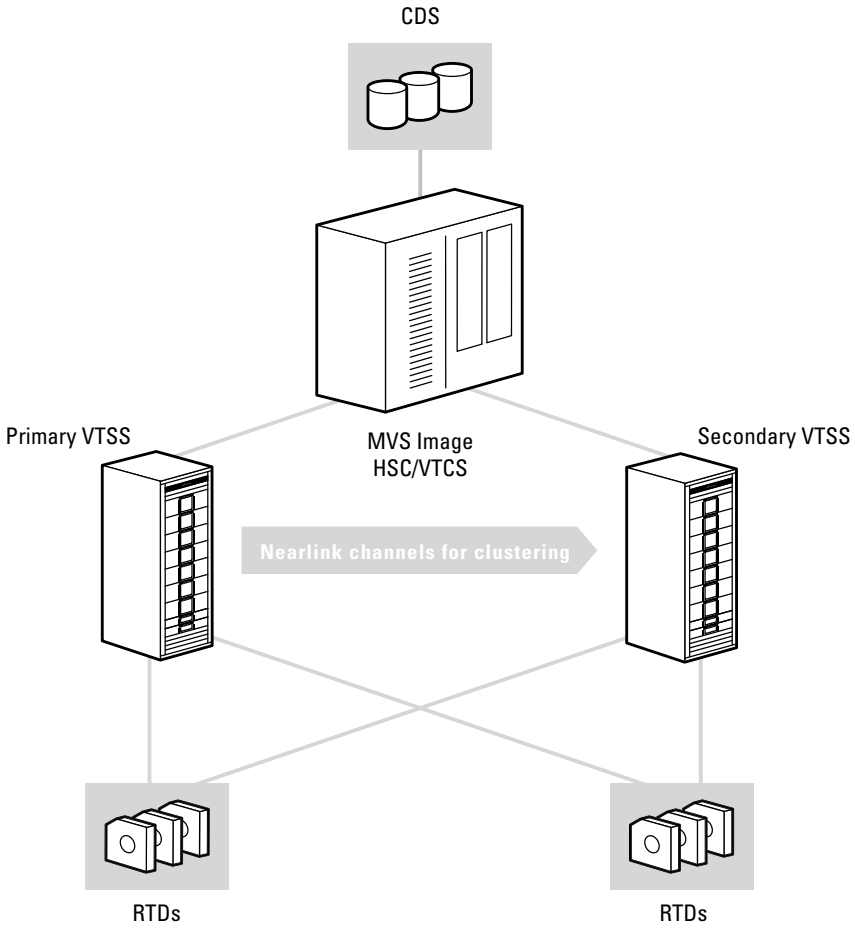


Figure 2. A clustered VTSS solution with two VTSSs sharing the same real tape drives

The secondary VTSS may be at the same location as the primary or it may be located remotely. The cluster may be connected to two separate ACS IDs to implement dual-ACS support in addition to clustering. Alternatively, the cluster may be connected to a single ACS. The two VTSSs can share the same real tape drives or for improved performance may be connected to dedicated real tape drives. The two VTSSs must connect to the same device types on the same ACS in order for both VTSSs to have access to all multiple volume cartridge copies for any VTV.

Nearline Control Solution (NCS) software 6.0 or above will be required in order to support clustered VTSS. A new hardware feature is required on both the primary and the secondary VTSS, which adds clustered support (StorageTek No. CLST). The prerequisite for this feature will be IMPX, which provides advanced management capability. In addition, the VTSSs in the cluster must be VSM3s. The total number of real tape drives that are library-attached to the two clustered VTSSs can be 30. Each VTSS can support 16 real tape drives. The number of connections used going from the primary to the secondary VTSS reduces this number.

With the introduction of NCS 6.1, clustered VTSS will allow users to create bi-directional clusters.

With the introduction of NCS 6.1, clustered VTSS includes the capability to provide bi-directional clusters, which consist of two peer VTSSs connected by one or more cluster links. In a bi-directional cluster, a system administrator can use the MGMTclas statement's REPLICAT parameter (which requires the advanced management feature) to allow replication in either direction from peer to peer across a cluster link. Each peer VTSS, therefore, acts as a "warm standby" to the other, and each can accept production work as well as send and receive VTVs.

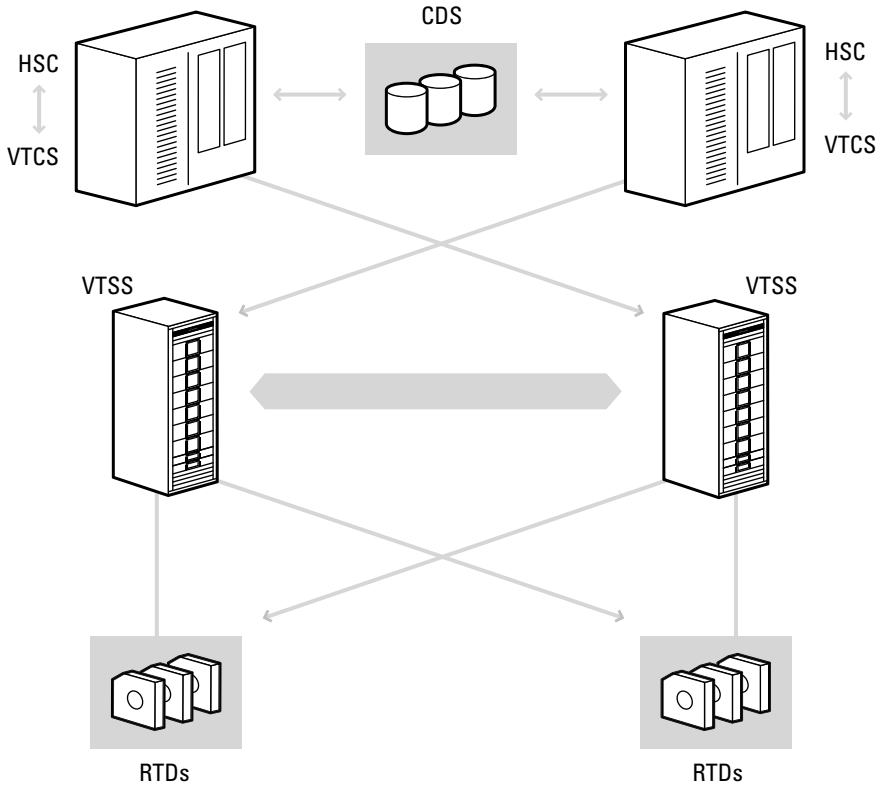


Figure 3. Overview of the bi-directional VSM clustered VTSS architecture

4 How does VTS peer-to-peer compare to clustered VTSS?

Clustered VTSS adds the further replication of VTVs to a secondary VTSS, which reduces the time required to restore the contents of a failed VTSS buffer.

IBM has introduced this peer-to-peer architecture in order to solve two fundamental requirements. The first requirement is to protect data against media failures. This allows a VTS to have a virtual tape volume residing on two separate pieces of media. StorageTek's VSM system already provides this capability with virtual tape volume (VTV) duplexing and or quadplexing.

The second requirement is the ability to place a second copy at a remote location. The VSM system provides that capability in the form of dual/quad ACS. Dual ACS was introduced with the VSM system in 1999 without adding additional VSM hardware and expense. Quad ACS provides for the duplication of virtual tape volumes to the physical media in more than one location — up to four locations.

Clustered VTSS expands upon this concept by adding the further replication of VTVs to a secondary VTSS, which reduces the time required to restore the contents of a failed VTSS buffer. All that is needed is a second VSM system (VTSS).

4.1 Investment needed for VTS

IBM's VTS system requires a significant investment in additional hardware.

IBM's VTS system, on the other hand, requires a significant investment in additional hardware. A typical peer-to-peer VTS configuration with 128 VTDs would have the following cost ($US):

- Virtual Tape Controller (ESCON): $100,000 ($25,000 x 4)
- Virtual Tape Controller (FICON): $240,000 ($60,000 x 4)
- Auxiliary Frame (CXO): $20,000
- Peer-to-Peer Base (B20 feature) $40,000 ($20,000 x 2)
- Peer-to-Peer Incremental (B20 feature) $20,000 ($10,000 x 2)

Total of $180,000 for ESCON

Total of $320,000 for FICON

If a company wishes to implement the 128 VTD support, the cost of this system would increase by $120,000 for either FICON or ESCON.

This is without considering the cost of the other hardware and software items that may be needed prior to implementing the peer-to-peer solution. Those costs could include:

- B20 Extended Higher Performance Option $118,120 ($59,060 x 2)
- B20 Performance Accelerator $200,000 ($100,000 x 2)
- Potentially more disk capacity in the B20
- An additional 3494 B20
- An additional 3494 library with 3590s
- Increased environmental requirements (floor space, power, etc.)

4.2 Investment needed for clustered VTSS

Compare the above solution price to that for clustered VTSS:

- VSM system's clustered feature $100,000 ($50,000 x 2)

Total $100,000

Other hardware and software items that may be needed pre-requirements for a clustered VTSS implementation include:

- Advanced Management Capability $50,000 ($25,000 x 2)
- An additional VSM system (VTSS)

In this comparison, the bottom line shows a significant cost differential. The IBM solution costs approximately $180,000 to $320,000, while the StorageTek solution costs about $150,000.

4.3 Other differences between the two systems

*Clustered VTSS allows the replication of data at a dataset by dataset level and can be changed **dynamically**.*

Other differences between the two systems include the following:

- Clustered VTSS allows the replication of data at a dataset by dataset level and can be changed dynamically. IBM provides "immediate copy" or "deferred copy" options. It should be understood that Immediate Copy mode is a request, that might not execute due to a number of factors. If there is a component failure, the Immediate Copy mode may operate as if in Deferred Copy mode until the failed component(s) are brought back online. This situation may occur when one of the Model B20 VTSs is not operational, links to a B20 are lost, or the tape library associated with a VTS is not operational. VSM system allows an ACS to be shut down while work continues. Only the shortage of a resource (e.g., recalling a VTV from a multiple volume cartridge) would cause an impact. Lack of access to an ACS is not a failure by default to a VSM system.
- Each VTC emulates 32 IBM 3490Es and has two ESCON interfaces to the host. This means that there can be at most two concurrent transfers to those 16 drives. The user will have to balance workload across the four VTCs. Clustered VTSS does not have this limitation.
- Each VTC has a single ESCON/FICON connection to each B20. This has significant performance implications, as both host interfaces now go across a single connection to each B20. Clustered VTSS does not have this limitation.
- If a primary VTSS becomes unavailable, the system administrator can vary the primary VTSS offline with work continuing on the secondary VTSS. Once the primary is available, then the administrator can vary the primary VTSS online, and workload will be shifted to the primary VTSS. If a VTS peer-to-peer site (VTCs, VTS and 3494s) suffers an outage and can no longer be used, the second VTS is rendered unusable until an IBM service representative has placed the surviving VTS into a "read/write disconnect mode." This allows the operations to continue with one VTS and a minimum of one VTC.
- Import/export functionality with IBM's VTS in peer-to-peer mode is lost and cannot be used. The VSM system, on the other hand, maintains its export/import capability.
- Should the master VTS require additional space in IBM's peer-to-peer configuration, a virtual tape volume must be migrated, recalled and then copied to the remote site with the secondary VTS in deferred mode. Since VSM system replicates a VTV first and has the capability to migrate from the secondary VTSS, this is much less likely to occur.

4.4 Comparison summary

The following matrix highlights some of the key differences between clustered VTSS and Peer-to-Peer implementations:

	Clustered VTSS	**Peer-to-Peer**
Number of virtual drives	64/256 per VTSS pair	64/128
Maximum number of real tape drives available	30	24
Minimum number of back-end tape libraries (ACSs)	1	2
Back-end library support	9740, 9360, 4410, 9310 or SL8500	3494
Back-end drive support	9490, 9840, T9940, in any combination	3590, 3592
Minimum additional hardware required in order to implement within an existing virtual tape environment	1 VSM system (VTSS)	1 B20 VTS 1 3494 library 4 3590 drives (6 recommended) 4 VTC Virtual Tape Controllers 1 CXO Auxiliary Frame
Billable virtual tape feature pre-requirements that need to be enabled prior to implementation	IMPX Feature = Import Export/Advanced Management Policy feature	3400 Feature = Extended High Performance Option 5236 Feature = Performance Accelerator Two 3412 = Extended Performance ESCON Channels One or more of features 4010, 4011, 4012, 4013 = Peer-to-Peer Copy
Maximum number of virtual tape volumes	Unlimited	250,000
Dual library support	Yes, dual/quad ACS	No, one per B20
SCSI host support	No	No
Ability to support export/import	Yes	No

5 Benefits of clustered VTSS

StorageTek's clustered VTSS provides advantages at a lower cost than some alternatives and is an integrated solution within the VSM architecture.

There are notable differences between the clustered VTSS implementation from StorageTek and the VTS Peer-to-Peer option from IBM.

StorageTek provides an integrated solution within the VSM architecture. This is a continuation of enhanced recovery capabilities that VSM system has delivered since its market introduction in 1998 — for example, VTV duplex and dual/quad ACS. Since IBM's VTS architecture does not allow for these types of capabilities, substantially more hardware components are required in order to provide similar functionality.

Ultimately, clustered VTSS provides certain advantages and does so at a lower cost than some alternatives. With IBM, you can have either PtP or import/export, but not both. With VSM, you have a wide range of options to choose from.

ABOUT STORAGETEK®

Storage Technology Corporation (NYSE: STK) is a $2 billion global company that enables businesses, through its information lifecycle management strategy, to align the cost of storage with the value of information. The company's innovative storage solutions manage the complexity and growth of information, lower costs, improve efficiency and protect investments. For more information, visit www.storagetek.com, or call 1.800.275.4785 or 01.303.673.2800.

WORLD HEADQUARTERS

Storage Technology Corporation
One StorageTek Drive
Louisville, Colorado 80028 USA
1.800.877.9220 or 01.303.673.5151

ET 0001 A 03/05