An Oracle Technical White Paper
December 2011

StorageTek T10000C Tape Drive FICON Performance

Table of Contents

Table of Figures

Table of Illustrations

Table of Tables

Introduction

Oracle's StorageTek T10000 FICON tape drive family offers Oracle customers industry-leading performance and capacity. The high performance of the T10000 reduces wall-clock time of batch production and backup applications while the increased capacity lowers the cost of storage and saves library slots. The StorageTek T10000C FICON drive is the third generation in the StorageTek T10000 FICON tape drive family and includes increased cartridge capacity to five terabytes (5TB), 4GB Fibre Channel and FICON interface, increased head-to-tape data transfer rate of 240 MB/s, and backwards read compatibility with the StorageTek T10000A and StorageTek T10000B tape drives. The first and second generations in the T10000 FICON tape drive family are referred to as the StorageTek T10000A and StorageTek T10000B tape drives, respectively.

Figure 1 - T10000 Tape Drive

Scope

This paper specifically describes the StorageTek T10000C performance characteristics when directly connected to a mainframe channel. Note that this paper does not include the performance of the drive when used in conjunction with Oracle's StorageTek Virtual Storage Manager (VSM).

Disclaimers

All tests were run with encryption enabled and encryption disabled. Where the included illustrations do not show plots for both, the results were sufficiently similar to simply show one line.

Due to variations in environments, applications, and microcode levels, customers may observe slightly different results in their own data centers.

Measurement Environment

To obtain accurate performance data, all performance measurements were conducted on a stand-alone z10 model 2096-X02 processor running z/OS version 1.11 equipped with 8Gb/s FICON Express channels. For the multiple drive tests, a Brocade SilkWorm 48000 FICON switch was used to attach all the StorageTek T10000Cs to a single FICON CHPID.

The microcode used in this testing was 1.53.311.

Tape Drive Description

The StorageTek T10000 uses a unique, single-reel cartridge. The file reel is located inside the cartridge while the machine reel resides inside the tape drive. The drive uses a technology called partial response, maximum likelihood (PRML) to provide a high-density data format. PRML enables recording and storing an uncompressed capacity of up to:

- 500 gigabytes (GB) with the StorageTek T10000A tape drive
- 1 terabyte (TB) with the StorageTek T10000B tape drive
- 5 terabytes (TB) with the StorageTek T10000C tape drive and StorageTek T10000C media

A StorageTek T10000A drive can read and reclaim a tape cartridge written by a StorageTek T10000A drive.

A StorageTek T10000B drive can:

- Read and reclaim a tape cartridge written by a StorageTek T10000A drive
- Write, read, and reclaim a tape cartridge written by a StorageTek T10000B drive

A StorageTek T10000C drive can:

- Read tape cartridges written by either a StorageTek T10000A or StorageTek T10000B drive
- Write, read, and reclaim a tape cartridge written by a StorageTek T10000C drive

The tape drive uses fiber optic host connections to provide a high data-transfer rate

StorageTek T10000 Cartridges

The StorageTek T10000 supports five types of cartridges:

- StorageTek T10000 T1 cartridge (T10000A/B drive):
  - Data: 500 gigabytes StorageTek T10000A or 1 terabyte StorageTek T10000B
  - Data, sport: 120 gigabytes StorageTek T10000A or 240 gigabytes StorageTek T10000B
  - Oracle's StorageTek VolSafe software, capacity: 500 gigabytes StorageTek T10000A or 1 terabyte StorageTek T10000B
  - StorageTek VolSafe, sport: 120 gigabytes StorageTek T10000A or 240 gigabytes StorageTek T10000B
  - Cleaning cartridge: 50 uses (CT or CL cartridge)
- StorageTek T10000 T2 cartridge (StorageTek T10000C tape drive):
  - Data, standard: 5 terabytes
  - Data, sport: 1 terabyte

- StorageTek VolSafe, capacity: 5 terabytes
- StorageTek VolSafe, sport: 1 terabyte
- Cleaning cartridge: 50 uses (CC or CL cartridge)

FICON Architecture Implementation

StorageTek T10000 FICON tape drives implement native FICON direct attach interfaces, often referred to as "1x1" architecture. This means each drive, such as the StorageTek T10000C FICON, can be connected directly to a FICON channel or FICON director/switch without intervening equipment (see Illustration **1**).

Oracle allows any number of drives to be attached to the FICON channel, within the architectural limits of FICON. This flexibility gives Oracle customers the greatest performance and availability while minimizing costs.

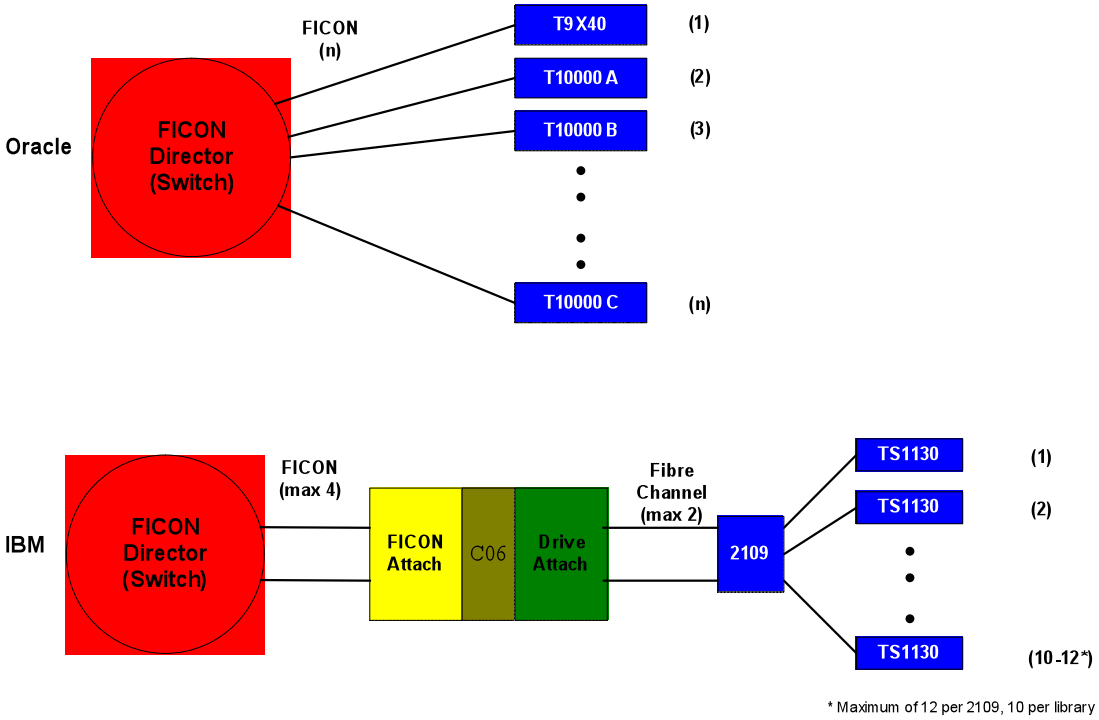


Illustration 1 - Oracle versus IBM FICON Implementation

Compare this with a shared control unit based architecture, which contains multiple protocol converters that introduce configuration limitations. The IBM solution introduces performance bottlenecks by allowing only four FICON channels into a solitary controller with only two internal fibre channel paths, and a limited number of IBM TS1120 and IBM TS1130 drives. All this added

hardware significantly increases the total cost of ownership and introduces failure points, which would make all the drives in the subsystem unavailable.

The Oracle 1x1 architecture has been the architecture preferred by customers since it was introduced with ESCON several years ago. With FICON, the 1x1 architecture continues to be the best choice for numerous reasons:

- Greatest flexibility in channel and subsystem configurations
- Greatest performance which reduces backup and recovery times
- Highest reliability since there is less equipment and no single point of failure to multiple drives
- Lower cost of ownership for the same throughput

StorageTek T10000C Performance

In this section, we will look at the performance of the StorageTek T10000C FICON drive in detail.

Block Size

Illustrations **2** and **3** show the read and write data transfer rates of the StorageTek T10000C using a queued sequential access method (QSAM) application with 4:1 data compression on a 2Gb channel connection. The graph shows read/write1 throughput measurements over a range of block sizes from 4 KiB to 256 KiB. As is the case with all I/O devices, we see that block size plays a large role in effective throughput.

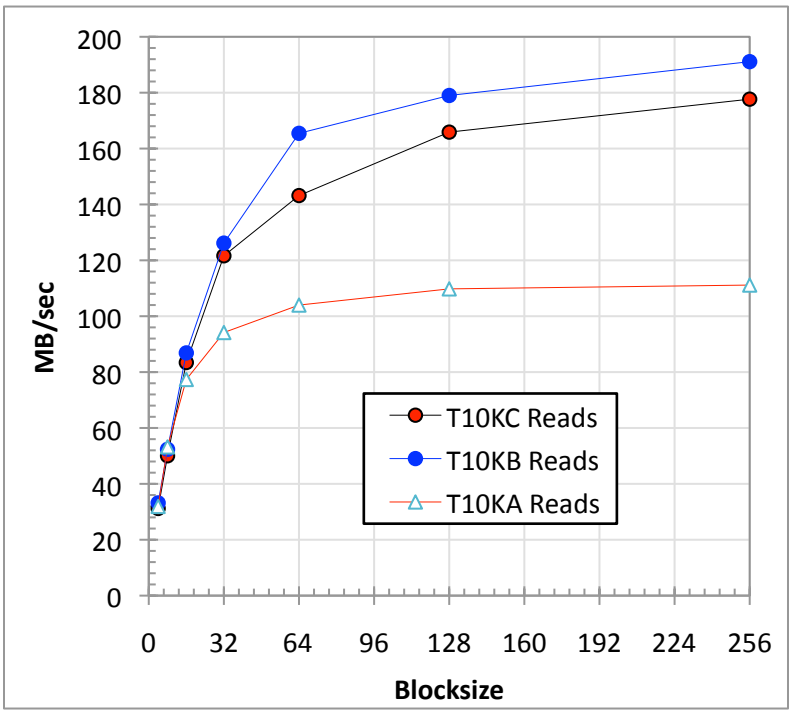


Illustration 2 - QSAM Read Transfer Rate, 4:1 Compression

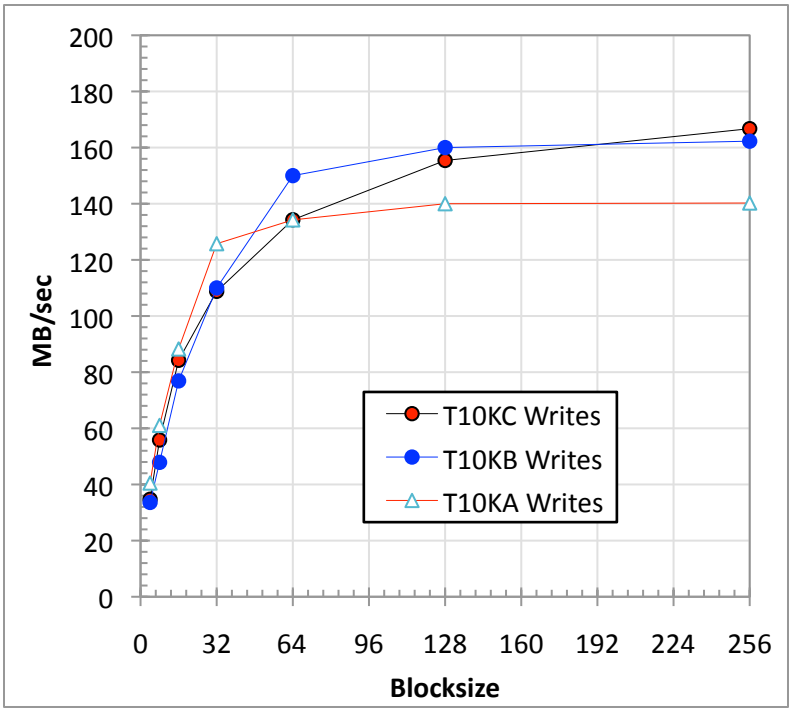


Illustration 3 - QSAM Write Transfer Rate, 4:1 Compression

Illustrations **2** and **3** can also be used to estimate the performance of applications based on their block sizes. Table **1** lists some applications and their respective block sizes. Note that maximum performance can be achieved using large block applications such as Oracle's Expert High Performance Data Mover (ExHPDM) utility.

APPLICATION	BLOCK SIZE
DFHSM	16 KiB
FDR	64 KiB
ExHPDM	256 KiB

Table 1 - Application Block Sizes

Illustration **4** is similar to Illustrations **2** and **3** except a 4 Gb FICON channel was used. When an execute channel program (EXCP) application was used, we obtained the results shown in Illustration **5**. Note: the differences between the 2 Gb and 4 Gb channel data were negligible.

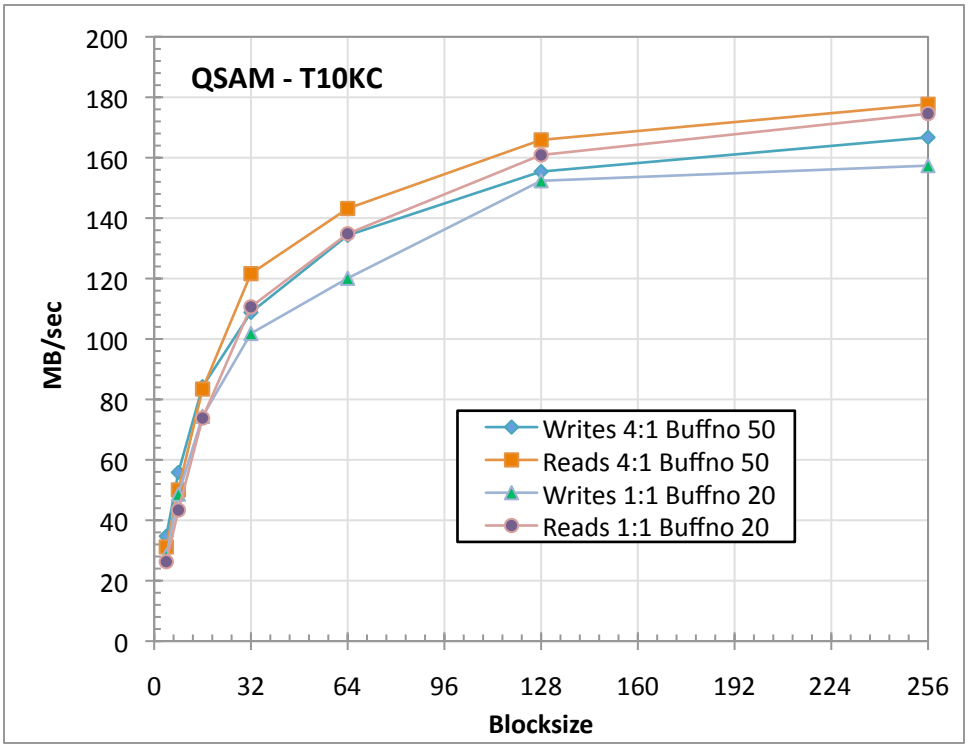


Illustration 4 - QSAM 4 Gb Channel, Block Size Chart

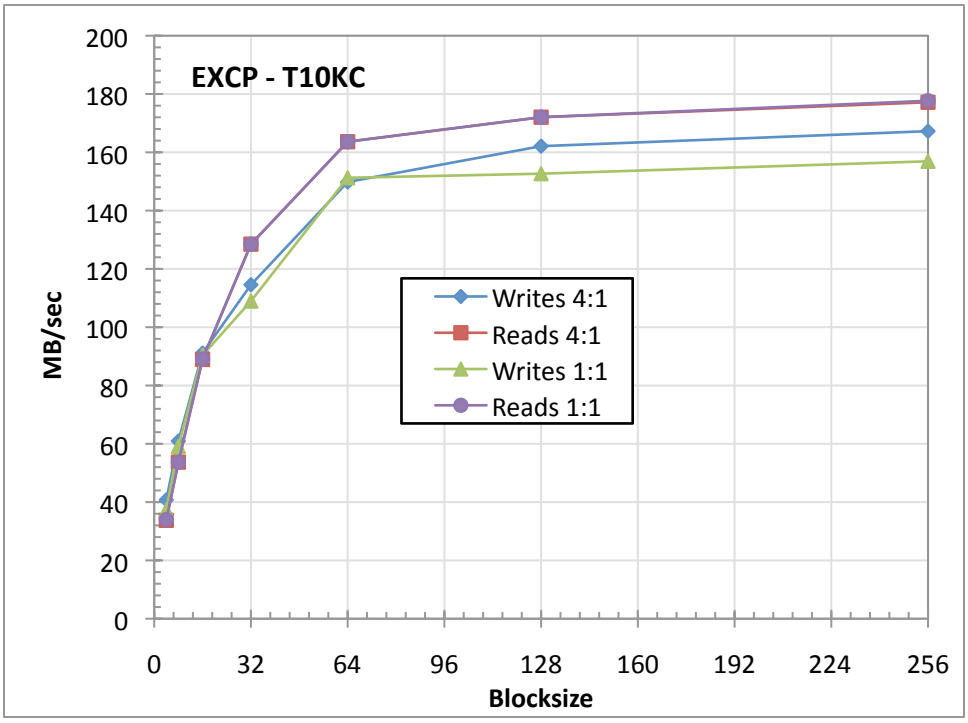


Illustration 5 - EXCP Block Size Chart

Chain Length

Illustration **6** shows the effects of command chaining length on throughput. In this chart, an EXCP application was used with 4:1 data compression and a block size of 32 KiB and 16 KiB. The chaining lengths are the number of read or write commands chained together in a single channel program, i.e., the number of blocks to be transferred in a single I/O operation. Adding more read or write commands to a channel program allows more blocks of data to be transferred per START SUBSHCNNEL instruction. This makes the I/O operation more efficient for the host processor, the channel subsystem, and the I/O device.

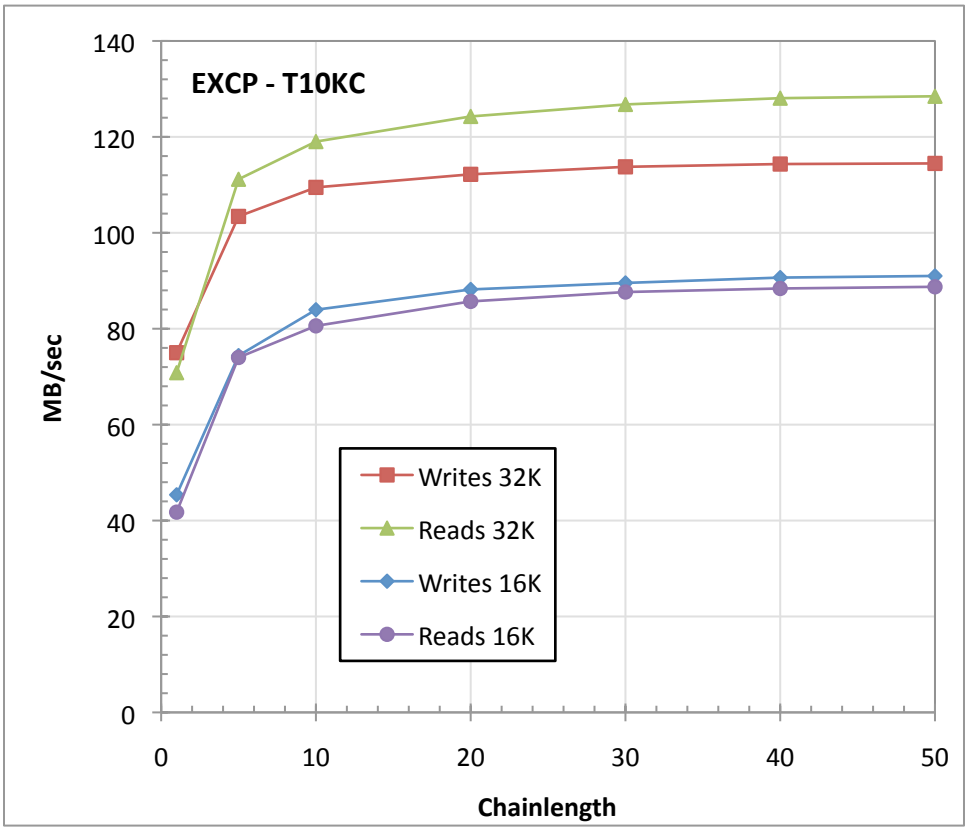


Illustration 6 - EXCP 2GB Chain Length with 4:1 Compression

Chain length can be easily controlled in many access methods by coding DCB=BUFNO=nnn in the DD statements of the application's JCL. For example, QSAM will, on average, transfer BUFNO/2 blocks per I/O operation from the application. The QSAM application was used in Illustration **7**.

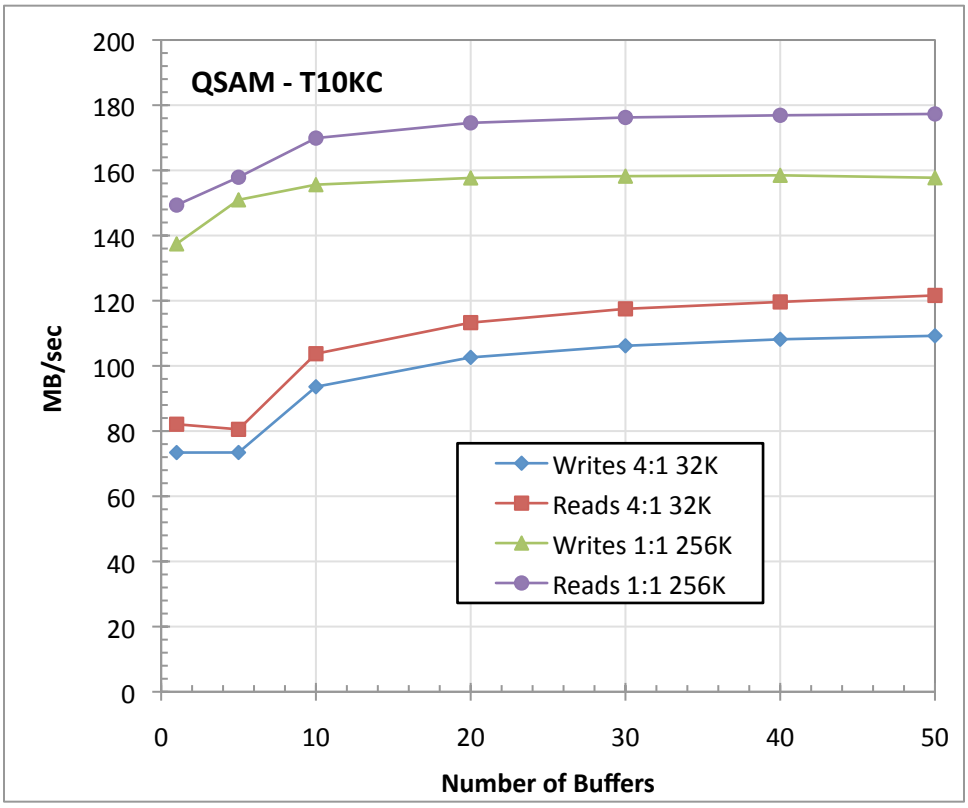


Illustration 7 - QSAM 2 GB Chain Length with 32700 Block Size

Note that larger chain lengths require more real storage and can monopolize the channel for longer periods of time. Therefore, it is prudent to find a chaining length that provides the highest performance with the fewest side-effects.

Compression

Data compression not only allows for greater effective storage capacities, but also improves throughput. By measuring performance at many different data compression ratios, we can see how this affects performance in Illustration **8**.

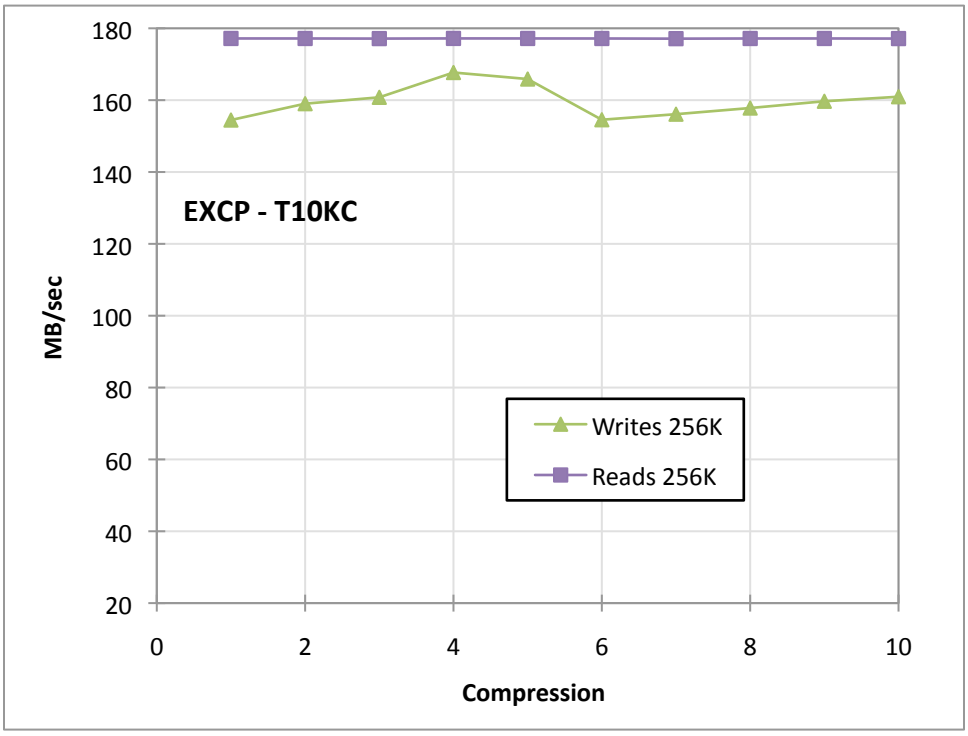


Illustration 8 - EXCP 4GB Compression Chart

Distance

We can see how distance affects the StorageTek T10000C in Illustrations **9** and **10**. Each line on the graph plots performance against blocksize for a range of distances from zero to 30KM. For this series of experiments, performance of the StorageTek T10000C was measured using an ANUE H Series V2 distance emulator.

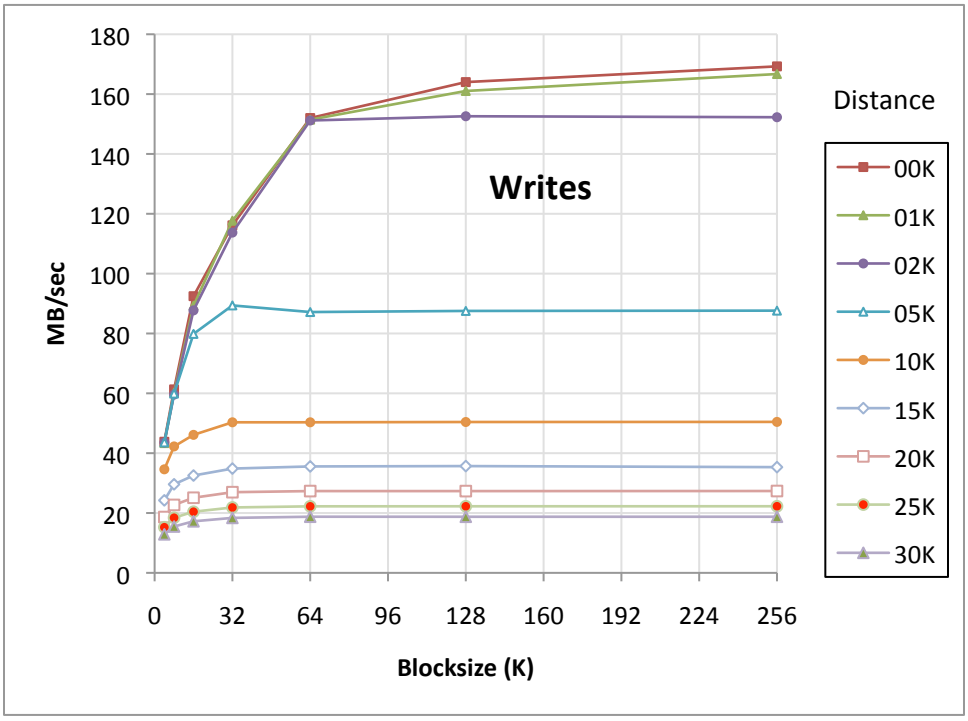


Illustration 9 - Affect of Distance on Write Performance

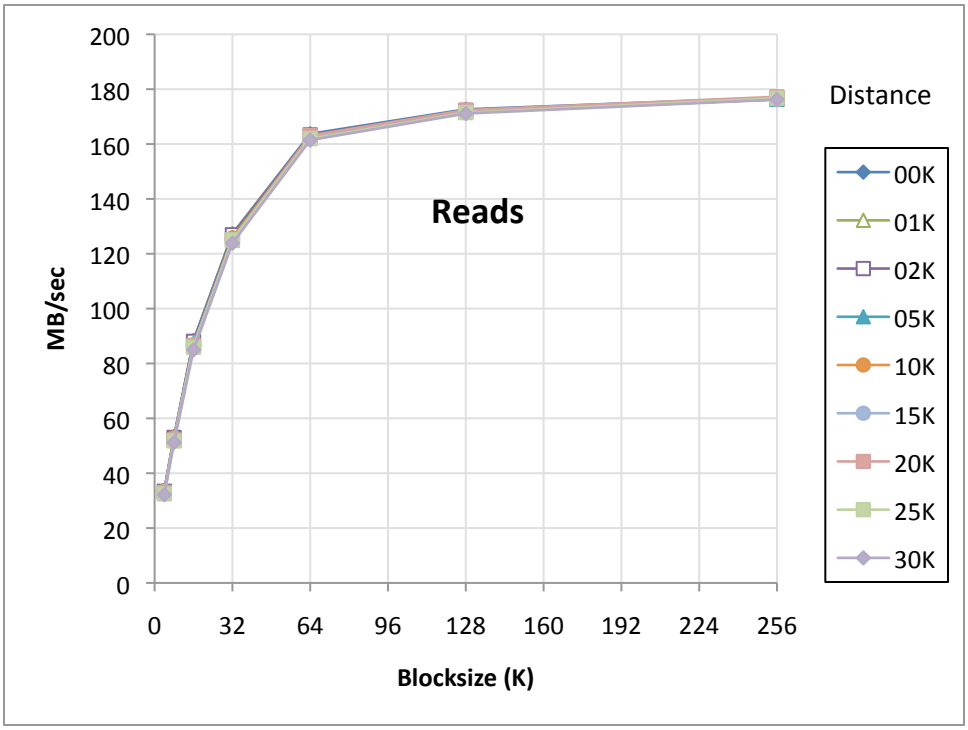


Illustration 10 - Affect of Distance on Read Performance

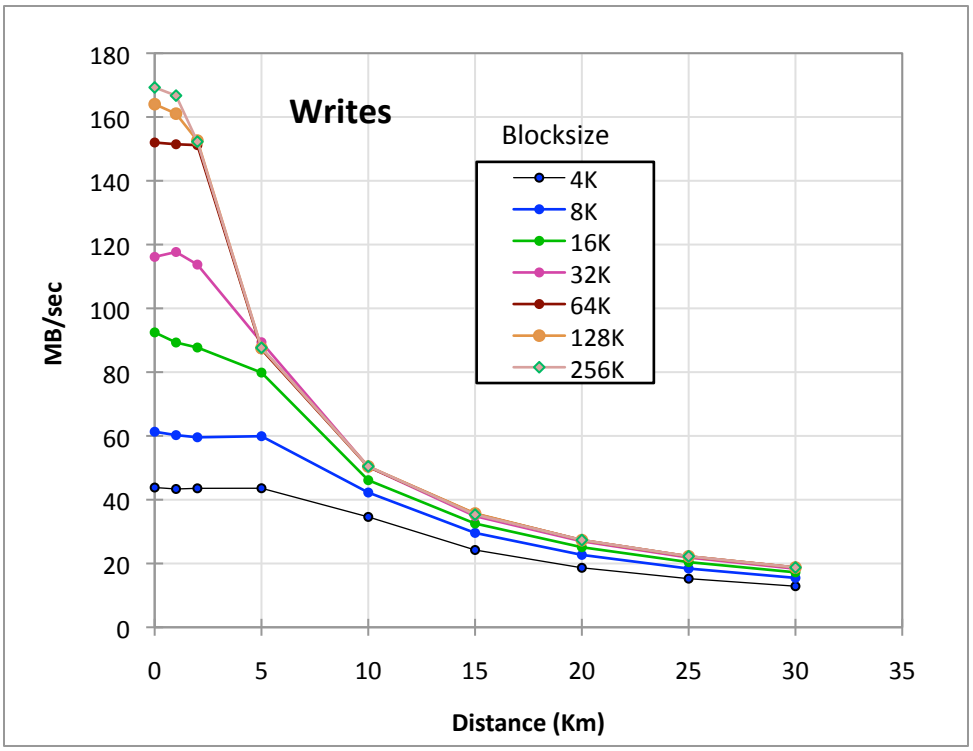
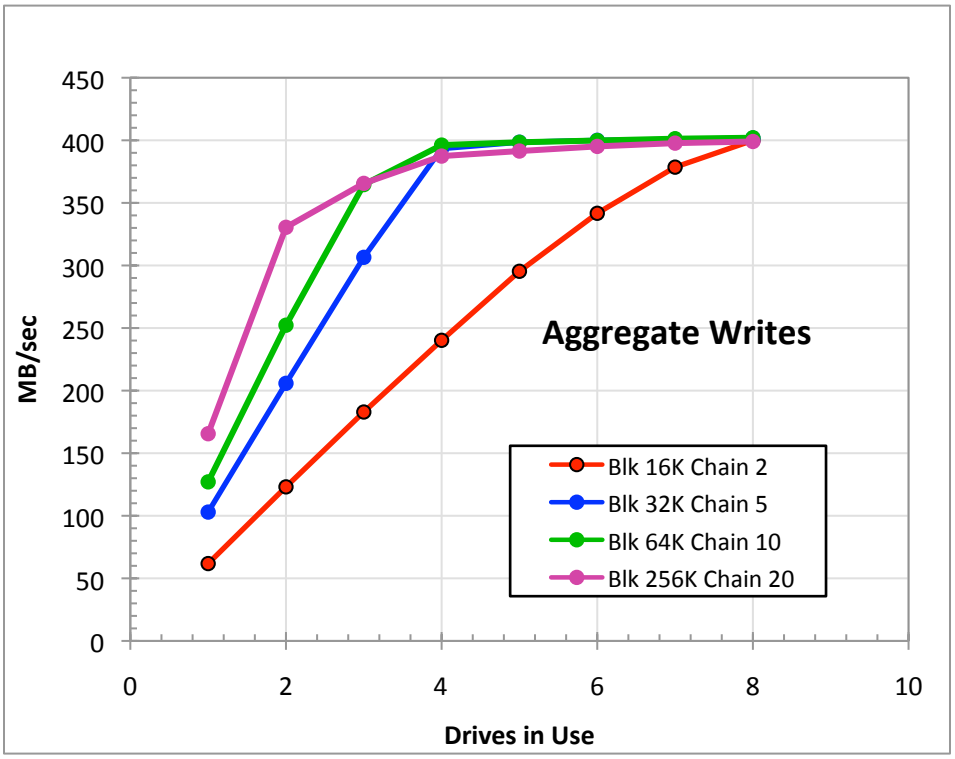


Illustration 11 - Affect on Performance by Block Size Over Distance

Channel Loading

To determine how many StorageTek T10000s perform on a single FICON channel, we conducted a series of experiments which produced the results shown in Illustration 12 for the StorageTek T10000C. For these experiments, we measured performance of the drives by writing 10GB of data to one drive, then 10GB of data to two drives, then 10GB of data to three drives, and so on as defined by the graph legend, using the block sizes shown on the x-axis of the graph. We repeated the measurements many times, with each iteration changing the number of drives concurrently active (each running the same measurement on a single 4 Gb/s FICON channel). We see that the channel becomes saturated at smaller and smaller blocksize as more StorageTek T10000Cs are made concurrently active and becomes completed saturated at three drives when the blocksize reaches 64K. This implies that it is pointless to connect more than 3 drives per channel if performance is the goal.



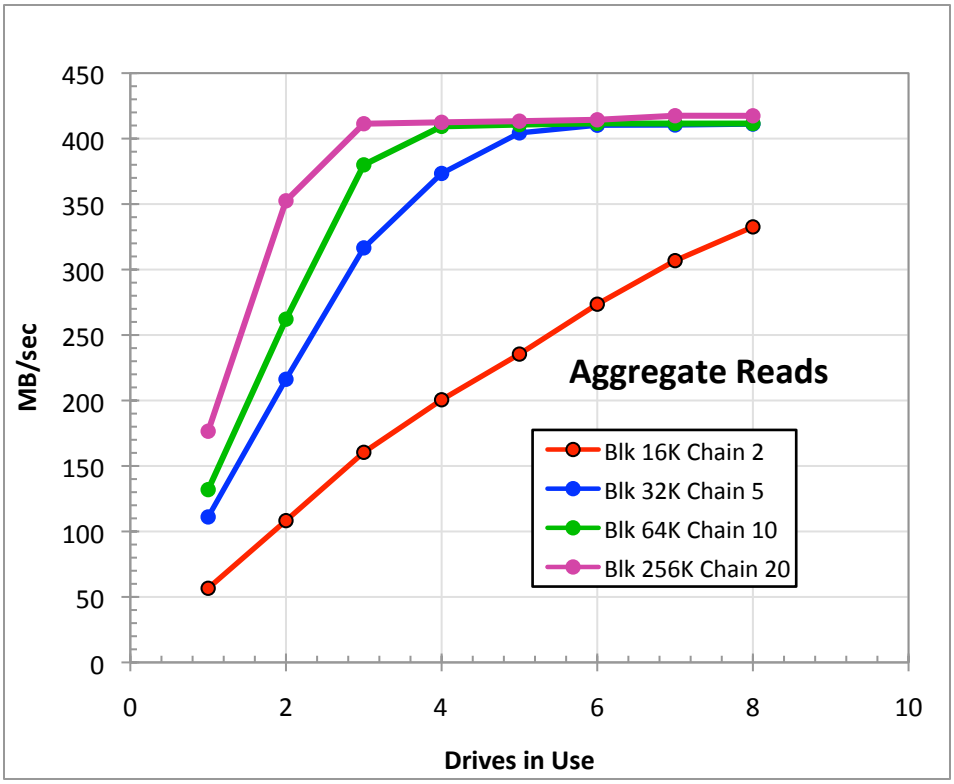


Illustration 12 - Affects of Channel Loading

While beyond the scope of this paper, Resource Measurement Facility (RMF) reports for the channel can be viewed to determine the nature of the saturation. Typically, many of the drives transferring small blocks and short chain lengths will saturate the microprocessor within the channel. Multiple drives performing large data transfers will typically saturate the I/O bus within the channel.

Comparisons

Like Oracle's StorageTek T9840D tape drive, the StorageTek T10000C has a similar FICON interface, but the native head-to-tape speed of the StorageTek T10000C is much faster than the StorageTek T9840D as shown in Illustrations **13** and **14**.

Unlike the dual-hub access-centric design of the StorageTek T9840 series drives, the StorageTek T10000 series utilizes a single-hub storage-centric design. This allows the StorageTek T10000C to store up to 5000 GB of uncompressed data, sixty-five times that of the StorageTek T9840D.

While the single hub cartridge of the StorageTek T10000 media allows more tape to be stored and greater capacity, it is not mid-point load/unload capable like the StorageTek T9840 media. Therefore, the access and rewinds times on the StorageTek T10000C are longer than the StorageTek T9840D. A comparison of access times is shown in Table **2**. Miscellaneous specifications are shown in Table **4**.

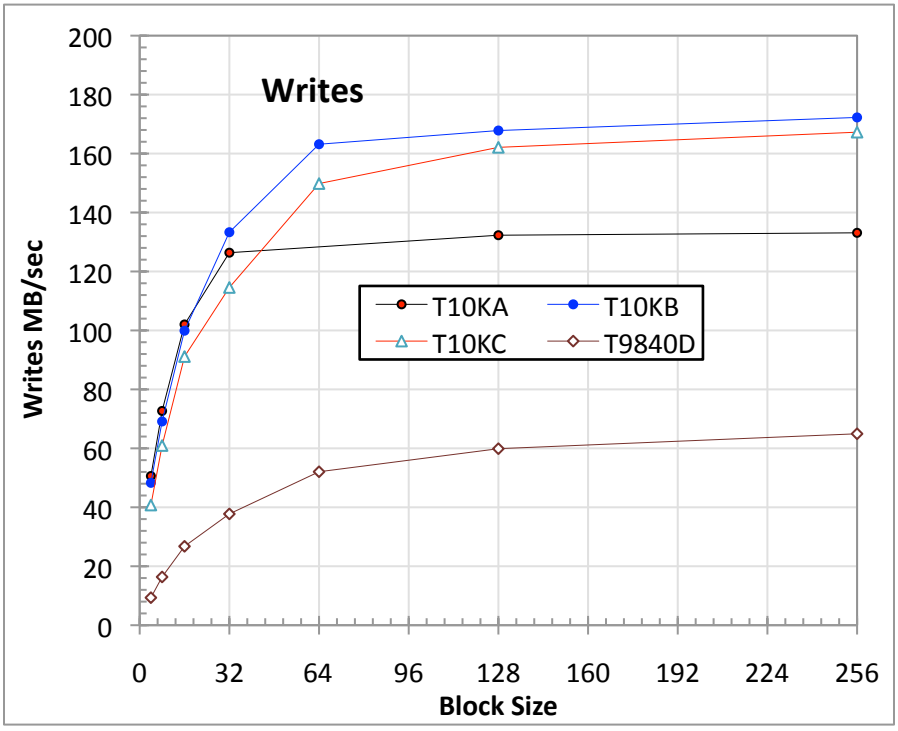


Illustration 13 - Write Performance Comparison

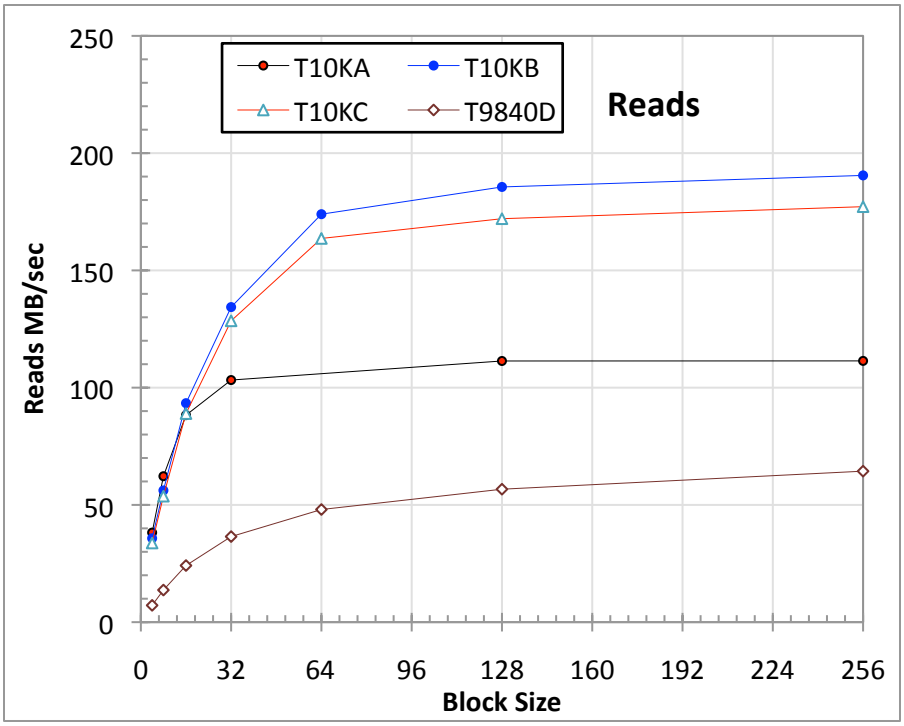


Illustration 14 - Read Performance Comparison

Metric	StorageTek T10000C		StorageTek T10000B		StorageTek T10000A		StorageTek T9840D
	Std	Sport	Std	Sport	Std	Sport	9840D native
Load/Thread to Ready (s)	13.1	13.1	16.5	16.5	16.5	16.5	8.5
Avg Search Time (to mid tape) (s)	57	17.5	46	14	46	14	8
Max Search Time (to EOT) (s)	115	32.5	91	23	91	23	16.5
Avg Rewind Time (from mid tape) (s)	57	17.5	48	13	48	13	8
Max Rewind Time (from EOT) (s)	115	32.5	91	23	91	23	16
Unload Time (s)	23		23		23		12.5

Table 2 - Access Time Comparison

Metric	StorageTek T10000C		StorageTek T10000B		StorageTek T10000A		StorageTek T10000A
	Std	Sport	Std	Sport	Std	Sport	9840
Recording Format	Linear Serpentine		Linear Serpentine		Linear Serpentine		Linear Serpentine
Capacity							
Uncompressed Native Capacity (GB)	5000	1000	1000	240	500	120	75
Tape Length (m)	1147	334	917	267	917	267	271
Number of Tracks	3584		1152		768		576
Recording Channels	32		32		32		16
Wraps	112		36		24		36
Areal Density (Mb/inch^2)	3140		800		403		114
Linear Bit Density (kb/inch^2)	367		285		215		162
Performance							
Data Transfer Rate, Native (MB/s)	240		120		120		30
Uncorrected Bit Error Rate	1x10^{-19}		1x10^{-19}		1x10^{-19}		1x10^{-18}
Data Buffer (MiB)	2000		256		256		3.4
Tape Speed – read/write (m/s)	5.62		2	3.74	2	4.95	3.4
Tape Speed – search/locate (m/s)	12		9		9.5		8.0 8.3 8.3

Tape Speed – high speed rewind (m/s)	12	9.5	9.5	8.3

Table 3 - Miscellaneous Specifications

Performance of FSA/TAA Feature Set

The StorageTek T10000C tape drive offers a drive acceleration feature to improve data transfer rates when running older applications that have not been fully tuned for newer tape equipment. This feature is called the Tape Applications Accelerator (TAA)

With each new tape drive model developed, the write speed of the drive has increased. This, in turn, has caused the tape movement system to take longer stopping and starting.

In the early days of tape, each drive had a capstan motor system that allowed the tape to be moved past the read/write head. These tape devices were known as start/stop devices. Tape densities were very low, in the 800-6250 bits per inch (BPI) range. Full tape capacities were at 200 megabytes per tape (on round reel). To start or stop, these drives took a mere 1 millisecond (.001 of a second). These were the 9-track (3420) drives and during this time, many of our current tape applications were developed.

As tape capacities increased the streaming tape drive came to be. These were smaller in size and had more read/write tracks, 18- or 36-tracks, on rectangular cartridges. This was 3490 technology (800 megabyte to 2GBs in cartridge capacities). These streaming drives removed the capstan motor system. Now, they moved the tape with their reel system and start/stop times greatly increased. They were now in the 0.5 to 1.0 second time frame. To compensate for these slower start/stop times, larger tape drive data buffers were added (16, 32, and 64 megabyte data buffers). The large buffers hid the stop/start time (called the drive's "football".)

The follow-on 20-75 GB cartridge 3590 technology had even higher capacities and even slower start/stop times – 1.0 to 1.5 seconds. Larger data buffers, 64-128 MBs, were used to hide this even larger football. Next was the 200-1000GB cartridge capacity 3592 with 128-256 MB data buffers to hide their footballs. Start/stop times were in the 2.0 to 3.5 second ranges. Older tape applications were now running slower on the newer tape drives.

These older tape applications have not been tuned to run on streaming devices. Syncing of the drive's data buffer to the tape media (start/stop) occurs too frequently and minimizes the drive's ability to hide the footballs. If a synchronize data buffer operation occurs at a frequency smaller than the size of that drive's data buffer size, then the stop/start times will be observed and visible to the tape application. When data buffer sizes were small, the sync frequency caused by smallish data set size was not an issue, but with the huge data buffer sizes now in use, syncing of the data to tape frequencies requires a solution.

IBM implemented new 3590 commands to allow drives to keep streaming and to limit the stop/starts. Buffered write tape marks for labels and data set separations were implemented within new tape applications and Virtual Back Hitch (VBH) was designed into the 3592 drive.

The StorageTek T10000C has a similar solution called File Sync Accelerator (FSA). However, it also has the same limitations, as does IBM's Virtual Back Hitch (VBH) - still too slow for applications that have been tuned for start/stop drives. That is, neither VBH nor FSA can take full advantage of the very high data rates the newer machines are capable of. Newer drives have 120 – 160 MB/sec data rate possibilities but VBH and FSA only help speed small workloads to the ~25 MB/sec range.

The Tape Application Accelerator (TAA) feature increases data rates, reference Illustration 15, by converting all write tape mark commands to buffered write tape mark commands and not syncing the data buffer to media each time (usually only once at the end). In order to utilize the drive's very large data buffer (2GBs), the tape drive decides when the best time to write to media, not the application. This inhibits the applications synchronization operations (stop/start events) from being seen at the application level. The drive's buffer size is now able, once again, to hide the effects of the drive's footballs. TAA allows the drive to stream for longer periods of time.

If the customer's tape applications cannot or will not change to use streaming tape drives, Oracle highly recommends that customers try this TAA feature in their mainframe environments and assess the tradeoffs.

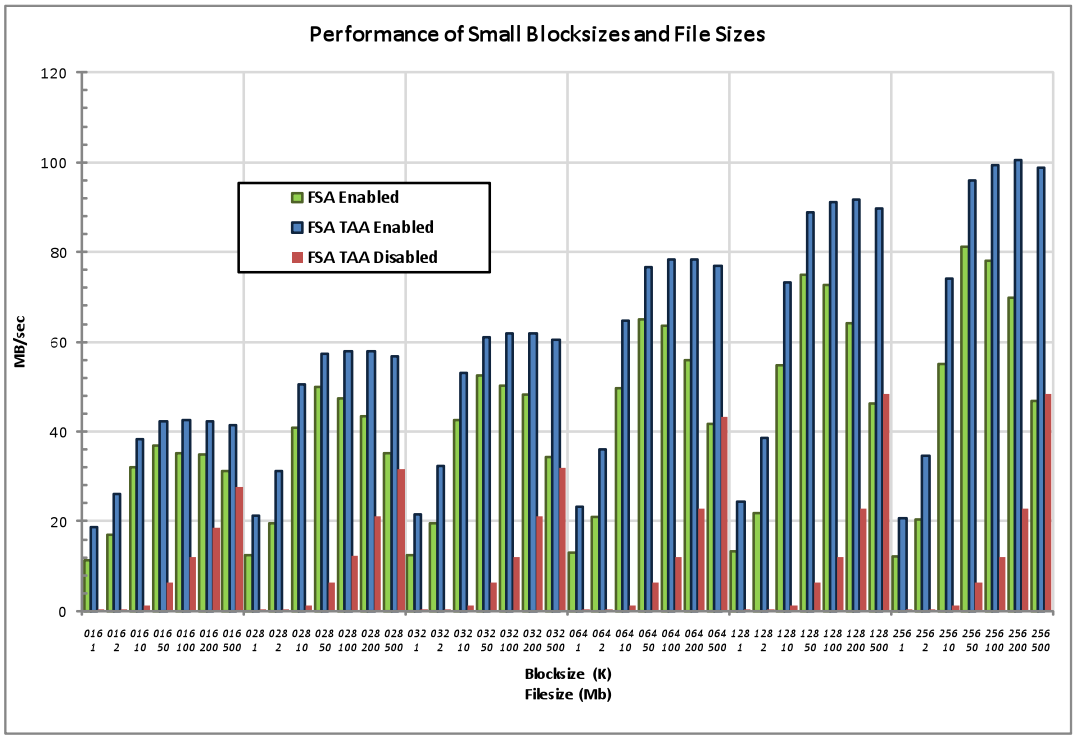


Illustration 15 - Small Blocksize Performance Characteristics

Additional Information

To obtain additional information about any of Oracle's tape drive products please access the following URL: http://www.oracle.com/technetwork/documentation/tape-storage-curr-187744.html

StorageTek T10000C Tape Drive FICON Performance

December 2011

Oracle Corporation
World Headquarters
500 Oracle Parkway
Redwood Shores, CA 94065
U.S.A.

Worldwide Inquiries:
Phone: +1.650.506.7000
Fax: +1.650.506.7200

oracle.com

Oracle is committed to developing practices and products that help protect the environment

Hardware and Software, Engineered to Work Together