IBM Magstar 3494 Virtual Tape Server

Performance White Paper

Version 2.1

IBM Corporation 1 June 1999

Introduction

This paper provides performance information on the IBM Magstar Virtual Tape Server (VTS), the 3494 model B18. It is intended for use by IBM field personnel and their customers in designing tape server solutions for their applications. The VTS, model B18, features the availability of larger tape cache, improved throughput performance with the larger caches, and, with the *ESCON high performance option* (EHPO), offers substantially improved throughput performance and effective tape cache capacity when operated with compressible data. This paper also includes data on the performance of the *SCSI Host Attachment* and *Import/Export* features available as of May 1999.

Product Description (VTS, model B18)

The general availability of VTS, 3494 model B18, was 28 August 1998. The model B18, compared to the previous B16 model, features a new more powerful RS/6000 processor and an optional larger *tape volume cache* (the disk storage) of 288 GB physical storage, doubling the maximum previously available. Additional improvements, available as of 13 October 1998, include the availability of four ESCON channel connectivity and an *ESCON high performance option* (EHPO) which markedly improves the server throughput performance when operating with compressible data.

Figure 1 shows the physical configuration of the VTS, model B18. The left hand cabinet includes the controller and tape cache (disk storage), while the right hand frames constitute the IBM 3494 library component of the system, including up to six 3590 tape drives.

Figure 2 shows a schematic of the principal functional units in the VTS, model B18. The host connectivity is provided by two or four ESCON adapters. Two ESCON channel attachments (Additional Enhanced ESCON Channel Attachments - FC #3302) can be replaced with one or two SCSI Host Attachment Features (FC #3422), as shown in Fig. 2, to provide up to four SCSI host attachments to the B18. The data compression feature in the ESCON adapter is part of the EHPO. With EHPO all data within the system below the ESCON adapters is in compressed form. Data compression is standard on the SCSI host adapters. All data passes through, and is managed, by the logical software components indicated in the RS/6000 logical block. Incoming data is then destaged to tape volume cache (HDDs) that is attached through two SSA adapters. The tape volume cache is available in 72 GB increments, up to 288 GB total. The HDDs are organized in RAID-5 units of 2 + parity + spare. Data is striped across all the 72 GB units as shown in Fig. 2.

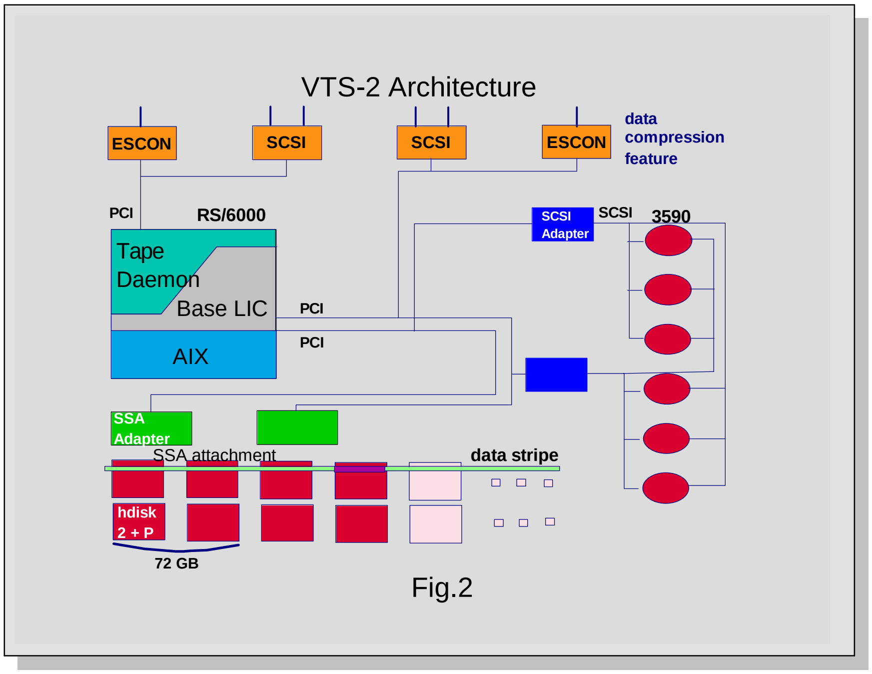

Fig.2

One of the significant features of the virtual tape server is that scratch mounts do not require the physical mounting of a tape cartridge. The tape mount request is handled at electronic speed and data transfer can begin immediately into the tape cache. Once a complete tape volume is received in tape volume cache, it is scheduled for migration to physical tape under the control of licensed internal code. The 3590 tape drives are part of the IBM 3494 tape library. Without the data compression feature, data is compressed at the 3590 and stacked as multiple compressed volumes per physical tape until the tape is full. With the data compression feature data is compressed at the ESCON adapter.

Host read requests are satisfied at either the tape volume cache (read hits), or by first staging the data from 3590 physical tape to the tape volume cache (read misses).

The base VTS, model B18, presents itself to the host as two tape control units (TCUs) with 16 logical tape drives each for a unit with two ESCON channels; with four ESCON channels, part of the EHPO option, the presentation is as 32 or 64 logical 3490E tape drives on two or four TCUs. The optional larger tape volume cache capacity improves throughput performance directly through providing read hits and also typically provides up to seven hours of buffering for write intensive periods. It also improves the effective bandwidth of the VTS by increasing the bandwidth of the tape cache.

The VTS configurations include a choice in the number of 3590 tape drives from three to six. A larger number of physical 3590 tape drives improves performance during periods of high specific mount request periods (read misses).

The Base VTS Subsystem Performance Metrics

In the following sections we discuss the performance of the VTS. The primary metric we use is the *sustained write* rate in MB/s. The opposite of this rate is the *read miss* (or *read recall*) rate, which closely mirrors the resources used by the sustained write. We promote the use of these metrics, especially the *sustained write*, as a way of succinctly characterizing VTS performance. The *sustained write* rate can only be obtained after the system has been run at least once on the order of several days to assure that a steady state has been reached between the tape cache and the I/O and migration activity. The *sustained write* rate measures the collaborative overlapped performance of the ESCON input, the RS/6000 processor and its internal paths, the file system, the disk adapters, the HDDs, and the 3590 tape drives, and it does it under circumstances where the system is managing a large number of volumes, with at least sixteen of them active. This is the net performance value of the VTS to the customer.

All of the measurements, unless specifically otherwise noted, were made with four ESCON channels.

We do also report *peak write/fill* rates and *read* hit rates because they are an integral part of VTS performance, and since they are higher, can be used to enhance higher overall average system throughput when they are taken advantage of appropriately.

It is tempting to quote only these latter, higher, MB/s performance numbers of specific tasks within the VTS or competing products when comparing them. We discourage this specsmanship because they can lead to misleading assessments of comparative performance value. We examine such a comparison in Appendix A.

Write Performance

Write performance quoted here has been derived from measurements that simulate typical user environments; namely sixteen applications writing 800 MB tape volumes simultaneously. The block size used is 32 KB and the BUFNO parameter is set at 20. The measurements quoted in these sections apply to a VTS without the EHPO option. We describe the effect of the EHPO data compression feature on these performance projections in a following section entitled “The Effect of EHPO Data Compression on VTS Performance.”

The tape volume cache is allowed to fill and the system is then operated for some time in a sustained mode with new data being added to the tape cache while older data in the tape cache are being copied to physical tape. We then characterize write performance in two ways: the “*peak write/fill*” rate and the “*sustained write*” rate as described below.

Peak Write/Fill Rate

The peak write/fill rate is defined for a tape server that has been in long term use, has most of its new data copied to tape, and now is subject to an incoming rate of writes that saturates its throughput. (Following the copying of data from tape cache to physical tape, the copy of the data in tape cache remains, providing for possible read hits in tape volume cache while making the

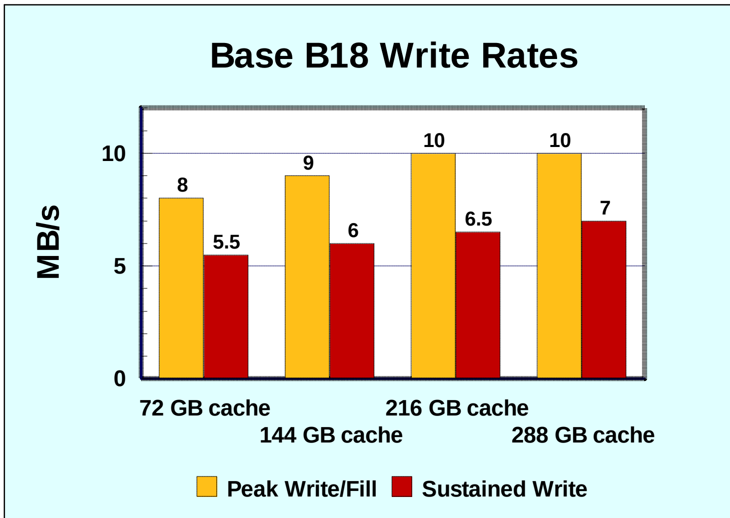

Fig. 3

space the logical volumes occupy immediately available for new writes, if necessary.) Under these conditions, the predominant activity of the virtual tape server is to place the incoming data into tape volume cache (the HDDs). Until a fill threshold of the tape volume cache is reached there is little copy activity of the data from tape volume cache to physical tape. This peak write rate (Peak Write/Fill) for the virtual tape server is shown in Fig. 3 as a function of the capacity of the tape volume cache. The peak write rate ceiling is between 8 and 10 MB per second, varying with the size of the cache. The peak write rate performance can be used to ride through periods of heavy write activity if the tape volume cache has been previously copied to physical tape. The approximate hours of peak write/fill rate possible with different tape cache sizes is shown in Fig. 4 for the base B18 model. With the EHPO option, the higher effective cache capacity and data rate with compressible data compensate each other to yield approximately the same time at peak fill rate as shown in Fig. 4.

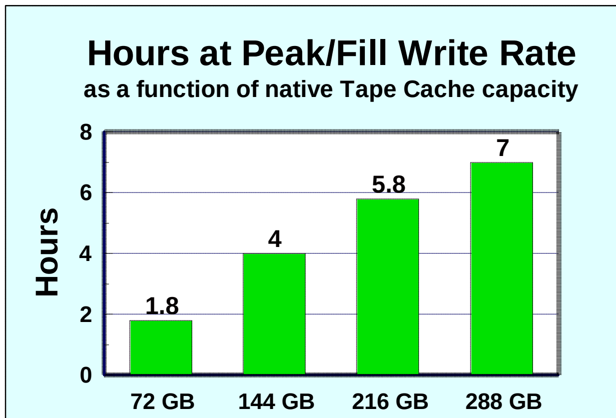

Fig. 4

Sustained Write Rate

The *sustained write rate* is defined for a tape server that has been in long term equilibrium operation with the rate of incoming data equaling the rate at which data is being copied from tape volume cache to physical tape. Thus for sustained write there are three separate IO activities going on simultaneously. New data is being written to the tape volume cache, and aged data is being read from the tape cache and then written to the tape library. Because of this additional activity the sustained write rate is lower than the *peak write/fill* rate. The sustained write rate varies with tape volume cache size between 5.5 and 7 MB per second (cf., Fig 3).

Read Performance

As with the write performance data, above, we quote the projected read throughput performance for a particular operation operating exclusively. As on the write data, the data quoted here have been obtained with 800 MB volumes, 32 KB blocking of data, and BUFNO=20.

Read Hit (Read from Tape Volume Cache)

A *read hit* is a host read request which finds the requested volume in tape volume cache (the HDDs), whether the data has been copied to physical tape or not. Thus the data transfer to the host involves mostly locating the data on the HDDs, staging them to the VTS CPU memory, and then transferring them to the channel adapter(s). The read hit throughput in MB/s, assuming several concurrent applications issuing overlapping read requests, is shown in Fig. 5.

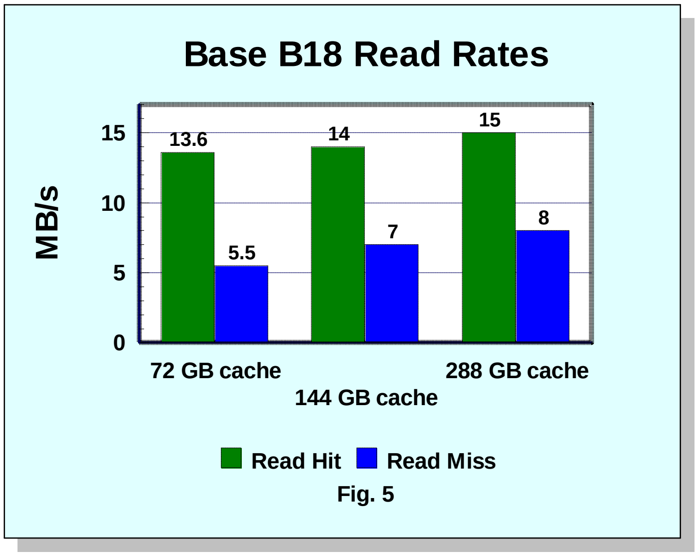

Fig. 5

Read Miss (Read from Tape Library)

A *read miss* (also referred to as *read recall*) is a host generated request to read a tape volume which has been copied to physical tape, and is no longer available in tape volume cache. Thus a read miss requires staging of the requested tape volume from the tape library to tape volume cache (*mount time*), the transfer from tape volume cache into VTS system memory, and then transfer from memory to the channel adapter. This process is approximately the reverse of the *sustained write* operation. The read miss throughput in the limit of several such overlapping operations operating exclusively is shown in Fig. 5.

With respect to read misses, it is necessary to observe that if these are the result of an application requesting the mounting of a large number of volumes nearly simultaneously, there will be a delay before all volumes are mounted. For example, if the VTS is configured with the maximum of six tape drives, and two of them are assigned to receiving of data being copied from tape volume cache, then only four specific mounts can be handled simultaneously. Additional mounts will be queued on those tape drives and will be handled sequentially. Thus average volume mount time statistics on the VTS may appear larger than expected.

Read Backwards

Reads backwards refers to a read request sequence which accesses the tape data in the opposite sequence order from the one in which they were written. For example, if the application requests the previous block of data from the one it just read back from tape. This is a very cumbersome, low performance type of tape data access on physical tape drives. This kind of tape access method is much more efficient on the VTS because the block IOs are actually executed on HDDs.

Volume Mount Response Time

The VTS has two characteristic mount response times; the time required for the VTS to respond with a “ready” after an host request to mount a tape volume. A *fast ready* is associated with the transfer of data which is already in tape cache and occurs in typically two to three seconds. A *normal ready* requires the mounting of a physical tape and transfer of the tape volume data involved to tape cache.

*Specific mount write update*s to tape cache (the writing of an updated tape volume which is still in the VTS tape cache), *scratch mounts* (the writing of a volume which does not already exist on the VTS to a public scratch tape), and *read hits* have a characteristic response time of two to three seconds (fast ready) unless the system is heavily occupied by other activity. These response times are much shorter than with conventional tape drives because they are done at electronic speeds.

The mount response time for *read misses* and *writes to specific mount* or *private scratch pool* volumes (if their data is not in tape cache) are longer because they require the staging of the requested tape volume to tape cache before a "ready" is presented to the host. The response times thus include the time to physically copy an entire volume of the requested data from tape to tape volume cache. (The additional time to destage the data from VTS memory to tape cache (HDDs) mostly overlaps the tape data transfer.) A best case response time for such mounts is on the order of two to three minutes, and would only happen when little else is active on the VTS.

The Effect of EHPO Data Compression on VTS Performance

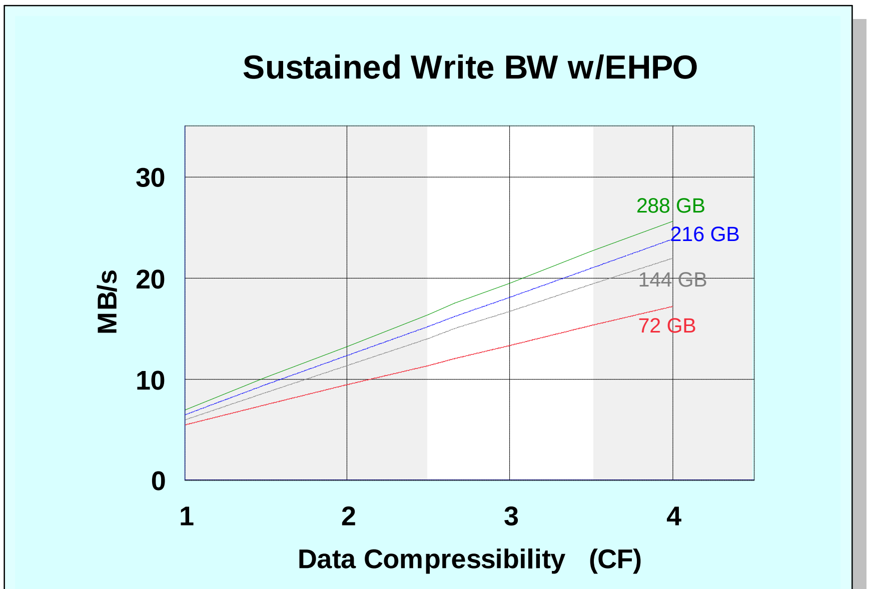

Fig. 6

The optional EHPO feature applies compression to data coming into the VTS and decompression to data leaving the VTS via the ESCON channels. The principal benefit of compression is to increase the throughput performance of the VTS and increase the number of virtual volumes in cache storage available for cache hits. Depending on the nature of the workload, throughput performance is constrained by the ESCON adapters to the VTS, the RS/6000, the SSA adapters (HDD attachment), the HDDs, or tape drives. For most realistic workloads, however, the throughput of the VTS with EHPO can be estimated from a consideration of the size of the tape cache . Figs. 6 and 7 show the dependence of sustained write and peak write MB/s throughput performance as a function of data compressibility, *CF*. The tape cache constraints are shown for the available physical tape cache sizes. The throughput shown in the figures has been measured with four ESCON channels/adapters. The data in Figs. 6 and 7 are based on selected measurements with a data compressibility factor, *CF*, up to four. Analytical expressions for the curves in Figs. 6 and 7 are given in Appendix B.

Most of the data transfers at the ESCON adapters and within the tape cache incur resource utilizations in proportion to the amount of data transferred; thus throughput performance, in the main, is increased approximately in proportion to the compressibility of the data.

For most regularly achievable data compression values the throughput is tape cache constrained to the values shown. Tape data compression experience indicates that many applications can expect an average data compressibility on the order of three, the highlighted region in Figs. 6 and 7. Thus many customers that fall within this profile can expect three times the performance with EHPO as compared to the 3494 B16.

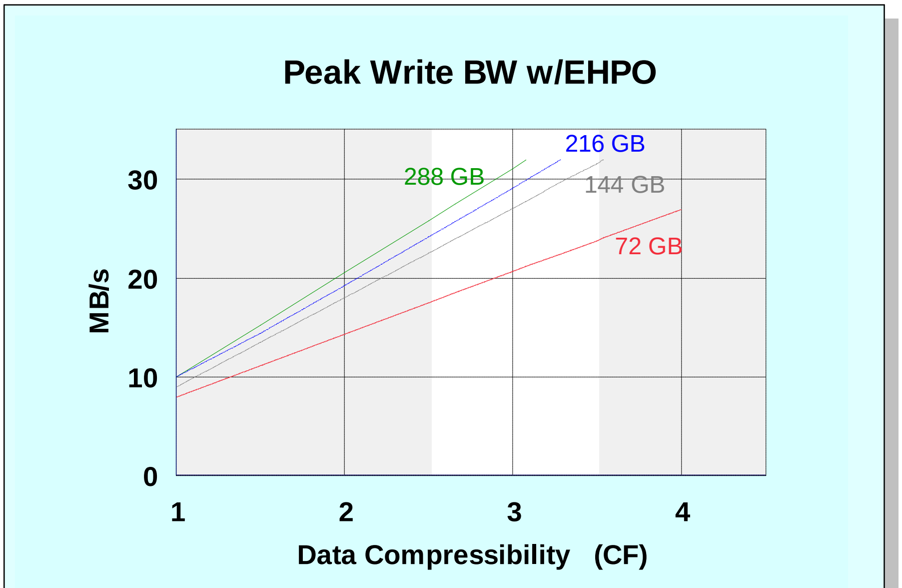

Fig. 7

We note that at an operating point with CF=3 the *peak write* throughput is on the order of 60% larger than the *sustained write* throughput. This additional bandwidth is available for up to seven hours, depending on the size of the tape cache (cf., Fig. 4), when used in conjunction with a tape cache that has been fully copied to physical tape. Thus the VTS provides a generous *peak write* throughput capability on top of its already significant *sustained write* rate.

Customers that require the maximum throughputs that are shown in Figs. 6 and 7 should configure the VTS with four ESCON channels/adapters.

The read hit and read miss (read recall) rates with compressed data correspond approximately to the peak write and sustained write rates, respectively. In most measurements we observe that the read rates are slightly higher than the corresponding write rates.

VTS Input Throttling

VTS input throttling is implemented to allow background processes, those not involving data transfer over the ESCON channels, to obtain a sufficient share of the resources to maintain the system in balance. Such processes are, for example, recall of data that are on physical tape and copying of data in tape cache to physical tape.

If the VTS has been in a peak throughput mode for some time, i.e., with the copy process from tape volume cache to physical tape not keeping up with the rate of new incoming data, it is possible for the cache to reach a fill level with new data where no more can be accepted safely without first copying some of the data in the tape volume cache to physical tape. The same situation can also occur with recalls (i.e., read misses or specific mount write updates) since the data being recalled has to displace data in tape cache. To avoid such situations, the VTS throttles the input. A fixed delay is added to each incoming write operation. This increases the available internal bandwidth for the background processes and allows it to stay up with new incoming data. If the VTS is observed to be often or persistently in a throttling mode (large input delays), the throughput of the VTS is being utilized at its maximum sustainable level and increased VTS capacity may be advisable.

B18 Performance using the optional SCSI Host Attachment

A configuration option for the B18, available beginning in May 1999, is to replace the Additional Enhanced ESCON Channel Attachments (FC #3302) with one or two SCSI Host Attachment Features (FC #3422). Each of these SCSI Attachment Features, also referred to as *adapters* here, provides connectivity for up to two fast/wide or ultra-SCSI busses. This section presents the performance of the VTS as viewed from the SCSI interface. At the end of this section we do also discuss the effect of SCSI attachment on the VTS performance as viewed at the ESCON interface.

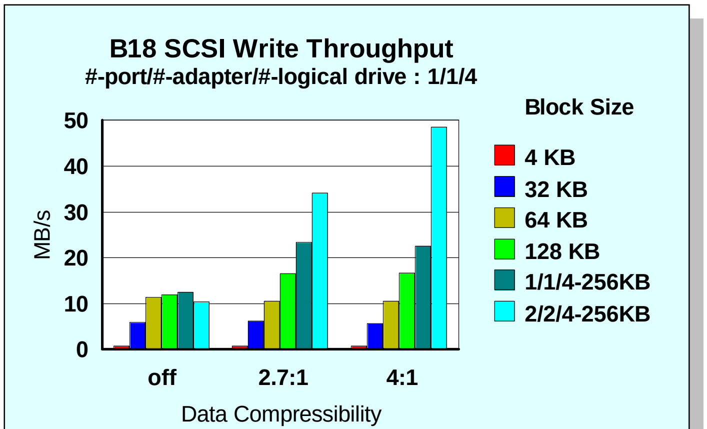

Fig. 8

A number of user controllable parameters determine the throughput performance of the SCSI attachment. Figure 8 shows the effect of block size on the write throughput of one adapter with data of various compressibility (a two adapter result at 256 KB is also shown for comparison). The measurements were done with no I/O load on the ESCON channels. The tape cache capacity was 288 GB and the experimental conditions are those for *peak write/fill*. For block sizes of 64 KB and below there is little effect of compressibility on throughput. At the larger block sizes throughput improves with data compressibility and the improvement goes with block size. The write throughput goes from near one MB/s with 4 KB blocking to about 23 MB/s with 256 KB blocks (at compressibility of 2.7 and above). For the 256 KB block case we also show a bar for the case with two logical tape drives being driven through separate SCSI busses to two adapter cards (2-ports/4-log drives on two adapter cards).

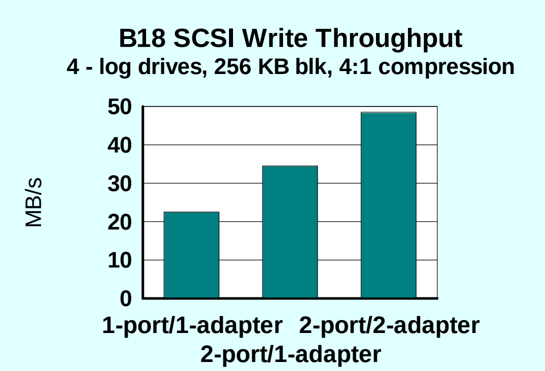

Fig. 9

The effect of multiple SCSI paths and adapter cards is further elaborated in Fig. 9 for the conditions indicated. The left and right bars correspond to the (1/1/4 and 2/2/4-256KB) cases in Fig. 8. For the 2-port/1-adapter case the number of logical drives was also varied. Four logical drives represents approximately the maximum throughput. Either decreasing or increasing (up to eight) the number of logical drives can reduce the throughput on the order of 10%. **For maximum throughput one should maximize the number of data paths: both SCSI busses and the number of host adapters.**

The read throughput generally follows the write throughput dependence shown in Figs. 8 and 9, except that it tends to be somewhat lower (up to twenty percent) for the larger block sizes.

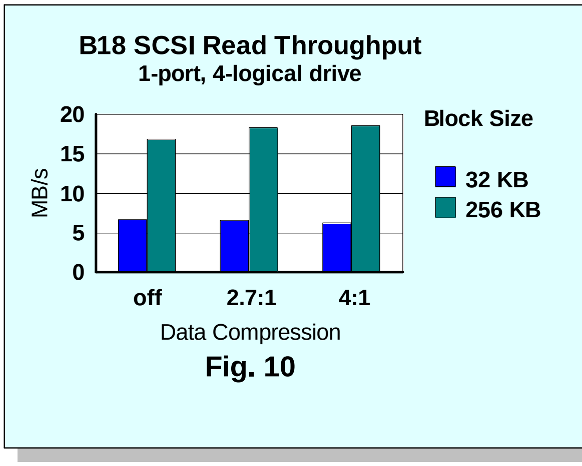

Fig. 10

Figure 10 shows the insensitivity of read throughput with data compressibility and the strong effect of block size for the same conditions as in Fig. 8 for writes.

The effect of SCSI input on ESCON throughput appears to be small, but the two throughputs share an aggregate maximum. For example, with a *peak write/fill* 2 x ESCON write input (32 KB block, 2.7:1 compression) saturated at about 16 MB/s, when a two-logical-drive/two-card SCSI write input (256 KB blocking, 4:1 compression) is added, the total VTS input rate , ESCON plus SCSI, is observed to reach a maximum of about 47 MB/s. This is about the same maximum rate as was observed with the SCSI input alone. Of the 47 MB/s about 15 MB/s is observed to be coming across the ESCON channels.

B18 Performance using the optional Import/Export Feature

Beginning in May of 1999, the B18 offers an optional tape cartridge data Import/Export Feature (Feature Code 4000). Using this feature, user specified tape data volumes can be assembled on a tape cartridge which can be physically removed from the VTS (export). Similarly data thus exported on this or other VTSs can be physically "imported" (by introducing the cartridge into the import/export station) to become part of the tape volume repository managed by this VTS.

Import/Export performance is measured by the elapsed time required to perform an export or import task consisting of a number of logical tape volumes on a number of physical cartridges. The considerations that determine the time required are:

1. Number of physical cartridge mounts/demounts in the aggregate task
2. Total amount of data transferred
3. If the tasks involve selected volume export/imports from the same physical volume, then tape drive repositioning time between volumes needs to be accounted for (avg. random repositioning time [search time] is about 21 seconds)
4. What the non-Export/Import workload is on the VTS

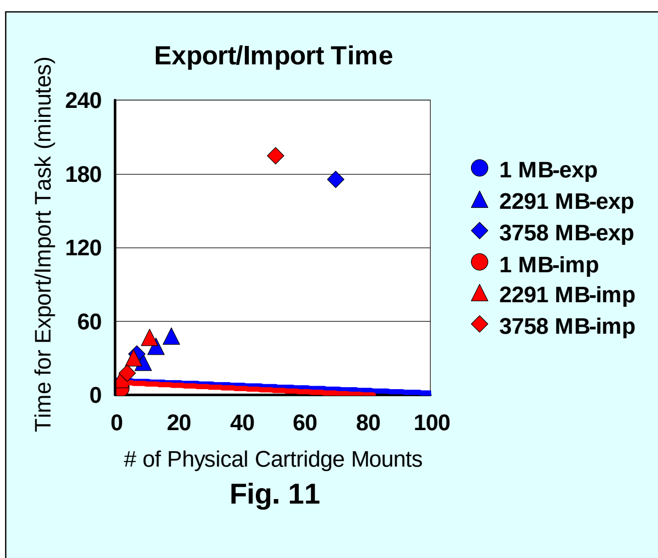

Fig. 11

In Fig. 11 shows some measurements on a variety of import/export tasks. The principal determinant of task time is seen to be the aggregate number of physical tape cartridge mounts in the task. The trend lines shown (blue line [lower] for export, red line [upper] for import) are based on points all involving eighteen or less total mounts/demounts and amounting to a total data transfer of about 2.3 GB (data as compressed). They describe well the outlying points, which involved about a 64% greater total data transfer.

At minimum, the export or import of a single logical tape volume requires two physical tape cartridge mounts; to mount the source and destination cartridges.

The trend lines in Fig 11 are given by:

$$\text{Export\_time(min)} = 0.00296 \times (\textit{total\_MB}) + 2.36 \times (\#\textit{mounts}) + 0.35 \times (\#\textit{search})$$
$$\text{Import\_time(min)} = 0.00276 \times (\textit{total\_MB}) + 3.75 \times (\#\textit{mounts}) + 0.35 \times (\#\textit{search})$$

where the *(#mounts)* is the sum of source and destination drive mounts. Generally, the *(#source_mounts)* + *(#search)* = the total number of logical volumes in the source of the export or import. However, in some cases the number can be larger than that. Although the data in Fig. 11 did not include search times, the formulas have been elaborated to also cover the cases of selective Export/Import which will involve search times. The *(total_MB)* are the net MB after VTS B18 EHPO feature compression.

The present results were obtained from measurements on a three frame 3494 library with one dual gripper accessor and floating home cell operation. Larger libraries will incur longer mount times. A single gripper accessor and fixed home cell operation would also increase the observed mount/demount time. Furthermore, the exact export or import task time will also depend on cartridge placement in the library and specific distribution of logical volumes on the physical cartridges. Thus the illustrated results should be used as a guideline, and not a guarantee of performance.

There is an overhead associated with the beginning of a task involving the import or export of a number of logical volumes. This overhead is not significant on the scale of Fig. 11, and is not included in the task time formulas above. However, when the formulas are used to derive the time for the import or export of a single volume, for example, a discrepancy of up to two minutes can occur, both plus and minus. This discrepancy is attributed to the overhead and the fact that the concept of a statistical distribution of cartridge locations and accessor position breaks down.

As one might expect, the import/export task time will increase if there is another I/O load on the VTS. In a measurement of the time to complete a ten logical volume export request requiring *thirteen* mounts and a ten logical volume import request requiring *six* mounts (total 2.3 GB, each) at a time when the VTS was saturated in a “peak” write mode from the host, the former (export) time was elongated by about 25% versus no host load; the latter (import) by about 50%. Thus, the less mount activity an import/export task requires the more it is affected (relatively) by contention with other I/O activity on the VTS. Also, import/export with simultaneous (saturated) recall activity increases the import/export task time by about 50%.

Conversely, neither “peak” or “sustained” write throughput nor read hit throughput as measured at the ESCON host were greatly affected by simultaneous import/export activity on the VTS B18. However, read miss (recall) maximum rate was reduced to approximately half.

Import/Export Recommendations

1. It is recommended that import/export tasks be run during non-peak I/O activity times to minimize contention.
2. In order to reduce physical mounts required for export, the export operation should be performed as close to virtual volume creation time as possible.
3. Virtual volumes with similar creation dates should be exported together. This increases the probability that multiple volumes will be found on the same physical source cartridge.

Performance Considerations and Tools

Workload Considerations

As noted in the introductory section on VTS, 3494 model B-18, one of the VTS objectives is to reduce the number of physical tape drives required and to achieve a more efficient use of tape by “stacking” multiple volumes on a single physical tape cartridge. The VTS design point is to satisfy the typical workload of an installation, including peak activity which can be buffered in the tape cache, but excluding some workload components that can not achieve a performance benefit from VTS; namely, excluding HSM volumes, DUMPs, and others that already tend to write full physical tapes. These latter will not achieve any net efficiency in going from system to tape though the VTS. And they will not achieve any benefit from residence in tape volume cache. On the contrary, they will consume a significant fraction of the VTS internal bandwidth and will monopolize a fraction of the VTS tape drives.

Our performance recommendation is that, in considering whether to commit workload components such as HSM and DUMPs to the VTS, the tools described below should be used to assess the performance impact. There will be cases in which allocating such work to native tape drives or native tape drives within the 3494 tape library can significantly enhance the performance of the VTS.

It is recognized that customers will want to assign their HSM and DUMP volumes to the VTS for simplicity and uniformity in data management. It is an objective of VTS development to accommodate these customer desires and requirements over time with minimal performance impact.

Effect of ESCON Distance

The performance results quoted above in this paper have been obtained with the ESCON connection between host and VTS within small distances compared to one kilometer (km). We have also measured the effect of ESCON distance at three points up to 18 km. On that basis we offer the following rule of thumb for the effect of ESCON distance using IBM ESCON directors:

The VTS throughput is reduced approximately 1.9% per kilometer of ESCON distance between the host and VTS for 32 KB block transfers.

Tape Magic

Tape Magic is a high-level tape subsystem configurator available to IBM customer representatives that is intended to give an initial prediction of a tape configuration that would satisfy a customer’s tape processing needs. Tape Magic predicts both native and volume-stacking configurations. Input to Tape Magic is answers to a half-dozen or so simple questions about basic customer tape workload characteristics, typically entered via a Thinkpad™ on a visit to the customer’s location. Because Tape Magic does not directly process any host-processor statistical data, such as MVS SMF records, it is also useful for host platforms that do not provide data that can be input to IBM's more detailed configuration tools.

Workload Analysis and Configuration Estimation Tools

A more accurate assessment of a VTS configuration than possible with Tape Magic can be made by a detailed analysis of the customer's workload as represented in SMF records, RMF data, and tape management system data. The current tool available to IBM representatives is called the Virtual Tape Subsystem Analyzer (VTSA) and provides a detailed analysis of existing customer tape workload characteristics and projects the required VTS configurations for a subset of that workload. VTSA uses as input an extract of selected SMF records, as provided by the DFSMS™ Volume Mount Analyzer (VMA). VMA is an analysis product whose main purpose is to project Tape Mount Management DASD buffer requirements and benefits, but also provides basic tape workload characteristics such as mount and drive allocation activity. VTSA augments VMA by providing additional workload statistics, such as both input and output tape data transfer activity by hour. To project a VTS configuration, the user first uses the extensive filtering capabilities of VMA to filter out tape activity, such as output files destined for trucking to a remote vault and tape activity that already efficiently utilizes native tape, that will not be volume-stacked. VTSA then projects required VTS configurations based on current workload and expected workload growth. VTSA also provides numerous statistics on expected VTS cache performance. VTSA is planned to eventually be phased out and replaced by a similar tool called Batch Magic, which would have more extensive data input capabilities.

Performance Monitoring Tools

VTS generates data that is transmitted each hour to the host processor, where the data is embodied in an SMF type 94 record. This SMF record also contains information on library performance associated with native tape drives. Information provided in the SMF type 94 record includes logical and physical drive usage, number of fast-ready (virtual scratch), read-hit, and recall mounts, channel and tape input and output data transfer activity, and cache usage statistics. IBM provides routines that give hourly and daily reports on these VTS statistics. This allows the customers to understand the level of activity of their VTS subsystems, and allows customers to also, with assistance provided by IBM field personnel, to determine when the limits of the VTS subsystems are being reached.

Conclusions

The virtual tape server, VTS model B16, has demonstrated a clear customer requirement for consolidated tape data management and automation, while taking advantage of technological advances that reduce hardware and floor-space requirements. The VTS, model B18, builds on this base to present significantly improved throughput performance with a maximum sustained write throughput performance approaching 32 MB per second with the EHPO option. At a data compressibility of three, and the 288 GB tape cache option, the sustained write throughput of the B18 is 19.4 MB/s, a threefold increase over the model B16. Read miss (recall) rates are improved similarly. Beginning in May of 1999 the VTS B18 also offers SCSI host attachment with up to four SCSI busses and Import/Export of VTS format data cartridges.

Appendix A

Comparison of "peak" and "sustained" performance metrics.

We examine here how one can compare virtual tape systems when some are specified in terms of peak read and write rates, while others specify sustained write performance.

Example: If a competing product specifies a pea k aggregate read rate of 55 MB/s and a peak aggregate write rate of 50 MB/s, using data that compresses by a factor of four. What is its effective sustained write throughput?

The first thing to realize is that *peak* transfer rates do not involve the tape subsystem at all. They are transfers from the ESCON interface to the disk subsystem and the reverse. These are basically specifications of a disk subsystem.

To estimate an equivalent sustained write rate that might be achieved by this disk subsystem when attached to a tape library, we can assume that the peak write accounts for the host write and the peak read accounts for the migration of the data to tape. Under these circumstances, if the read and write are mostly competing for the same constraining resource within the disk subsystem, one can estimate the sustained write rate in the form:

$$(susWr) = 1 + ((1/peakWr) + (1/peakRd)) = 26.1 MB/s.$$

This is a number comparable to that of the VTS with the EHPO option (cf. Fig. 6) at a data compressibility of four. However, it does not account for the additional overheads associated with migration and any other bottlenecks along the path from the disk subsystem (tape cache) to physical tape. Thus it is merely an upper bound for *sustained write* provided by the disk subsystem.

Table 1. Linear Coefficients for the GB/hr Curves in Figs. 6 and 7.

	Tape Cache:	72 GB	144 GB	216 GB	288 GB
Sustained Write	*m(GB)*	3.91	5.34	5.79	6.23
	b(GB)	1.58	0.71	0.76	0.82
Peak Write	*m(GB)*	6.33	9.04	9.75	10.54
	b(GB)	1.67	-0.04	-0.1	-0.54

Appendix B

Analytical data on VTS throughput as a function of tape cache capacity and data compressibility.

The B18 w/EHPO sustained and peak throughput rates shown in Figs. 6 and 7 can be reproduced from linear relationships of the form:

$$(MB/s) = [m(GB) \% CF] + b(GB),$$

where the parameters *m* and *b* are functions of the size of the Tape Cache in GB, and *CF* is the data compressibility. The values of *m* and *b* are given in Table 1.

Care should be taken to make sure that the input, *CF*, and output, (MB/s), of this formula is maintained within the bounds shown in Figs. 6 and 7; namely $CF [4$, and $(MB/s) [32$. Performance claims at larger compression factors, but still with $(MB/s) [32$, should be considered non-validated extrapolations.

DISCLAIMERS

The performance information contained in this document was derived under specific operating and environmental conditions. While the information has been reviewed by IBM for accuracy under the given conditions, the results obtained in your operating environments may vary significantly. Accordingly, IBM does not provide any representations, assurances, guarantees or warranties regarding performance. Please contact your IBM marketing representative for assistance in assessing the performance implications of the product in your specific environment.

All statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.

TRADEMARKS and COPYRIGHTS

Table of Contents