IBM System Storage

April 2025

IBM® TS7780, TS7780T, and TS7780C

Release 6.0

Performance White Paper

Version 1.0

By Sammy Utter

Virtual Tape Performance

IBM Tucson

Table of Contents

Table of Figures

Introduction

This paper provides performance information for the IBM TS7780, TS7780T, and TS7780C, which are the three current product configurations in the TS7700 family of storage products for the 6.0 release. This paper is intended for use by IBM field personnel and their customers in designing virtual tape solutions for their applications.

This is an update to the previous TS7700 paper dated April 02, 2024, and reflects changes for release 6.0. This white paper only shows runs for the VEF model, although 6.0 supports both VED and VEF models. Major new function and improvements delivered with release 6.0 include:

- 32 Gb FICON adapters.
- New common cache controller drawers that currently have no active capacity. This controller will be called the 3948-CFD and be used for both HDD and SSD based TS7780. The cache controller has 32 Gbps fiber adapters for VEF connectivity.
- A choice of either SSD (XFD) flash based or HDD (XSD) based solutions both featuring the new 3948-VEF virtualization engine.
- New 3948-XSD HDD based expansion cache drawers with 12 x 20 TB HDDs in each drawer
- New 3948-XFD SSD based expansion cache drawers with 24 x 15 TB SSD drives in each drawer
- New 4 x 6.4 TB internal NVMe flash drives for the 3948 VEF.
- New 240 TB (157 TB Usable) HDD Storage (#7120) for use with 3948-XSD
- New 360 TB (260 TB Usable) SSD flash Storage (# 7123) for use with 3948-XFD

Hardware Configuration

The following hardware was used in performance measurements. Performance workloads are driven from IBM System z16 host with eight 32 Gb FICON channels.

Standalone Hardware Setup

TS7700	Drawer count	Tape Lib/ Tape Drives	Cloud	IBM System z™ Host
TS7780 VEF 3948 CFD/XSD	2, 4, 6, 8, 10	N/A	N/A	z16
TS7780T VEF-T 3948 CFD/XSD	2, 4, 6, 8, 10	TS4500/ 12 TS1160	N/A	
TS7780T VEF-C 3948 CFD/XSD	2, 4, 6, 8, 10	N/A	IBM COS 3403	

Grid Hardware Setup

TS7700	Drawer count	Tape Lib/ Tape Drives	Cloud	Grid links (Gb)	IBM System z™ Host
TS7780 VEF \3948 CFD/XSD	10	N/A	N/A	4x10	z16
TS7780T VEF-T 3948 CFD/XSD	10	TS4500/ 12 TS1160	N/A	4x10	
TS7780T VEF-C 3948 CFD/XSD	10	N/A	IBM COS 3403	4x10	

IBM COS 3403

Manager	Accesser	Slicestor	Gridlinks/Accesser	Gridlinks/Slicestore
M01	2xA00	9xS01	2x10Gb	2x10Gb

The following conventions are used in this paper:

Binary			***Decimal***		
Name	*Symbol*	*Values in Bytes*	*Name*	*Symbol*	*Values in Bytes*
kibibyte	*KiB*	2^{10}	*kilobyte*	*KB*	10^{3}
mebibyte	*MiB*	2^{20}	*megabyte*	*MB*	10^{6}
gibibyte	*GiB*	2^{30}	*gigabyte*	*GB*	10^{9}
tebibyte	*TiB*	2^{40}	*terabyte*	*TB*	10^{12}

TS7700 Performance Overview

Performance Workloads and Metrics

Performance shown in this paper has been derived from measurements that generally attempt to simulate common user environments, namely many concurrent jobs writing and/or reading multiple tape volumes simultaneously. Unless otherwise noted, all the measurements were made with 64 simultaneously active virtual tape jobs per active cluster. Each tape job was writing or reading 10.7 GB of uncompressed data using 32 KiB blocks and QSAM BUFNO=20 that compresses within the TS7780 at 5.351 using ZSTD compression. Measurements were made with eight 32-gigabit (Gb) FICON channels on a z16 host. All runs begin with the virtual tape subsystem inactive.

Unless otherwise stated, all runs were made with tuning values:

- DCOPYT=125,
- DCTAVGTD=100,
- ICOPYT=ENABLED,
- LINKSPEED=1000
- CPYPRIOR=DISABLED,
- CRCSET generation (CRC32) and all the CRC checks are enabled (see section 3.1.31 in the White Paper – TS7700 Library Request Command V5.3 for details

For TS7780T, there are additional settings:

- PMPRIOR=7000, PMTHLVL=8000,
- Reclaim disabled,
- Number of premigration tape drives per pool=10,

For TS7780C, there are additional settings:

- CPMCNTL=0
- CPMCNTH=60
- CDELDNT=16
- CLDPRIOR=3600 with only 4 FC5274 installed

Refer to the **IBM® TS7700 Series Best Practices - Understanding, Monitoring and Tuning the TS7700 Performance** white paper for detailed description of some of the different tuning settings.

Types of Throughputs

The TS7780 or TS7780T$_{cp0}$ is a disk-cache only cluster, therefore read and write data rates have been found to be consistent throughout a given workload.

The TS7780T$_{cp1->7}$ contains physical tapes to which the cache data will be periodically written and read, and therefore it exhibits four basic throughput rates: peak write, sustained write, read-hit, and recall.

The TS7780C$_{cp1->7}$ connects to cloud (object store) to which the cache data will be periodically written and read, and therefore it also exhibits four basic throughput rates: peak write, sustained write, read-hit, and recall.

Peak and Sustained Write Throughout

For all TS7780T$_{cp1->7}$ measurements, any previous workloads have been allowed to quiesce with respect to pre-migration to backend tape and replication to other clusters in the grid. In other words, the test is started with the grid in an idle state. Starting with this initial idle state, data from the host is first written into the TS7780T$_{cp1->7}$ disk cache with little if any premigration activity taking place. This allows for a higher initial data rate and is termed the "peak" data rate. Once a pre-established threshold is reached of non-premigrated compressed data, the amount of premigration is increased, which can reduce the host write data rate. This threshold is called the premigration priority threshold (PMPRIOR) and has default value of 1600 gigabytes (GB). When a second threshold of non-premigrated compressed data is reached, the incoming host activity is actively throttled to allow for increased premigration activity. This throttling mechanism operates to achieve a balance between the amount of data coming in from the host and the amount of data being copied to physical tape. The resulting data rate for this mode of behavior is called the "sustained" data rate, and could theoretically continue forever, given a constant supply of logical and physical scratch tapes. This second threshold is called the premigration throttling threshold (PMTHLVL) and has a default value of 2000 gigabytes (GB). These two thresholds can be used in conjunction with the peak data rate to project the duration of the peak period. Note that both the priority and throttling thresholds can be increased or decreased via a host command line request. For all the run in this white paper, PMPRIOR and PMTHLVL were set to 7000 and 8000 respective to achieve a longer peak duration.

For all TS7780C$_{cp1->7}$ measurements, CLDPRIOR was set to 3600 to establish the cloud premigration priority threshold. The premigration throttle threshold was determined by the number the FC5274 installed. I only installed 4 FC5274 so that there were reasonable peak and sustained periods in a 6-hour run).

Read Hit and Recall Throughput

Like write activity, there are two types of TS7780T$_{cp1->7}$ read performance: “read-hit” (also referred to as “peak”) and “recall” (also referred to as “read-miss”). A read hit occurs when the data requested by the host is currently in the local disk cache. A recall occurs when the data requested is no longer in the disk cache and must be first read in from physical tape (or from the cloud). Read-hit data rates are typically higher than recall data rates.

TS7780T$_{cp1->7}$ tape recall performance is dependent on several factors that can vary greatly from installation to installation, such as number of physical tape drives, spread of requested logical volumes over physical volumes, location of the logical volumes on the physical volumes, length of the physical media, and the logical volume size. Because these factors are hard to control in the laboratory environment, recall is not part of lab measurement.

TS7780C$_{cp1->7}$ cloud recall depends on TS7780 and IBM COS hardware configuration as well as workload characteristics such as the number of concurrent job counts (see figure 27 for more information on cloud recall performance).

Grid Considerations

Up to six TS7700 clusters can be linked together to form a grid configuration. Seven- and eight-way grid configurations are available via iRPQ. The connection between these clusters is provided by two 1-Gb, four 1-Gb links, two 10-Gb links, or four 10-Gb TCP/IP links. Data written to one TS7700 cluster can be optionally copied to one or more other clusters in the grid.

Data can be copied between the clusters in either deferred, RUN (also known as “Immediate”), or sync mode copy. When using the RUN copy mode the rewind-unload response at job end is held up until the received data is copied to all peer clusters with a RUN copy consistency point. In deferred copy mode data is queued for copying, but the copy does not have to occur prior to job end if DCT is set to zero (default). Deferred copy mode allows for a temporarily higher host data rate than RUN copy mode because copies to the peer cluster(s) can be delayed, which can be useful for meeting peak workload demands. Care must be taken, however, to be certain that there is sufficient recovery time for deferred copy mode so that the deferred copies can be completed prior to the next peak demand. Whether delay occurs and by how much is configurable through the Library Request command. In sync mode copy, data synchronization is up to implicit or explicit sync point granularity across two clusters within a grid configuration. To provide a redundant copy of these items with a zero-recovery point objective (RPO), the sync mode copy function will duplex the host record writes to two clusters simultaneously.

TS7780 Basic Performance

The following sets of graphs show basic TS7780 bandwidths. The graphs show single cluster, standalone configurations. Unless otherwise stated, the performance metric shown in these and all other data rate charts in this paper is host-view (uncompressed) MB/sec.

TS7780 XSD Standalone Performance

To-Do

All runs were made with 64 concurrent jobs, each job writing and/or reading 2 GB (with 1:1 compression) or 10.7 GB (with 5.35:1 compression) using 32KiB blocks, QSAM BUFNO = 20, using eight 32Gb (8x32Gb).

Notes:

Mixed 1:1 workload refers to a host pattern made up of 50% jobs which read and 50% jobs which write. The resulting read and write activity measured in the TS7780 varied and was rarely exactly 50/50.

Mixed 5:35:1 workload refers to host pattern made up of 25% read jobs and 75% write jobs.

TS7780 XFD Standalone Performance

To-Do

TS7780T$_{cp1}$ XSD Standalone Performance

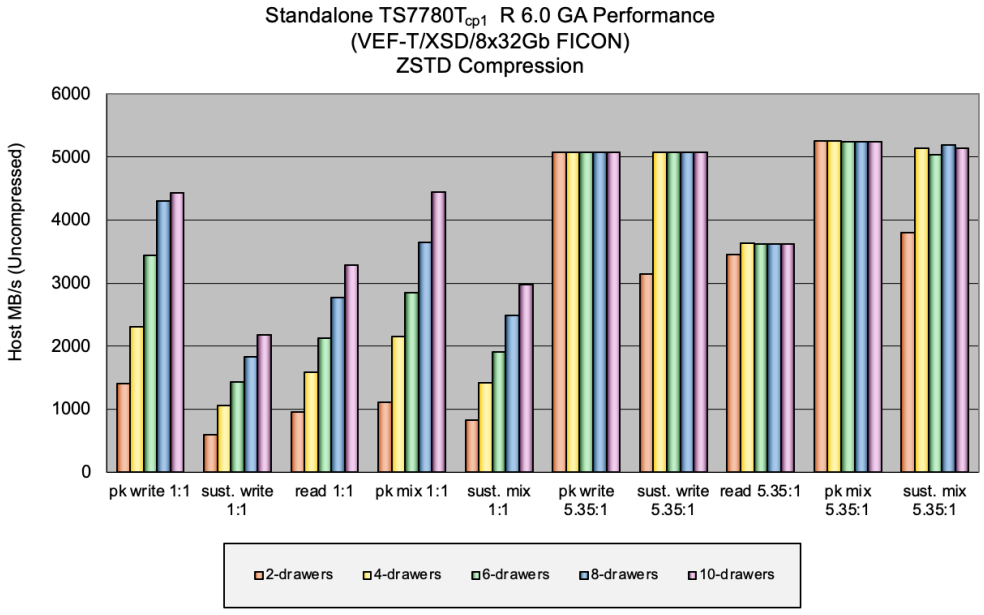
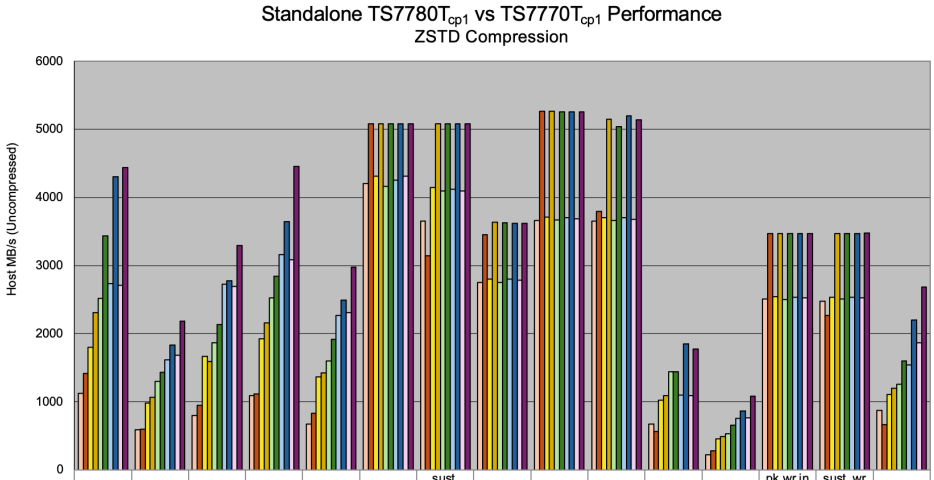


Figure 1: TS7780Tcp1 Standalone Maximum Host Throughput.

All runs were made with 64 concurrent jobs, each job writing and/or reading 2 GB (with 1:1 compression) or 10.7 GB (with 5.35:1 compression) using 32KiB blocks, QSAM BUFNO = 20, using eight 32Gb (8x32Gb).



	pk write 1:1	sust. write 1:1	read 1:1	pk mix 1:1	sust. mix 1:1	pk write 5.35:1	sust. write 5.35:1	read 5.35:1	pk mix 5.35:1	sust. mix 5.35:1	pk wr in mix 1:1	sust. wr in mix 1:1	pk wr in mix 5.35:1	sust. wr in mix 5.35:1	pre-mig
TS7770 - 2 Drawer	1119	589	797	1088	671	4199	3650	2747	3663	3649	674	219	2507	2473	871
TS7780 - 2 Drawer	1413	595	951	1114	827	5079	3142	3450	5260	3797	566	276	3471	2270	663
TS7770 - 4 Drawer	1802	980	1666	1920	1365	4311	4143	2798	3709	3706	1022	452	2540	2536	1107
TS7780 - 4 Drawer	2307	1066	1592	2155	1424	5078	5078	3636	5259	5147	1092	485	3467	3467	1194
TS7770 - 6 Drawer	2518	1298	1866	2524	1595	4163	4098	2751	3668	3663	1444	528	2503	2506	1258
TS7780 - 6 Drawer	3439	1435	2130	2845	1915	5081	5081	3624	5251	5040	1442	655	3469	3467	1601
TS7770 - 8 Drawer	2733	1612	2724	3158	2268	4249	4120	2797	3706	3699	1097	758	2537	2531	1539
TS7780 - 8 Drawer	4306	1831	2774	3643	2488	5076	5079	3620	5251	5198	1846	864	3466	3466	2203
TS7770 - 10 Drawer	2709	1682	2694	3084	2307	4311	4097	2782	3687	3677	1090	768	2526	2524	1865
TS7780 - 10 Drawer	4436	2182	3293	4452	2979	5081	5080	3620	5251	5140	1773	1079	3469	3473	2681

TS7770 Machine Specs: VED-T/12 EH8/2x10Gb/8x16Gb - CP1

TS7780 Machine Specs: VEF-T/ XSD/12 TS1160/8x32Gb/z16- CP1

***Figure 2**: Figure 1: TS7780Tcp1 Standalone Maximum Host Throughput- TS7780 versus TS7770*

Notes:

Mixed 1:1 workload refers to a host pattern made up of 50% jobs which read and 50% jobs which write. The resulting read and write activity measured in the TS7780T varied and was rarely exactly 50/50.

Mixed 5:35:1 workload refers to host pattern made up of 25% read jobs and 75% write jobs.

TS7780T$_{cp1}$ XFD Standalone Performance

To-Do

TS7780C$_{cp1}$ XSD Standalone Performance

To-Do

All runs were made with 64 concurrent jobs, each job writing and/or reading 2 GB (with 1:1 compression) or 10.7 GB (with 5.35:1 compression) using 32KiB blocks, QSAM BUFNO = 20, using eight 32Gb (8x32Gb).

Notes:

Mixed 1:1 workload refers to a host pattern made up of 50% jobs which read and 50% jobs which write. The resulting read and write activity measured in the TS7780C varied and was rarely exactly 50/50.

Mixed 5:35:1 workload refers to host pattern made up of 25% read jobs and 75% write jobs

TS7780DA$_{cp1}$ XSD Standalone Performance

To-Do

TS7780 Grid Performance

Figures 4 through 9 display the performance for TS7780 grid configurations.

For these charts “D” stands for deferred copy mode, “S” stands for sync mode copy and “R” stands for RUN (immediate) copy mode. For example, in Figure 4, RR represents RUN for cluster 0 and RUN for cluster 1. SS refers to synchronous copies for both clusters.

All measurements for these graphs were made at zero or near-zero distance between clusters.

Two-way TS7780 Grid with Single Active Cluster Performance

To-Do

Unless otherwise stated, all runs were made with 64 concurrent jobs, each job writing 10.7 GB (2 GB volumes @ 5.35:1 compression) using 32KiB block size, QSAM BUFFNO = 20, using eight 32Gb FICON channels from z16 LPAR. Clusters are located at zero or near zero distance to each other in laboratory setup. DCT=125.

Notes:

SDT/AES-256 (Security Data Transfer with TLS 1.2 AES256): Encrypted user data for grid replication.

Two-way TS7780 T_{cp1} Grid with Single Active Cluster Performance

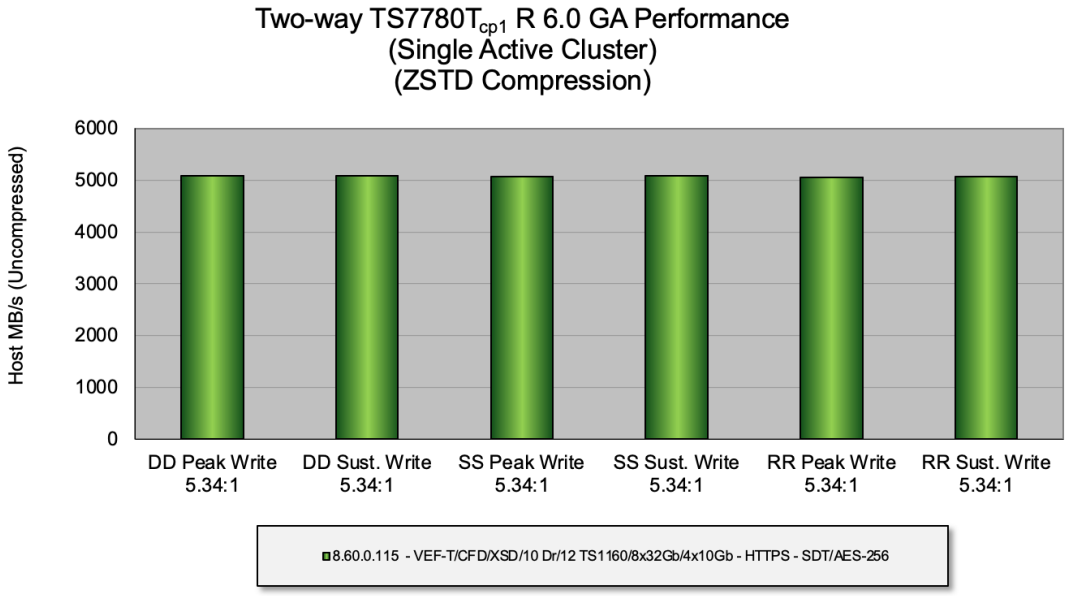
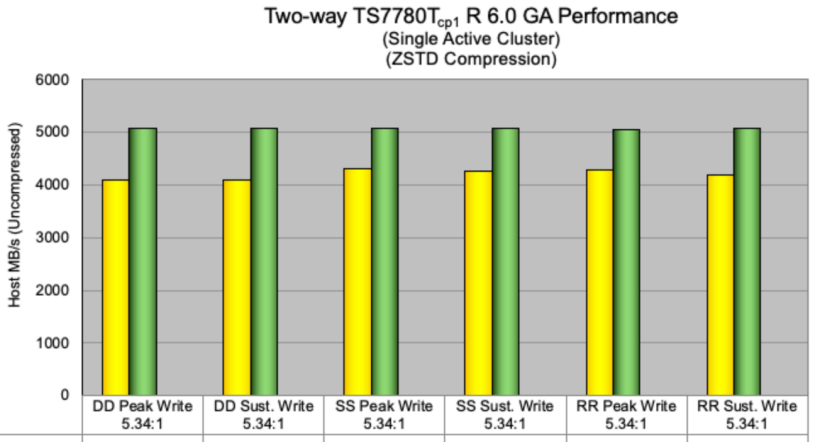


***Figure 3**: Two-way TS7780Tcp1 Single Active Maximum Host Throughput*



	DD Peak Write 5.34:1	DD Sust. Write 5.34:1	SS Peak Write 5.34:1	SS Sust. Write 5.34:1	RR Peak Write 5.34:1	RR Sust. Write 5.34:1
R 5.1 - VED-T/CSB-S/10 DRs/12EH8/8x16Gb/2x10Gb	4094	4094	4316	4254	4290	4180
R 6.0 - VEF-T/CFD/XSD/10 Dr/12 TS1160/8x32Gb/4x10Gb - HTTPS - SDT/AES-256	5078	5080	5076	5078	5058	5077
perf change (TS7780T-R 6.0 vs TS7770T-R 5.4)	24%	24%	18%	19%	18%	21%

***Figure 4**: Two-way TS7780Tcp1 Single Active Maximum Host Throughput R6.0 versus R 5.1*

Unless otherwise stated, all runs were made with 64 concurrent jobs, each job writing 10.7 GB (2 GB volumes @ 5.35:1 compression) using 32KiB block size, QSAM BUFFNO = 20, using eight 32Gb FICON channels from z16 LPAR. Clusters are located at zero or near zero distance to each other in laboratory setup. DCT=125.

Notes:

SDT/AES-256 (Security Data Transfer with TLS 1.2 AES256): Encrypted user data for grid replication.

Two-way TS7780 C_{cp1} Grid with Single Active Cluster Performance

To-Do

Unless otherwise stated, all runs were made with 64 concurrent jobs, each job writing 10.7 GB (2 GB volumes @ 5.35:1 compression) using 32KiB block size, QSAM BUFFNO = 20, using eight 32Gb FICON channels from z16 LPAR. Clusters are located at zero or near zero distance to each other in laboratory setup. DCT=125.

Notes:

* HTTPs: Communication protocol between TS7780C and cloud.

* SDT/AES-256 (Security Data Transfer with TLS 1.2 AES256): Encrypted user data for grid replication.

Two-way TS7700 Grid with Dual Active Clusters Performance

To-Do

Unless otherwise stated, all runs were made with 128 concurrent jobs (64 jobs per active cluster), each job writing 10.7 GB (2 GB volumes @ 5.35:1 compression) using 32KiB block size, QSAM BUFFNO = 20, using eight 32Gb FICON channels from a z16 LPAR. Clusters are located at zero or near zero distance to each other in laboratory setup. DCT=125.

Notes:

SDT/AES-256 (Security Data Transfer with TLS 1.2 AES256): Encrypted user data for grid replication.

Two-way TS7700 T_{cp1} Grid with Dual Active Clusters Performance

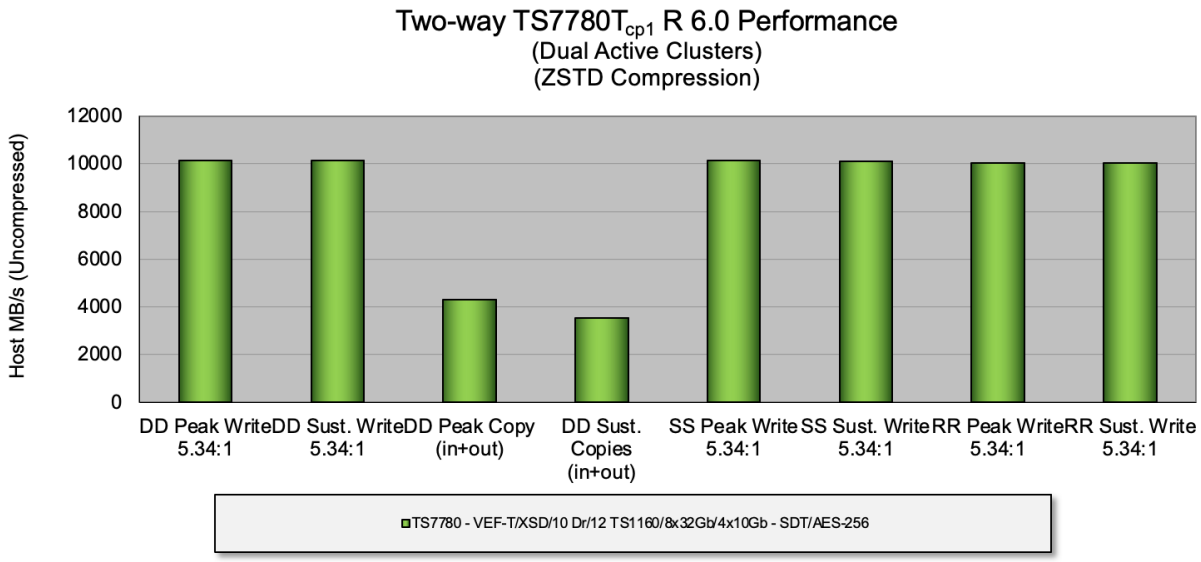
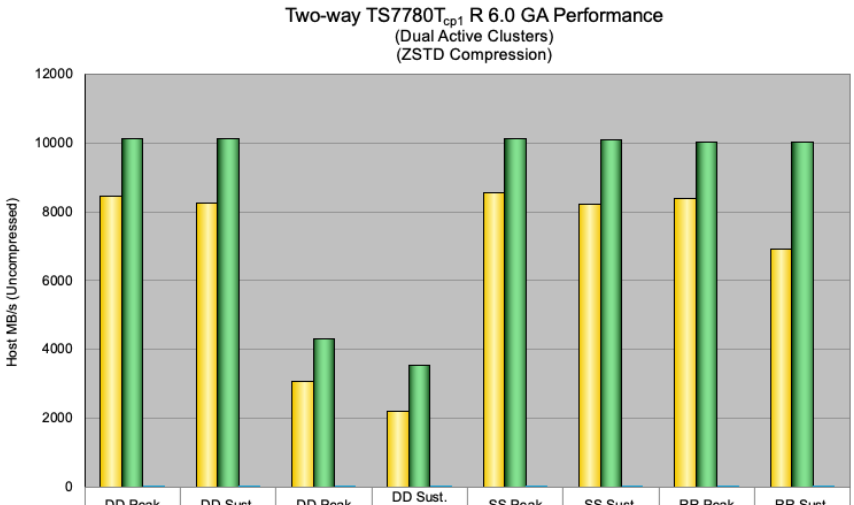


***Figure 5**: Two-way TS7780Tcp1 Dual Active Maximum Host Throughput*

Unless otherwise stated, all runs were made with 128 concurrent jobs (64 jobs per active cluster), each job writing 10.7 GB (2 GB volumes @ 5.35:1 compression) using 32 KiB block size, QSAM BUFFNO = 20, using eight 32Gb FICON channels from a z16 LPAR. Clusters are located at zero or near zero distance to each other in laboratory setup. DCT=125.



	DD Peak Write 5.34:1	DD Sust. Write 5.34:1	DD Peak Copy (in+out)	DD Sust. Copies (in+out)	SS Peak Write 5.34:1	SS Sust. Write 5.34:1	RR Peak Write 5.34:1	RR Sust. Write 5.34:1
R 5.1 VED-T/CSB-S/10 DRs/8x16Gb/12EH8/2x10Gb	8462	8265	3082	2197	8557	8236	8397	6926
R 6.0 - VEF-T/CFD/XSD/10 Dr/12 TS1160/8x32Gb/4x10Gb - SDT/AES-256	10126	10144	4309	3545	10125	10088	10029	10019
perf change (TS7780T-8.60.0.115 vs TS7770T-R5.1)	20%	23%	40%	61%	18%	22%	19%	45%

***Figure 6**: Two-way TS7780Tcp1 Dual Active R 6.0 versus R 5.1*

Notes:

SDT/AES-256 (Security Data Transfer with TLS 1.2 AES256): Encrypted user data for grid replication.

Two-way TS7700 C$_{cp1}$ Grid with Dual Active Clusters Performance

To-Do

Unless otherwise stated, all runs were made with 128 concurrent jobs (64 jobs per active cluster), each job writing 10.7 GB (2 GB volumes @ 5.35:1 compression) using 32KiB block size, QSAM BUFFNO = 20, using eight 32Gb FICON channels from a z16 LPAR. Clusters are located at zero or near zero distance to each other in laboratory setup. DCT=125.

Notes:

* HTTPs: Communication protocol between TS7780C and cloud.

* SDT/AES-256 (Security Data Transfer with TLS 1.2 AES256): Encrypted user data for grid replication.

Four-way TS7700 Grid with Single Active Clusters Performance

To-Do

Four-way TS7700 Grid with Dual Active Clusters Performance

To-Do

Additional Performance Metrics

TS7780 Performance vs. FICON Channel Configuration

The below figure shows how the number and/or configuration of the FICON channels affects host throughput.

To-Do

Performance vs. Block Size and Number of Concurrent Jobs

The below figure shows data rates on a standalone TS7780- VEF/CFD/XSD/10 drawers/8x32Gb FICON with different job counts driven from a z16 host using different channel block sizes.

To-Do

Conclusions

The TS7700 has provided significant performance improvement, increased capacity, and new functionality over the years. Release 4.1 introduced 16Gb FICON channel support which increased the maximum channel performance from 2500 MB/s to over 4000 MB/s. Release R 4.1.2 introduced software compression LZ4 and ZSTD which increase the compression ratio very significantly as compared to the traditional hardware compression at the FICON adapter level (FICON compression). Release 4.2 introduced cloud support. Release 5.0 introduced new TS7770, TS7770T, and TS7770C models with new Power 9 server and V5000 cache. Release 5.2 introduced new TS7770, TS7770T and TS7770C models with new SSD cache (CFD/XSD) that improve cache performance very dramatically with CFD/1 drawer outperform the prior CSB/10 drawer (HDD disk cache). Release 5.2 the CFD only supports 1- and 2-drawer configurations. Release 5.3 added support for CFD 3- and 4-drawer configurations. Release 6.0 supports up to 10-drawer configuration. When evaluating the number of drawers required for a given box, consider recent advancements in flash capacity. This ensures sufficient spindle count for optimal performance while also helping to identify the most cost-effective solution. The TS7700 architecture provides a base for product growth in both performance and functionality.

Appendix 1: Generalizing a Grid Configuration by Using Units of Work

The TS7700 performance behavior, including cloud object storage data, depends on the configuration. First, it is important to understand the data flow within a TS7700 grid.

Figure 7 shows a sample data flow in a two-cluster grid that consists of a TS7700C (Cluster 0 (CL0)) and a TS7700T (cluster 1(CL1)). It is an example of a near worst case scenario in which all data is replicated and premigrated to tape and cloud.

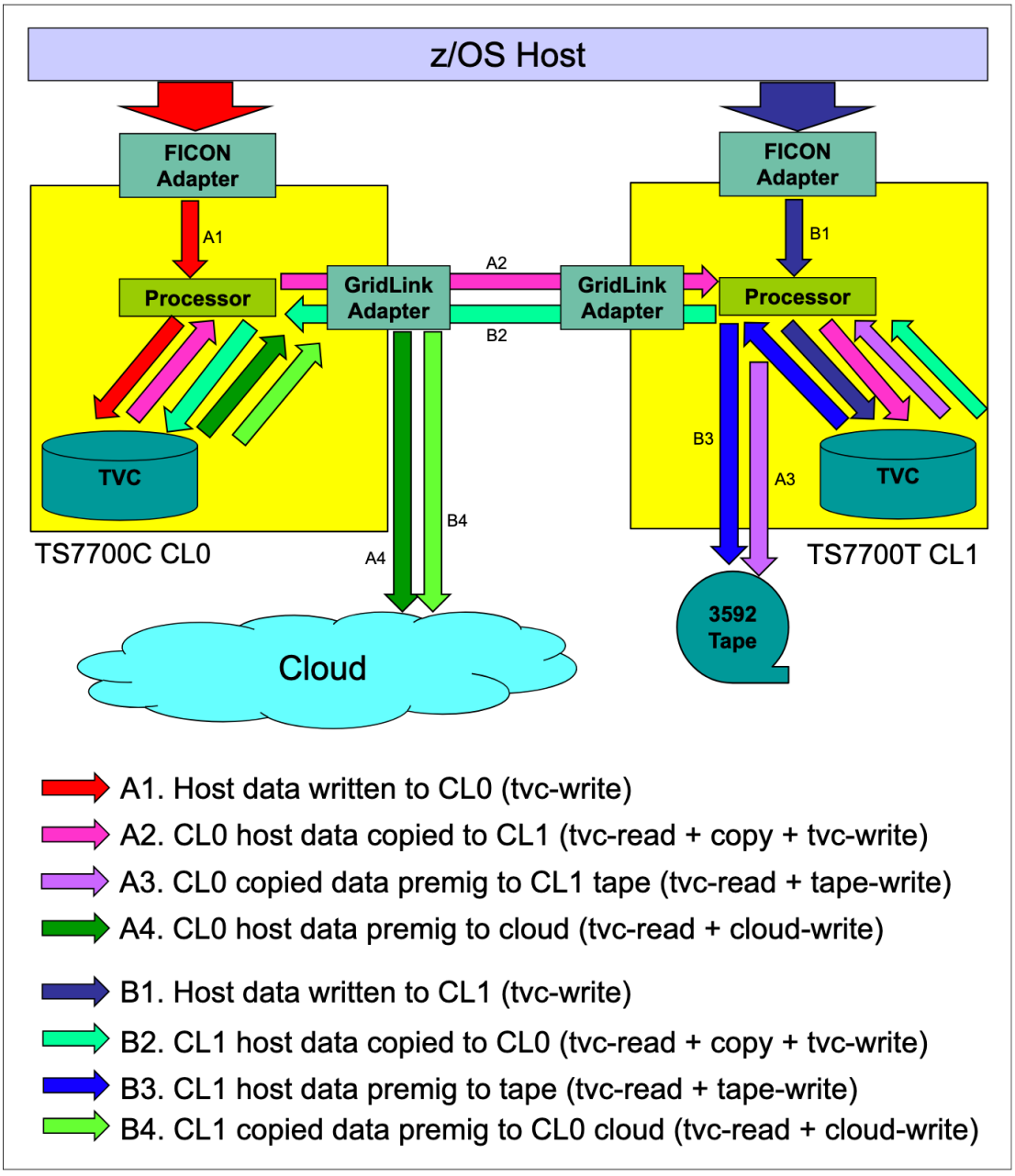


***Figure 7**: TS7700 Data Flow Sample in a Two-Cluster Grid*

The goal of this example is to inform you of how disk cache or tape volume cache (TVC) disk cache cumulative throughputs can be a factor in configuration's performance. In addition, total bandwidth on the grid links can be a factor in the configuration's throughputs. The example that is shown in Figure 7 attempts to break each activity to and from the disk TVC disk and each activity on the grid network as units of work.

This list describes the assumptions of the example in **Figure 7**:

- Each cluster receives its own 300 MBps decompressed from its connected hosts.
- All logical volumes (LVOLs) include RUN or Deferred copy mode with a zero deferred copy throttling (DCT).
- All LVOLs are premigrated to cloud on CL0 and 3592 tape on CL1.
- The data compression ratio is 3:1.
- None to minimal LVOLs are read from the host in this example.

With a 300 MBps channel speed, the A1 and B1 units of work each are 100 MBps after compression. If all things are at equilibrium in a sustained state of operation, each arrow or unit of work must match the 100 MBps throughput.

CL0 includes five total arrows (TVC reads and writes) coming into or out of its TVC disk cache. Therefore, its disk cache must sustain a total of 500 MBps of raw compressed 1:1 mixed read/write throughput. CL1 also must sustain the same rate because it also has five units of work or arrows into and out of the disk cache.

If DCT is enabled, the replication component can be deferred allowing fewer units of work into and out of the TVC disk cache. Premigration to the cloud or tape might be delayed or skipped, which also reduces the total demand on disk cache throughput.

The most complex grid configurations can be generalized by using this basic unit of work concept. It can help determine whether disk cache is potentially a performance limiter.

By using the TS7700 performance white paper, you can determine the maximum mixed 1:1 throughput of your disk cache configuration that is based on how many physical drawers are installed. It can then be used to determine the expected maximum sustained states of operation of the solution. If remote copies to a third location are also occurring, those copies also add units of work to the TVC disk cache and grid network links.

For this same example, the total units of work on the CL0 grid network are four: one outbound for replication, one inbound for replication, and two for outbound cloud premigration. The cumulative read/write rate of the grid network at CL0 must be 400 MBps to sustain the worst-case scenario in this example. Again, limiting replication or deferring or skipping premigration can reduce the workload on the links.

Acknowledgements

The author would like to thank Joseph Swingler, Ramon Antonio Minjares Campos, Josue Cazarez Aguilar, Christine Teale, and Vanessa Sobik for their review comments and insight.

The author would also like to thank Kymberly Slade, Dennis Martinez, Albert Veerland, George Venech, and Jeffrey Watson for hardware/zOS/network support.

IBM Systems

9000 South Rita Road

Tucson, AZ 85744

Printed in the United States of America

04-2025