

TS7700 Technical Update

TS7720 Tape Attach Deep Dive

Ralph Beeston
TS7700 Architecture
IBM

#SHAREorg



SHARE is an independent volunteer-run information technology association that provides **education, professional networking and industry influence.**

Copyright (c) 2015 by SHARE Inc.  Except where otherwise noted, this work is licensed under <http://creativecommons.org/licenses/by-nc-sa/3.0/>



Session objectives

- Brief Overview TS7700
 - Quick background of TS7700
 - TS7720T Overview
- TS7720T Deep Dive
 - TS7720T Introduction Points
 - TS7720T Expected Workload Types
 - Management Interface
 - Setting up partitions
 - Changing partition settings
 - Assigning workloads to partitions
 - TS7720T Premigration Concepts
 - How does it work?
 - How to monitor?
 - Changing a volume's partition
 - Overcommit a partition
 - Flash Copy for Disaster Recovery Testing
 - MES – Upgrade existing TS7700 clusters
 - Grid Examples

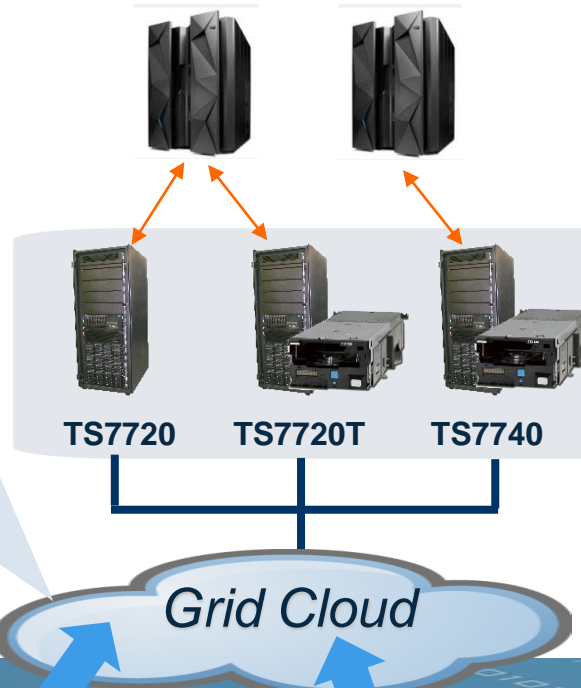
Complete your session evaluations online at www.SHARE.org/Orlando-Eval

IBM TS7700 System Z Virtualization Engine

Leveraging Grid as Cloud Storage for System z

Grid Cloud

- 8Gb FICON
- Disk Speeds
- TS1100 Tape Integration
- Transparent Business Continuance
- AES 256 bit Encryption



Cumulative FICON throughput of over 13GB/s

System z hosts view up to 2,976 equivalent devices

Grid access to all data independent of where it exists

TS7700 Solutions

- **TS7720 Virtualization Engine**
 - Access-centric applications (image data, report servers, critical backups, HSM ML2)
 - Cost-efficient applications (HSM, general backups, master-in, master-out, GDGs, archive)
 - **Up to 1PB** disk cache per system prior to compression. (3 PB with 3:1 compression)
 - **Optionally, up to 80PB backend physical tape storage**
 - AES 256bit Encryption (disk cache & tape)
 - 3TB NL-SAS 7.2K RAID6



- **TS7740 Virtualization Engine**
 - Cost-efficient applications (HSM, general backups, master-in, master-out, GDGs, archive)
 - **Up to 28TB** disk cache per system prior to compression
 - **84TB with 3:1 compression**
 - **Up to 80PB** backend physical tape storage
 - **240PB with 3:1 compression**
 - AES 256bit Encryption (disk cache & tape)
 - 600GB SAS 10K RAID6



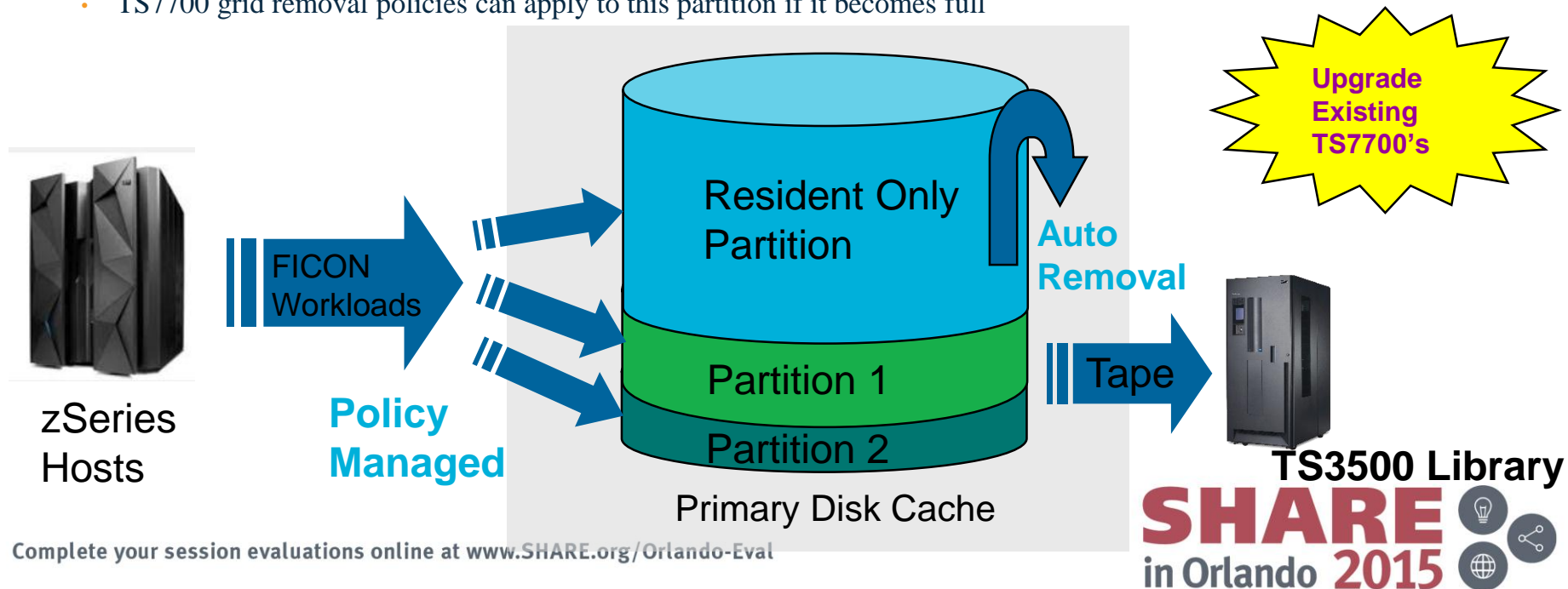
TS7720 Tape Attach (TS7720T)

- **Custom define up to seven tape attached partitions**

- Each partition utilizes a user defined fixed amount of capacity from the total configured available capacity
- Each partition is independently hierarchal storage managed to back end tape
 - Contains its own PG0 and PG1 content.
 - Deferred Premigration (retain copy in disk cache only until time from creation/last-access expires)
 - Perfect for Archive!
- Policy managed as to which partition a workload targets

- **One resident only partition**

- What space remains after creating one to seven tape managed partitions is viewed as the resident only partition
- All data in the resident only data partition remains there indefinitely and doesn't offload to tape
- TS7700 grid removal policies can apply to this partition if it becomes full



TS7720T – Intro Points

- **TS7700 with physical tape and deep cache**
 - Machine type is “VEB” but with physical tape drives and library.
 - Support large TVC (same as existing TS7720) and backend tape library (like TS7740)
- **Logical cache partitions on a cluster**
 - 1 to 7 tape attached partitions (CPx) : as if it is TS7740, but with more flexibility and cache capacity
 - Allows users to be selective about how much content resides in cache for independent workloads
 - One resident only partition (CP0) : as if it is TS7720
- **Enablement**
 - FC5273 : TS7720 Tape Attach enablement
 - Requires no additional hardware within the VEB server (uses existing FC ports reserved for 4th string of disk cache)
 - Can be installed in the field as an MES to an existing VEB based TS7720 (with 1 to 3 existing strings of disk cache)
 - FC5274 : 1TB increments of maximum amount of queued premigration content
 - At least 1 FC5274 is required
 - Up to 10 x FC5274 can be installed.
 - At least 2 increments recommended to match the default settings of a TS7740.
- **Supports all existing functions of both the TS7720 and TS7740**
 - Support Flash Copy for DR testing
 - Support Copy Export and Copy Export Restore/Merge

TS7720T – Four Workload Types

- **Disk Cache Only Content – using CP0**
 - These workloads remain only in disk cache indefinitely
 - Workloads that require a 100% cache hit rate and no local copy on tape
 - e.g. DFSMSHsm ML2 or OAM Object Support
 - May be susceptible to grid removal policies when CP0 becomes full if auto-removal is enabled (enabled by default)
- **Tape Only Content – using CPx and PG0**
 - Use disk cache as pass through
 - Workloads that don't require disk cache residency. Write once, read never...
 - Once data is premigrated to tape and peer replicated, remove the instance from disk cache immediately.
- **Disk Cache and Tape Content – using CPx and PG1**
 - Similar to TS7740 PG1 LRU, but now only relative to common workload targeting the same unique sized CPx partition
 - Workloads don't push unrelated workloads out of cache when they can have their own dedicated partition
 - Partitions can be viewed as a quality of service device with respect to guaranteed workload cache footprint.
- **Delayed to Tape Content – using CPx and delayed-premigration function (new)**
 - Tape industry is evolving.
 - Though physical tape continues to provide value add, customers are becoming more selective as to what and when data moves to physical tape.
 - Delaying data to physical tape provides a grace period allowing only that data which remains after the grace period to be moved to physical tape. Short lived or short retention content never makes it to tape.
 - Physical tape can become an automatic archive target.
 - Similar to Time Delayed Replication
 - Can be used in combination with Time Delayed Replication when configured in a grid

Complete your session evaluations online at www.SHARE.org/Orlando-Eval

R3.2 TS7720T – Why Partitions?

- **Residency only workloads can have dedicated space through CP0**
 - The size of CP0 can be limited in size thus providing a method to limit the amount of residency only data present before auto-removal kicks in
- **Different groups of workloads targeting physical tape may have different cache residency requirements too, thus the tape managed portion of the disk cache can be further divided into partitions too**
 - A given workload may require N days or X TB of content to be resident in cache
 - A dedicated partition can be defined to accommodate such a workload
 - Other workloads with different cache residency requirements can target other partitions without affecting the cache residency of adjacent partitions. For example:
 - Workload A has a 100TB residency requirement, so a 100TB partition A is defined
 - Workload B has a 50TB residency requirement, so a 50TB partition B is defined
 - Workload A creates 10TB of new content a day.
 - Workload B creates 20TB of new content a day
 - Both workloads A and B are setup as PG1.
 - As long as each workload targets their own partition, the 10TB workload for workload A only pushes out workload A content when the partition becomes full. The migration LRU algorithm for partition A is only applicable to workload A data. The same applies for workload B content. Both workloads and their associated daily activity will never alter the adjacent partitions workload's disk cache footprint.

TS7720T Creating/Updating Partitions

+ Create Partition		Actions	Filter		
Partition	Name	Used	Capacity	Volumes	
0	Resident Partition	8.05 TB	82.78 TB	51732	
	Prefer Keep	8.05 TB		51732	
1	General Tape	1.07 TB	50 TB	1032	
	PG1 premigrated	1.07 TB		1032	
2	HSM ML2	3.73 TB	100 TB	711	
3	HSM Backup	0 TB	5 TB	0	
4	Development	0 TB	2 TB	0	

Create Tape Partition

Max 72 TB

Name:

Allocated size: TB

Delayed Premigration Limit: GB

Primary tape partition:

Rename General Tape

Name:

Resize General Tape

Max 122 TB

Allocated size: TB

Delayed Premigration Limit: GB


TS7720T Viewing Partition Properties



Partition	Name	Used	Capacity	Volumes
0	Resident Partition	8.05 TB		
	Prefer Keep	8.05 TB		
1	General Tape	1.07 TB		
	PG1 premigrated	1.07 TB		

- Resize
- Rename
- Designate Primary Partition
- Delete
- Show in binary units
- View Statistics
- Properties

Properties for General Tape




Partition Name [Number]: General Tape [1]
 Volumes: 1032
 Capacity: 50 TB
 Used: 1.07 TB
 % of total cache: 20.85%

Delay Premigration size: 0 TB
 Cache to Copy: 0 TB

PG0 premigrated: 0 TB
 PG1 premigrated: 1.07 TB
 PG0 in premigration queue: 0 TB
 PG1 in premigration queue: 0 TB
 PG0 waiting for delay premigration: 0 TB
 PG1 waiting for delay premigration: 0 TB
 Number of virtual volumes migrated: 2590

[Close](#)

Properties for Resident Partition



Partition Name [Number]: Resident Partition [0]
 Volumes: 51732
 Capacity: 82.78 TB
 Used: 8.05 TB
 % of total cache: 34.52%

Flash copy size: 0 TB
 Cache Removal Threshold: 3 TB
 Cache to Copy: 0 TB

Prefer keep size: 8.05 TB
 Prefer remove size: 0 TB
 Pinned size: 0 TB

[Close](#)

TS7720T Assigning Workloads To Partitions

Create Storage Class ☰ Actions ▼			
Name	Partition	Storage Preference	Premigration Delay Time
-----	Resident Partition(0)	Prefer Keep	
PART1L0	General Tape(1)	Level 0	0 hours
PART1L1	General Tape(1)	Level 1	0 hours
PART2L0	HSM ML2(2)	Level 0	0 hours
PART2L1	HSM ML2(2)	Level 1	0 hours
PART3L0	Resident Partition(0)	Prefer Remove	

Create Storage Class

Example: Tape Partition Attributes

Name:

Description:

Partition:

Tape Volume Cache Preference:

Premigration Delay Time: hours

Premigration Delay Reference:

Create Storage Class

Example: Resident Only Attributes

Name:

Description:

Partition:

Volume Copy Retention Group:

Volume Copy Retention Time: hours

Volume Copy Retention Reference:

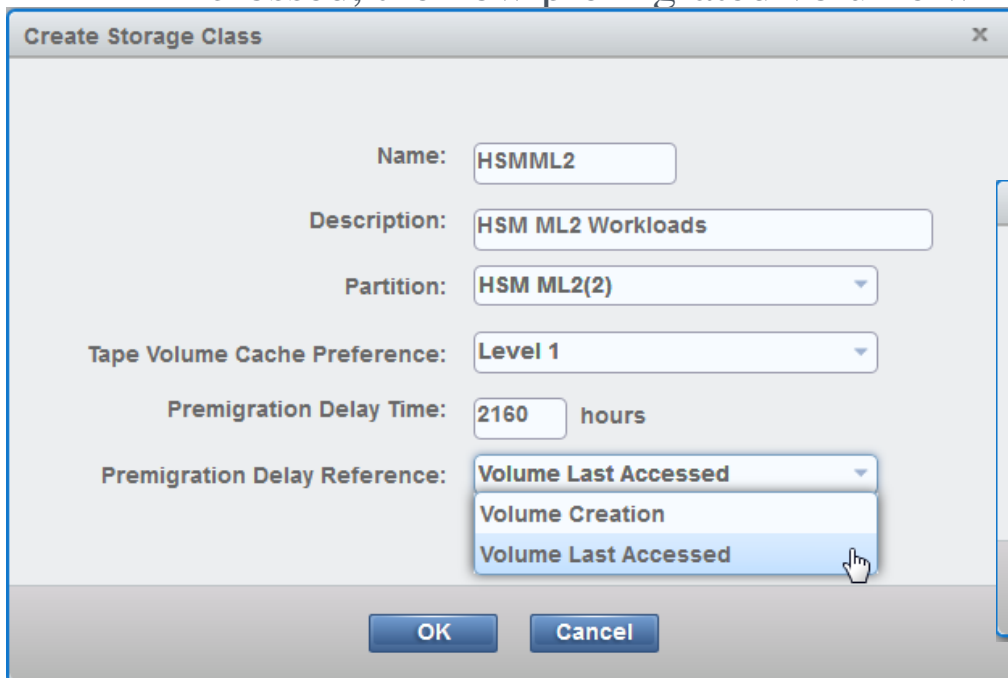
TS7720T Time-delayed Premigration (New Concept!)

- **Move workloads to tape after a timed delay**
 - Retain data only in disk cache until delay criteria is met
 - Delayed criteria in hours from creation or last access
 - Any content which expires prior to delay period never goes to tape
- **Archive made easy**
 - Have workloads automatically age to tape
 - Limited knowledge needed to determine what is considered archive
 - If it still exists after N hours or hasn't been accessed in N hours, then put it on tape.
- **Once on tape, honor tape based policies**
 - PG0/PG1 honored once a copy is put on tape
 - PG0 will be removed from disk cache immediately
 - PG1 will be retained using a lease recently used algorithm based on LRU of CPx's content
- **Limit how much data can be delayed to tape**
 - When sizing a tape partition, a maximum delay premigration limit is defined
 - If delay premigration content exceeds this limit, volumes will begin to move to tape early



TS7720T Time-delayed Premigration

- **Storage Class Attribute**
 - Premigration Delay Time – 0~65535 in hour (0 implies no delay)
 - Premigration Delay Reference – Volume Creation (write from BOT) or Volume Last Accessed (Read or Write)
- **Early expiration of time-delayed premigration threshold per partition**
 - If amount of resident lvols in the targeted CPx partition waiting for the delayed time exceeds the threshold, delayed primgration volumes which have been in the CPx partition the longest will be queued early until threshold no longer exceeded.
- **PG0/PG1 still honored**
 - Once the delay grace period has expired or queued early due to the threshold being crossed, the now premigrated volume will then honor the PG0 or PG1 preference



Create Storage Class

Name:

Description:

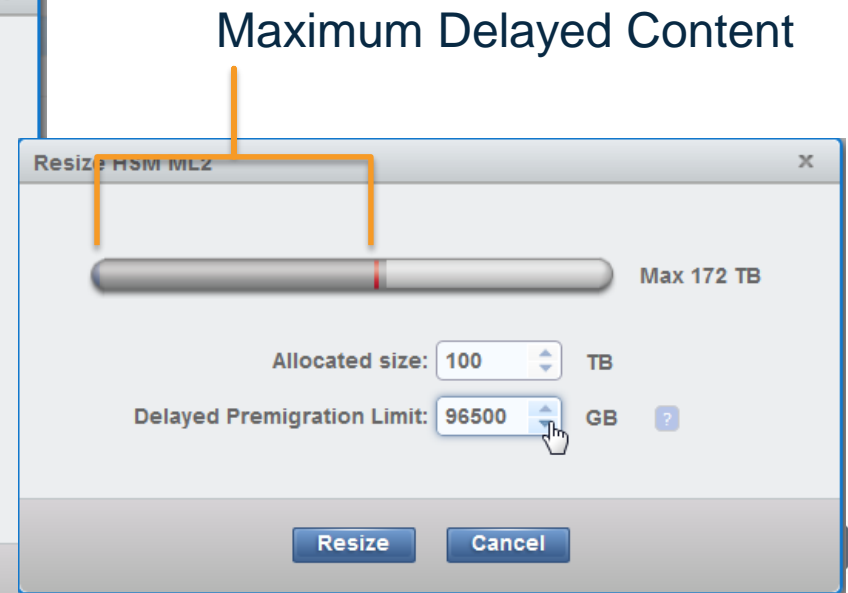
Partition:

Tape Volume Cache Preference:

Premigration Delay Time: hours

Premigration Delay Reference:

Maximum Delayed Content



Resize HSM ML2

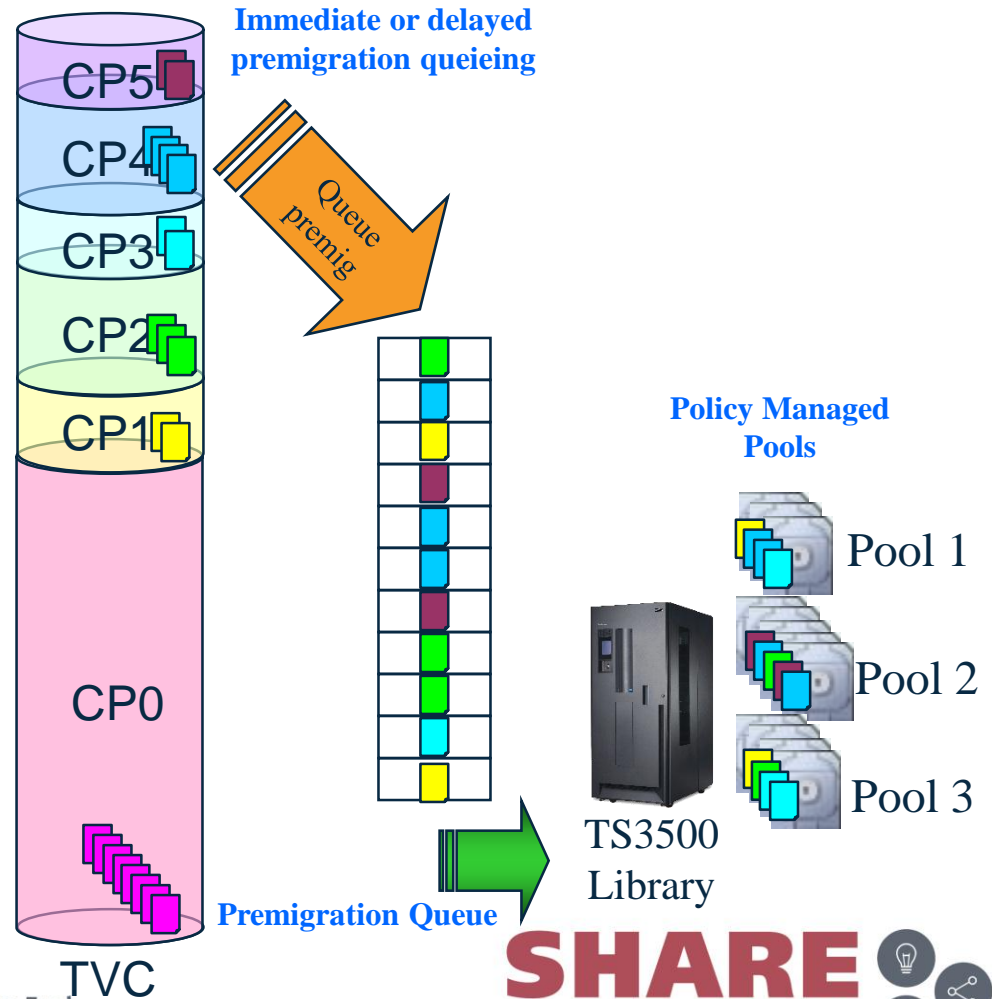
Max 172 TB

Allocated size: TB

Delayed Premigration Limit: GB

TS7720T Cache Partitions and Premigration

- Single premigration queue for all tape manage partitions
- PMTHROT and PMPRIOR thresholds apply to the global cumulative used disk cache space for all tape managed partitions
- If premigration throttling occurs, it's applied to all workloads targeting any CPx partition.
- Premigration queue size is limited by FC5274 Increments
 - Cumulative post-compressed size of all volumes contained in queue
 - PMTHROT is limited to this same limit
- Workloads targeting CP0 are not throttled due to premigration thresholds

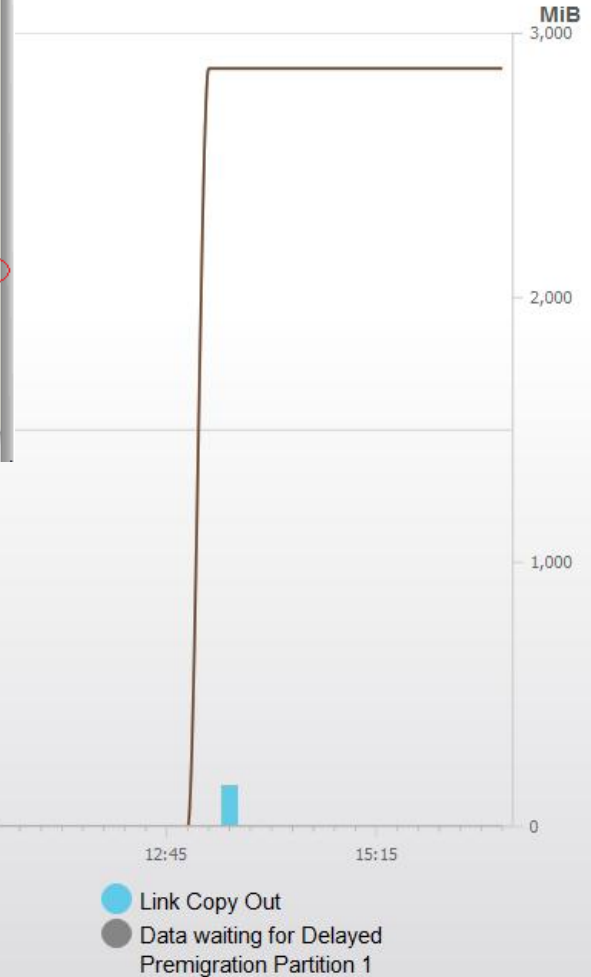


TS7720T – Premigration Size Monitoring

Management Interface Partition Summary

Partition Nu...	Name	Used	Capacity	Volumes
0	Resident Partition	2.34 TiB		
★ 1	Tape Partition 1	0 TiB		
2	Tape Partition 2	0 TiB		

Partition Number	Name	Used	Capacity	Volumes	% of Cache	Premigration Size	Cache to Copy	Delayed Premigration Limit
<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>



Management Interface Historical Summary

Select Metrics	
Throughput	6 Selected
Storage	4 Selected
Partition 5	<input type="checkbox"/> Migrated (MiB)
Partition 6	<input type="checkbox"/> Migrated (MiB)
Partition 7	<input type="checkbox"/> Migrated (MiB)
Data Waiting for	
Partition 1	<input checked="" type="checkbox"/> Premigration (MiB) <input checked="" type="checkbox"/> Delayed Premigration (MiB)
Partition 2	<input type="checkbox"/> Premigration (MiB) <input type="checkbox"/> Delayed Premigration (MiB)

TS7720T – Premigration Size Monitoring

z/OS LIBRARY REQUEST

```
"TAPE VOLUME CACHE STATE V4 .0                                "
```

```
"TS7700 VIRTUALIZATION ENGINE MODEL: TS7720 TAPE ATTACH      "
```

```
"TOTAL INSTALLED/ENABLED GBS: 47850 / 47850                  "
```

```
"TOTAL ADJUSTED CACHE USED GBS: 2630                         "
```

```
"CACHE ENCRYPTION STATUS: CAPABLE                            "
```

```
"PRIMARY CACHE RESIDENT ONLY INFORMATION                     "
```

```
" PRIVATE CACHE USED GBS: 9317                               "
```

```
" SCRATCH CACHE USED GBS: 0                                  "
```

CP	ALLOC	USED	PIN	PKP	PRM	COPY	CPYT
0	4850	2584	541	1672	353	0	0

```
" FLASH COPY INFORMATION                                     "
```

INDEX	ENABLED	SIZE
1	NO	0
2	NO	0
3	NO	0
4	NO	0
5	NO	0
6	NO	0
7	NO	0
8	NO	0

```
"PRIMARY TAPE MANAGED PARTITIONS                             "
```

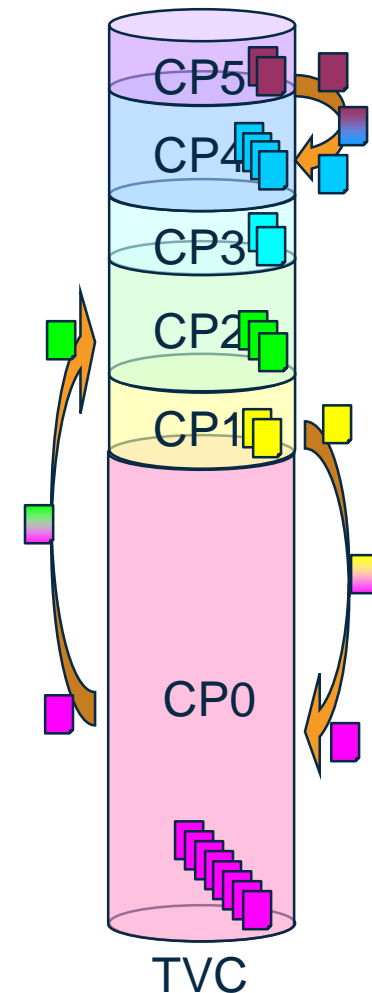
CP	ALLOC	USED	PG0	PG1	PMIGR	D_PMIGR	COPY	PMT	CPYT
1	42000	0	0	0	0	0	0	0	0
2	1000	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0

VEHSTATS Reports

Complete your session evaluations online at www.SHARE.org/Orlando-Eval

TS7720T Partition Assignment Change

- **Partition assignment is defined by storage class**
- **Two methods to change partition assignment for an existing volume that is already stored in the TS7720T**
 - Mount/demount to reflect the new storage class rules
 - LI REQ PARTREFSH (excludes any PVOL pool movement)
- **No actual physical volume data movement when moving to CP0**
 - When changing CPx->CP0 for a migrated lvol, it stays as migrated. When the host mounts the volume, a recall takes place and the physical copy is removed at that time.
 - If CPx->CP0 for a cache resident volume (cache only or pre-migrated), the volume is moved to CP0 and the copy on tape is removed immediately.
 - Changing CP0->CPx, premigration for the volume is queued.
- **Allow overcommitted state**
 - The volumes can be moved even if the target partition exceeds its currently defined size



TS7720T PARTRFSH

- Think COPYRFSH for Partitions!
- Setup a new Storage Class rule
- Issue PARTRFSH to a TS7720T for a specific volume
- The volume will move to the new partition based on the storage class rule

```
LIBRARY REQUEST,VTSCOMP1,PARTRFSH,TD0000
```

```
CBR1280I Library VTSCOMP1 request.
```

```
Keywords: PARTRFSH,TD0000
```

```
-----
```

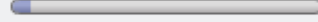


```
-----
```

```
PARTITION REFRESH V1 .0
```

```
UPDATED CACHE PARTITION OF LOGICAL VOLUME TD0000
```

TS7720T Over Commit State

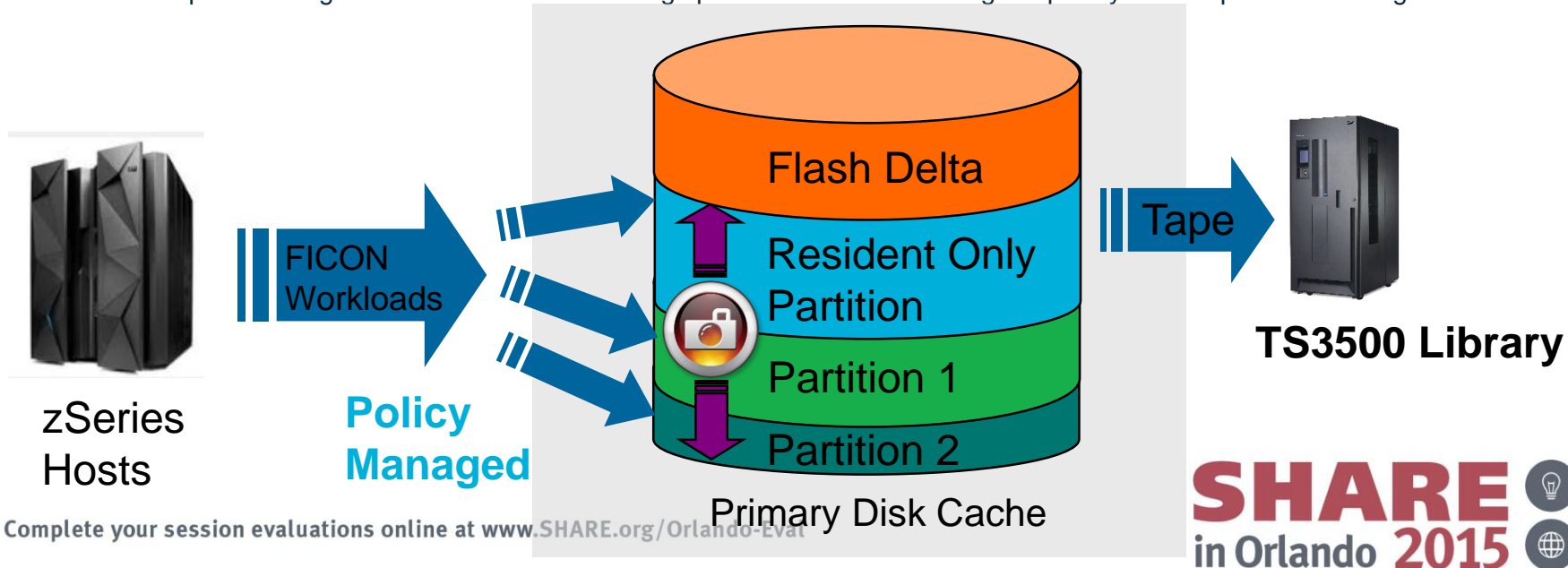
- If CPx used size is greater than configured capacity, the partition is in an over commit state.

Partition Nu...	Name	Used	Capacity
⊕ 0	Resident Partition	2.57 TB	 44.85 TB
⊕ ★ 1	Tape Par	This partition is in an over commit state 	
			 2 TB

- Following cases can cause over commit state for CPx
 - Moving resident or premigrated lvols to CPx via mount/demount or PARTREFSH when CPx is already full
 - Reduce configured capacity of the CPx below its current active size
 - Inbound host or copy content exceeds configured capacity without also reaching PMTHLVL
 - Allows temporary excess until pre-migration and migration catches up
- Amount of data exceeding capacity of all CPx partitions uses free space from CP0 until the over commit state can be resolved.
- Migration works to resolve the over commit state for whichever CPx partitions are over committed.

TS7720T and Flash Copy for DR Testing

- Flash snapshot works across the entire disk cache, including tape managed partitions.
 - ▶ Content only on tape is still accessible via the LIVECOPY attribute (live copy)
 - ▶ LIVECOPY is only applicable if copy was present in TS7720T ahead of time zero or start of DR test snapshot
- CPx partitions will continue to manage content virtually via PG0 and PG1 LRU algorithm
 - ▶ Virtually, it will be contained to CPx size including all delay premigration thresholds
 - ▶ Resulting migration of content written ahead of time zero virtually frees up CPx space, but physically it is still present in the flash region of the file system. Content written after time zero will be physically removed when it becomes migrated. CP0 content present at time zero is also retained within the flash delta.
- Flash Delta content
 - ▶ All delta modifications, deletions through migration and scratch expiration of CPx content is retained within the Flash Delta
 - ▶ Flash Delta borrows capacity from CP0 and thus CP0 available capacity will shrink as delta/migration increases
 - ▶ Worse case is CP0's available space is decreased by the sum of all CPx partition sizes plus CP0 new workload
 - ▶ Space management can be handled through procedures such as using temporary smaller partitions during DR tests



TS7720 to TS7720T MES

- Full support to upgrade existing VEB based TS7720s to TS7720T
 - Requires a service outage (non-concurrent for specific TS7720 being targeted)
 - Any existing data on the TS7720 can remain.
 - Requires a newly defined library partition within the TS3500 with dedicated media and drives
 - TS3500 requires two switches for FC attached drives (same as TS7740)
 - Two 8Gb FC cables are connected from two switches in the TS3500 directly to existing available ports on the TS7720 (using FC ports normally reserved for 4th cache string)
 - TS7720T or Four Disk Cache Strings are Mutually Exclusive
 - If 4 strings of disk cache already exist, TS7720T is not possible
 - Once TS7720T is supported, adding 4th string of disk cache is not possible
 - TS7720T initialize with all previous content in CP0 and an empty 3TB CP1 partition
 - CP1 size can be changed once the TS7720T is online
 - Additional CPx partitions can also be defined once the TS7720T is online
 - All previously created content which resides in CP0 can be moved to CPx partitions via mount/demount requests with policy changes and/or using the LI REQ PARTREFSH command.

TS7740 to TS7720T Frame Swap (push pull) MES

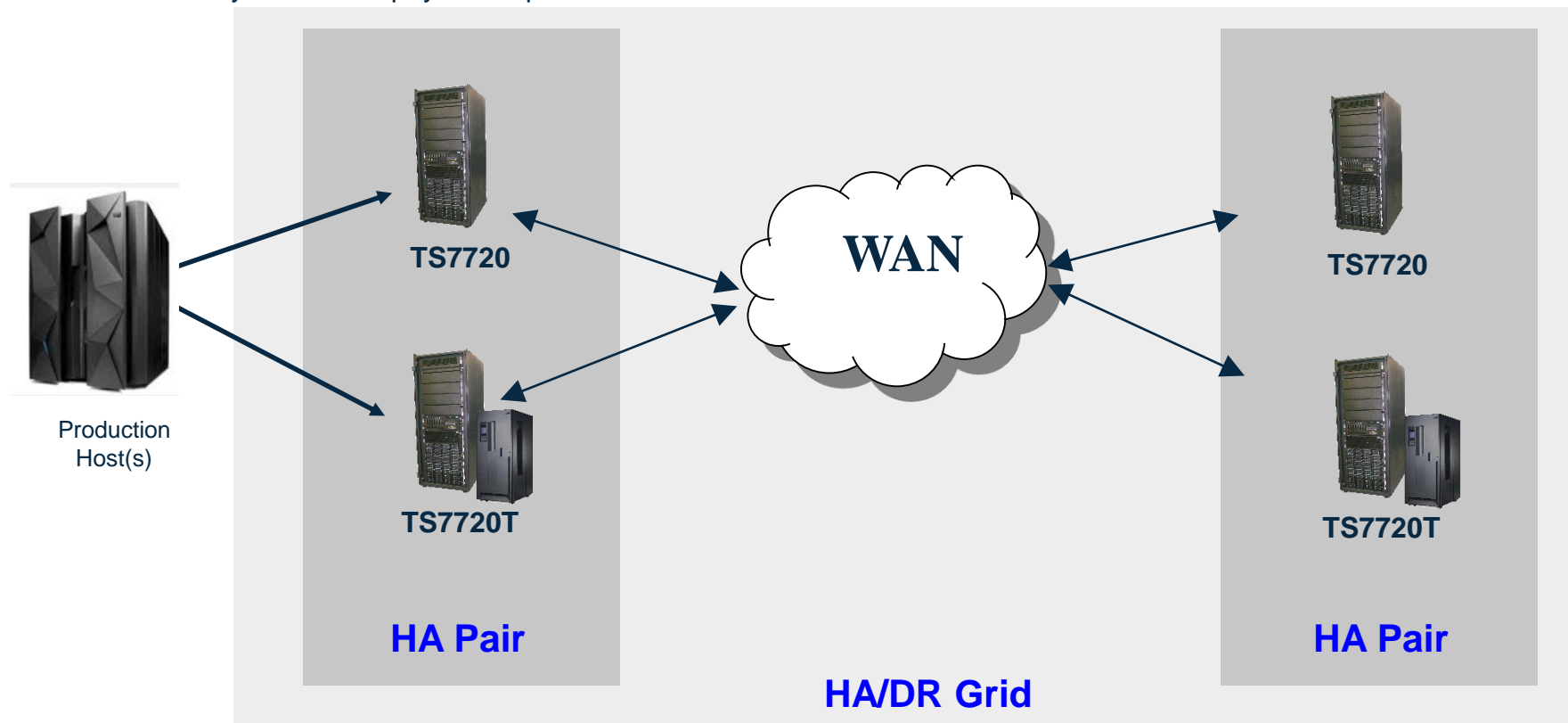
- Replace TS7740 with TS7720T and maintain content on physical tape
 - Requires a service outage (non-concurrent for specific TS7740 being targeted)
 - Only TS3500 attached libraries supported.
 - All existing data in TS7740 cache will be flushed to physical tape
 - Multiple backups of DB and configuration data DVDs
 - Push out TS7740, pull in TS7720T and attach existing FC cables from TS3500 to new TS7720T
 - Restore DB and configuration data from backup DVD
 - The TS7720T will come online with a 3TB CP1 partition and the remaining space assigned to CP0
 - All previous tape content is assigned to CP1 partition and available via recall
 - CP1 size can be altered in size once the TS7720T is online
 - Additional CPx partitions can be defined once the TS7720T is online
 - All previously created content which resides in CP1 can be moved to CP* partitions via mount/demount requests with policy changes and/or using the LI REQ PARTRFSH command.

4-way Hybrid Use Case (Similar 7720/7740)

- TS7720 is configured with large disk cache
- TS7720T is configured with small to large disk cache
- As host data arrives in TS7720, it copies to TS7720T
- As host data arrives in TS7720T, it copies to TS7720
- TS7720 uses auto removal to maintain space
- TS7720T relies on migration to maintain space
- Oldest data only resides on physical tape

Benefits

- Similar behavior to existing 7720/7740 hybrid configurations (some don't like change)
- Similar 4 copies of youngest data
- 4 young copies can all reside in disk cache for a longer period of time (NEW)
 - TS7720T may use a large tape managed partition which will provide much more local disk cache residency for local HA without needing to do a remote mount.

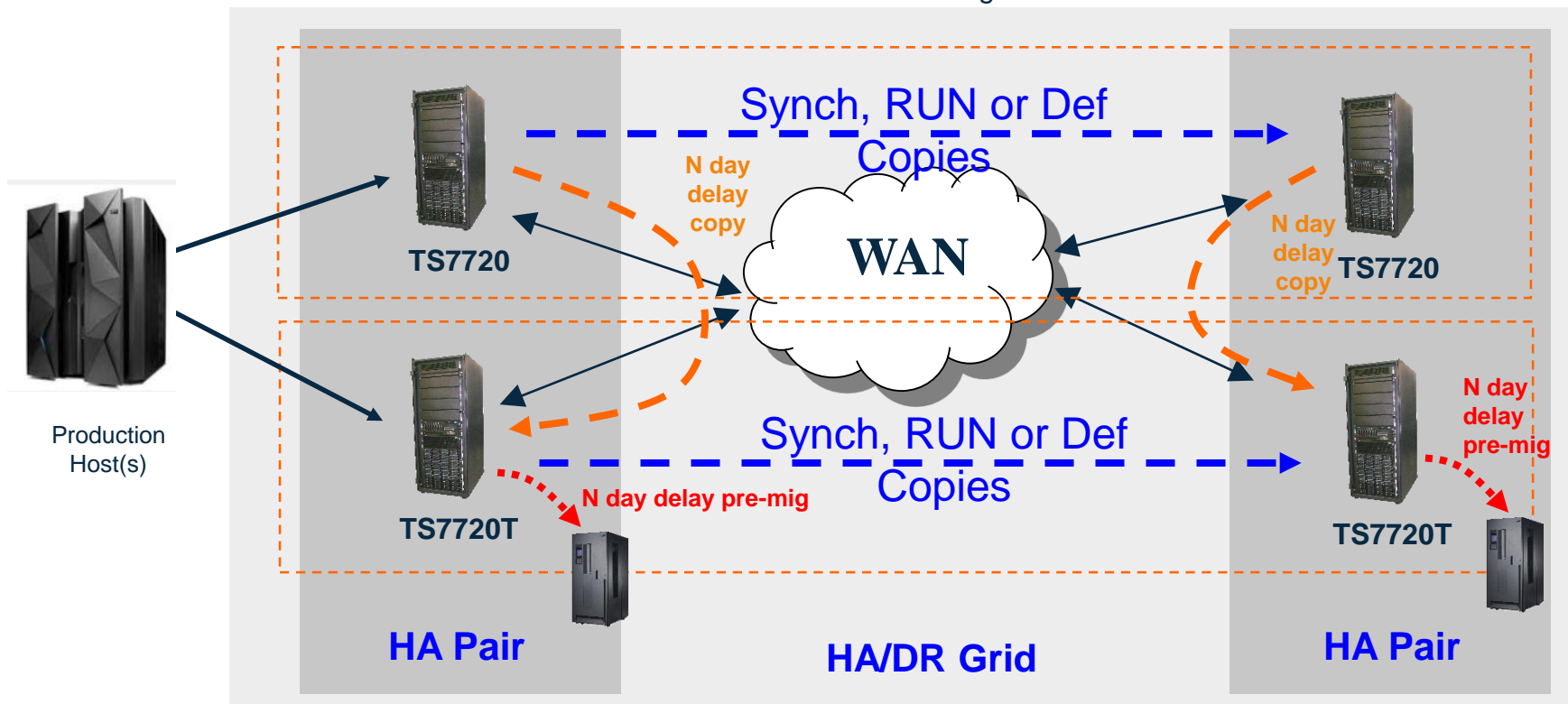


4-way Hybrid Use Case (Data Archive To Tape)

- Both TS7720 and TS7720T are configured with similar disk cache sizes
- New host workload only replicates to paired DR box
- TS7720 does a delay replicate to TS7720T by N days allowing it to auto-remove archive data
- TS7720T delay pre-migs to tape by same N days allowing it to migrate archive data
- Both delay replication and delay pre-mig use same reference point (creation or last access)

Benefits

- Using tape only as archive target
- Same HA and disk cache footprint as a 4-way TS7720 configuration
- No need to put tape behind both TS7720 clusters
- All data new and old has 2 or more copies
- Very attractive to disk only customers who are re-thinking having “some tape” in their configurations.



Q&A?

- Additional Questions?

Excellent References!

- Techdocs
 - <http://www-03.ibm.com/support/techdocs/atmastr.nsf/Web/TechDocs>
 - Search on TS7700
- TS7700 R3.2 Redbook (1st draft)
 - <http://www.redbooks.ibm.com/redpieces/abstracts/sg248122.html>
- TS7700 Infocenters
 - Customer - <http://pic.dhe.ibm.com/infocenter/ts7700/cust/index.jsp>
 - Service - <http://pic.dhe.ibm.com/infocenter/ts7700/serv/index.jsp>

Thank You