Under the Covers

Benefits of Disk Library for Mainframe Tape Replacement

Session 17971

#SHAREorg

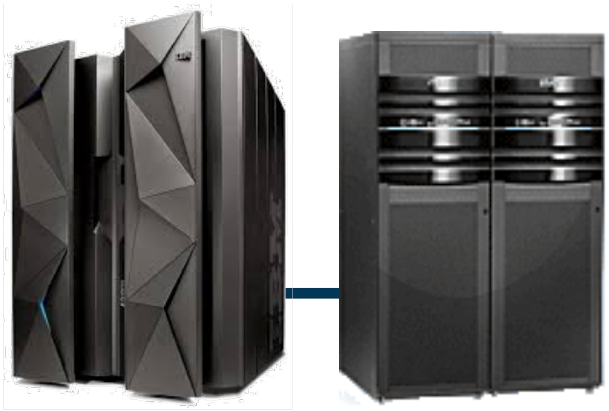
Session Overview

- DLm System Architecture
- Virtual Library Architecture
- VOLSER Handling

  Formats
  Allocating/Mounting
  Scratching
  Fast Locate
  WORM tape
  Data Encryption

- Guaranteed Replication
- Universal Data Consistency

WHAT IS DISK LIBRARY FOR MAINFRAME (DLM)?

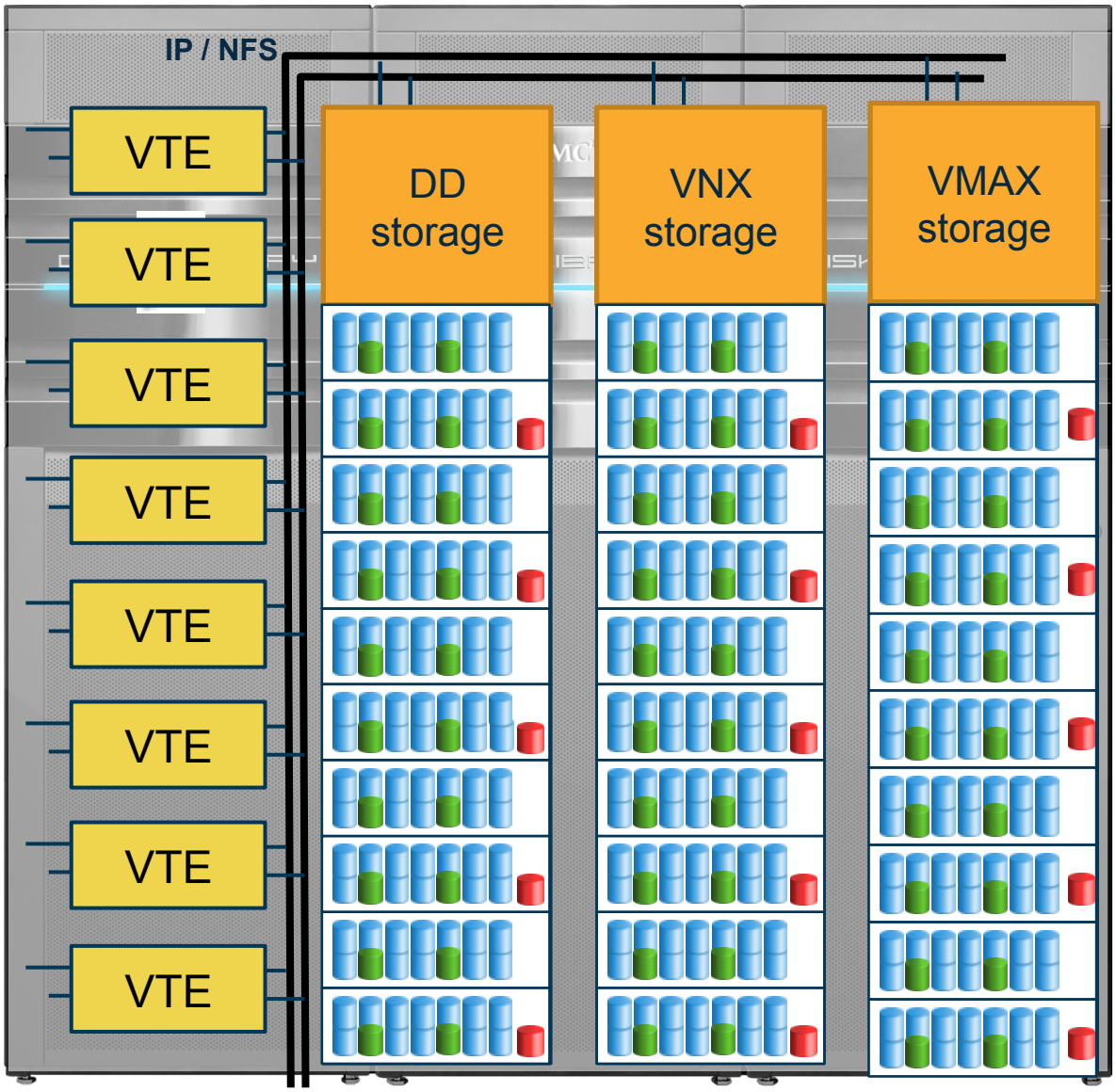


- A virtual tape library for IBM and Unisys mainframes
- DLm emulates tape drives to mainframe and writes tape volume images to disk
- A single DLm can provide up to 16 FICON, 2,048 drives, and > 3PB storage

DLM STARTS WITH EMC'S THREE INDUSTRY LEADING STORAGE OFFERINGS

And allows you to leverage them all

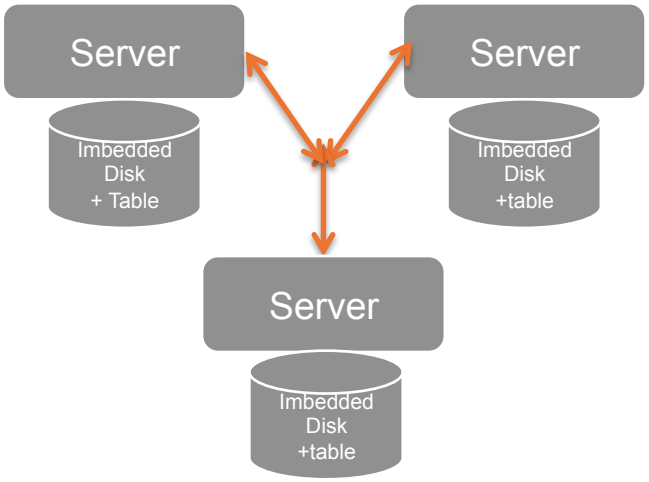
DLM THEN ADDS A RESILIENT REDUNDANT ARCHICHTECTURE



- 2 to 8 VTEs eliminate single point of failure, providing high-availability
- Each VTE delivers 2 8 Gbit FICON and up to 256 tape drives
- All VTEs see all tape volumes
- NFS insures that tape volumes are available through alternate VTEs
- Data is protected by RAID 6 arrays with sparing

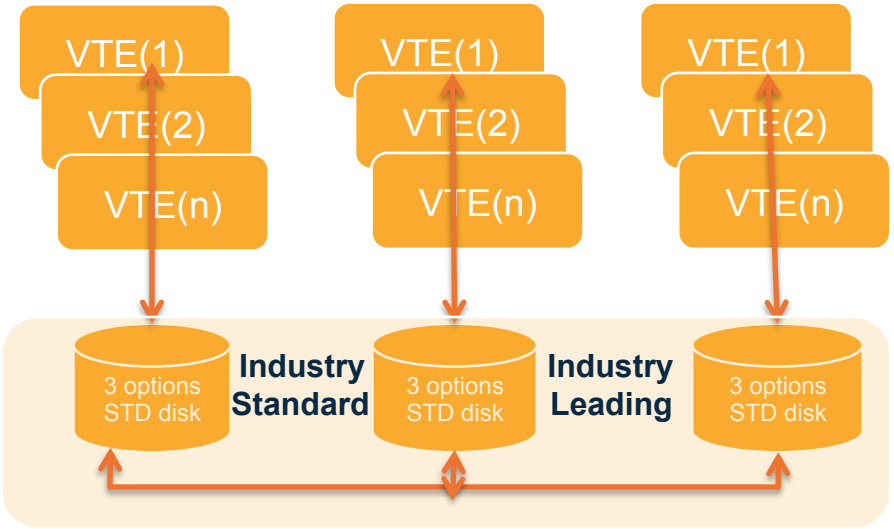
A DIFFERENT VIRTUAL TAPE ARCHITECTURE

Clustered-Server-scaled Architecture



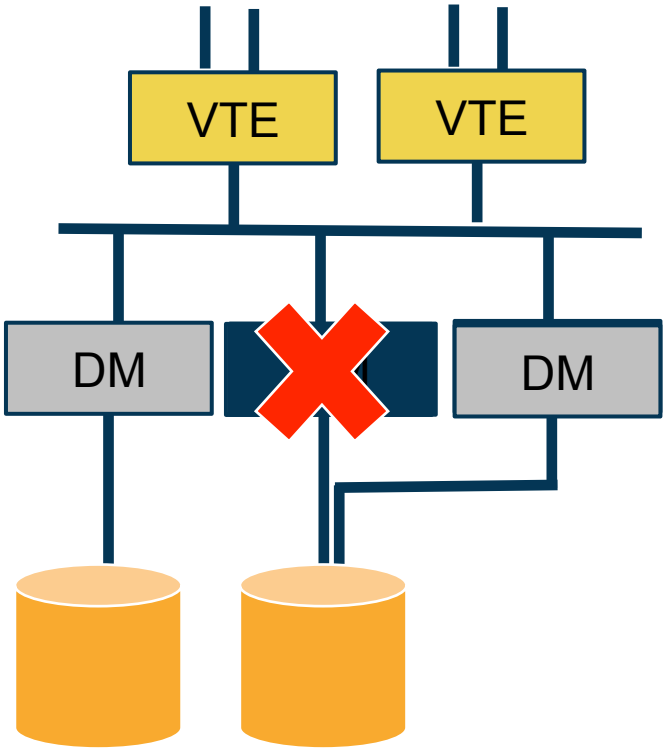
- Individual servers each with dedicated data storage
- Network connectivity between servers for inter-server communication
- Servers can ONLY fail-over to one another, providing availability
- Environment scales only by adding servers **and** storage

Shared-Storage scaled Architecture



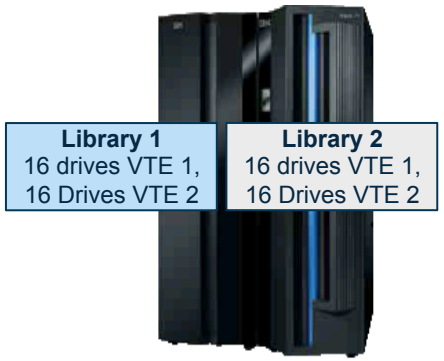
- Individual servers with shared data storage
- Servers operate independent of one another
- Surviving servers continue to operate when one or more servers fail, providing availability
- Environment scales by adding servers **or** storage

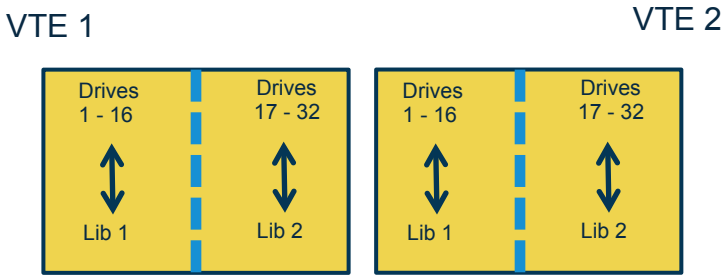
VNX AND VG8 ARCHITECTURE

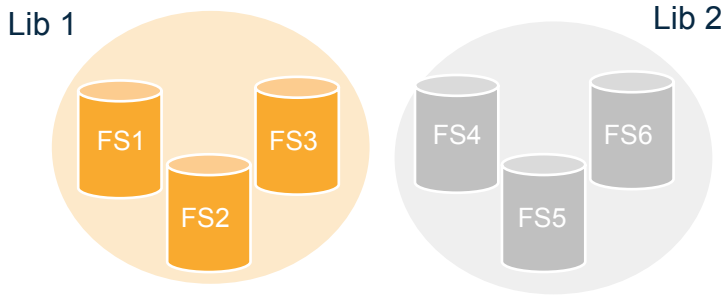


- Up to 8 data movers make up a VNX / VG8
- Each data mover (DM) manages 256 TB raw
- All VTEs can access all storage managed by all data movers
- A standby DM will take over for a failed unit
- No disruption of service occurs

VIRTUAL LIBRARY ARCHITECTURE

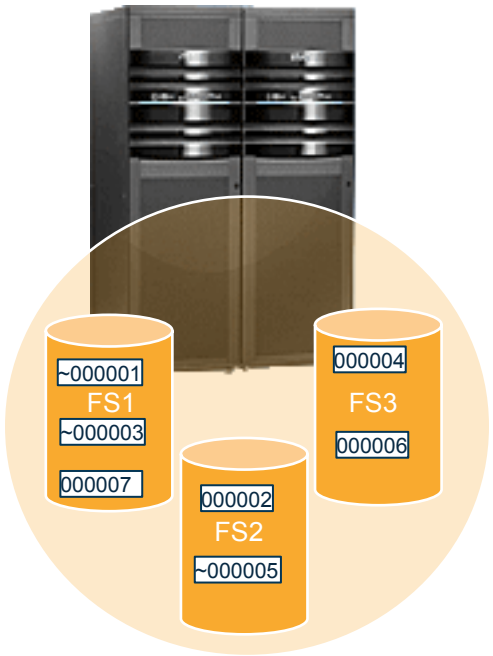






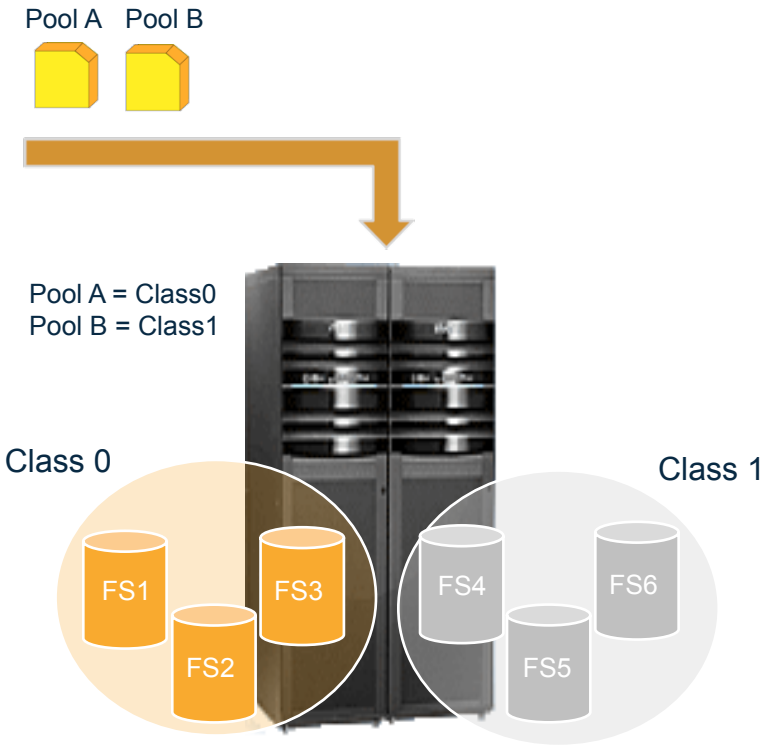
- Multiple filesystems (directories) are defined on the storage.
- filesystems are grouped into virtual libraries
- Each drive in a VTE points to 1 and only 1 library.
- Drives in each VTE point to each library to provide high-availability
- A single DLm8100 can emulate up to 2048 drives and multiple virtual libraries

VOLUMES IN THE LIBRARY



- No database / index is kept on storage
- Each VOLSER is single file within a filesystem
- Volumes are named with their serial number
- Scratch volumes are identified by “~”
- Initialize (init) command creates scratch volumes in library with VOL1 record.
- New allocations go to any filesystem in the class using 1 of 2 algorithms
  - Round-Robin
  - Most Freespace
- Scratch tapes will be moved between filesystems within class as needed.

STORAGE CLASSES



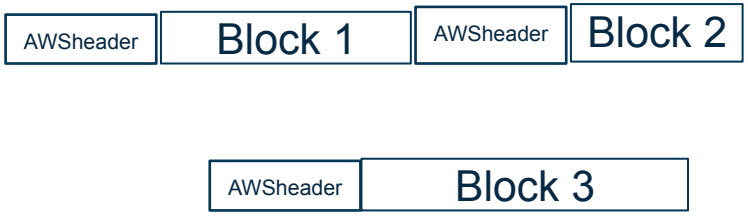
- Filesystems in a library are assigned to a storage class.
- TMS tape pools point to a storage class
- Allowing individual tapes to be allocated and written to a specific class
- Most frequent use case is for replicated vs. local only tapes

Class 0 = Replicated Filesystems
Class 1 = Local Only Filesystems

TAPE VOLUME FORMATTING

Volume on Tape

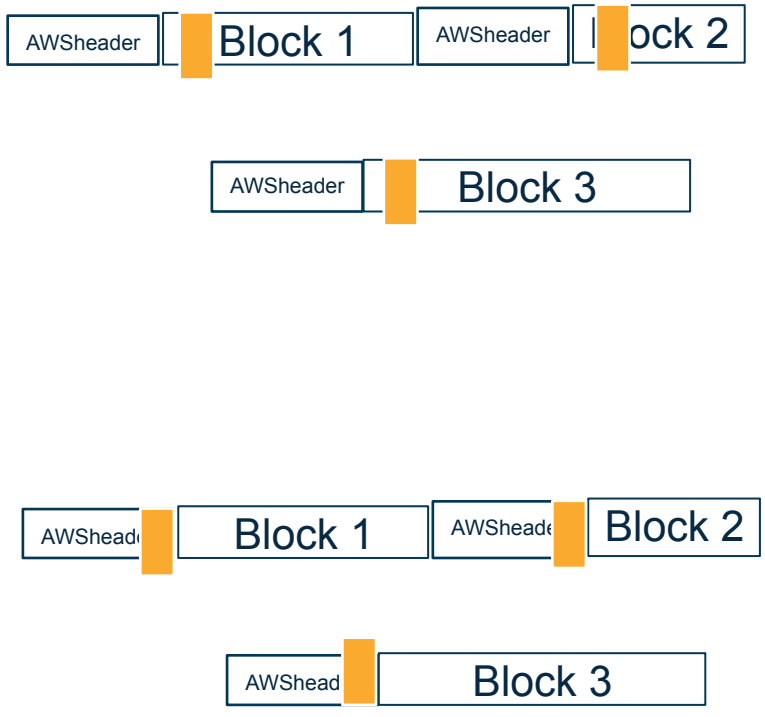
Volume on Disk



- Mainframe tapes write variable length blocks
- On physical tape blocks are separated by Inter-Record Gap (IRG)
- Variable length reads read until hitting IRG
- Tape on disk must record block lengths on write to insure block integrity on read
- An EMC-enhanced AWStape format is used on disk.
- Except for DD, VTEs compress the data before writing it to disk

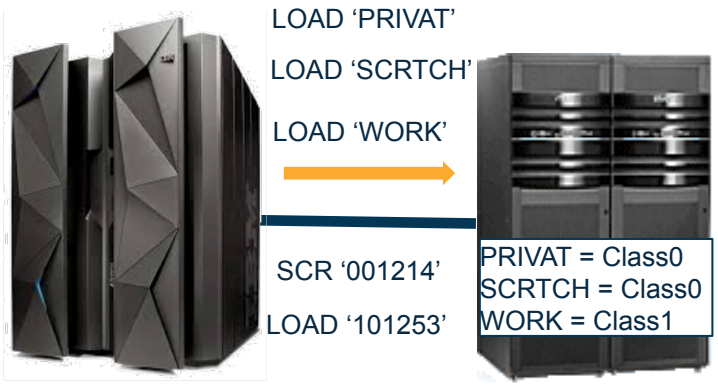
ASSISTED MAINFRAME DATA DEDUPLICATION

Before AMDD



- AMDD moves variable data to the AWSheader to improve data deduplication results
- AMDD looks at first few blocks for supported applications
  - DFSMSdss
  - FDR
  - Upstream
- On read data is restored to original format

ALLOCATING / MOUNTING / SCRATCHING

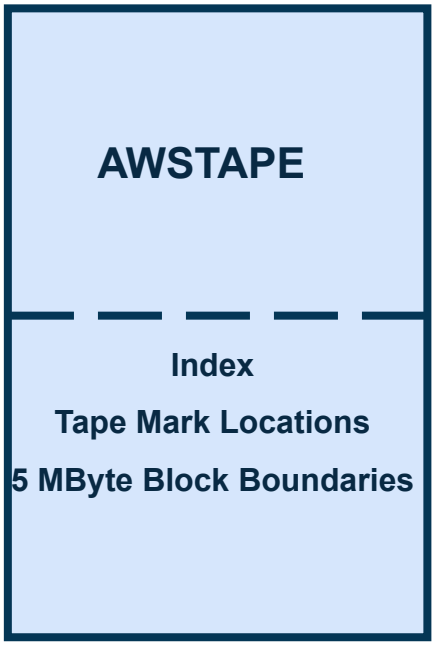


- All allocations/mounts/scratches are initiated by a LOAD CCW from the host
  - z/OS allocates a drive and then sends LOAD
  - EMC provides screen scraper for VSE
  - IBM DMSG is used in VM
  - Unisys Tape Manager issues LOAD CCW
  - EMC code for TPF
- Scratch synonyms defined in DLm identify scratch allocations vs. named mounts
- VTE will identify a filesystem and then look for a scratch using an LRU algorithm
- Scratch requests are a custom LOAD CCW

SCRATCHING ACTIVE TAPES

- Scratching is ALWAYS controlled by the Mainframe
- DLm provides a scratch utility for z/OS, VSE, and VM
- Scratch utility processes scratch reports from all leading commercial TMS software
- Unisys Tape Manager sends scratch requests
- Scratch requests cause DLm to rename file (123456 -> ~123456); data left in tact.
- Scratch space recovered under 3 conditions:
  - Re-use of scratched VOLSER
  - Scratch Erase command in scratch utility
  - Filesystem Space Threshold

FAST LOCATE CAPABILITIES



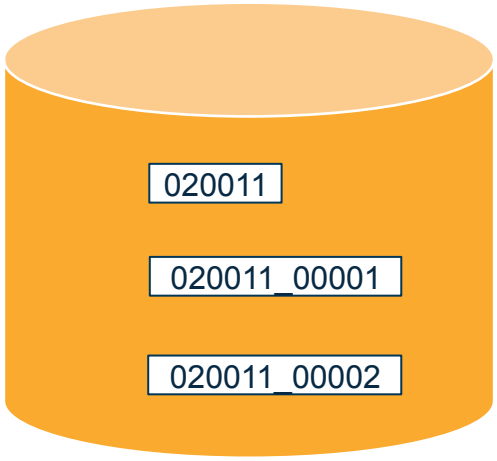
File Image on Disk

- As a VTE Writes a tape it builds an index for future fast locates.
- The index is stored at the end of the volume's disk file.
- The index contains a pointer to:
  - Each tape mark
  - Each 5 MByte block boundary
- On read the VTE can go directly to any tape mark
- For block locates the VTE will go to the nearest boundary and read forward or backward to find the block
- Block locates are accomplished with a maximum read of 2.5 MBytes (Avg. 1.25) of data

WORM TAPE SUPPORT

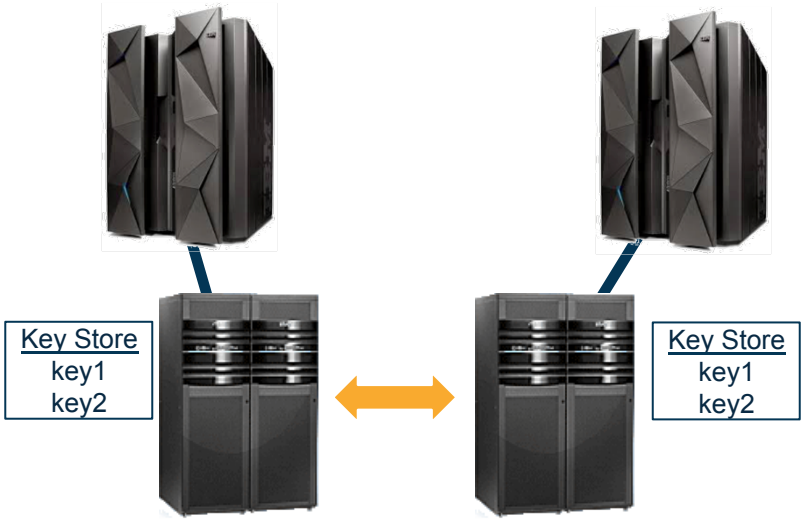
- DLm uses VNX File Lock Retention (FLR) or Data Domain Retention Lock (DDRL) to provide Write-Once-Read-Many (WORM) tape.
- DLm uses the expiration date in the HDR1 label to set expiration.
- On tape volume close, DLm sets the NFS “LAT” (last accessed date & time) fields to the future expiration date and then makes the file Read-Only.
- Once locked, the data on the VOLSER cannot be updated and/or deleted until the expiration date has expired.
- WORM volumes can be extended.

EXTENDING A WORM TAPE VOLUME



- Normally each VOLSER is a single file in a filesystem.
- FLR and DDRL prevent the file from being modified
- To allow WORM tape extension, DLm creates additional files for the VOLSER
- On Read all files will be presented as one logical volume.

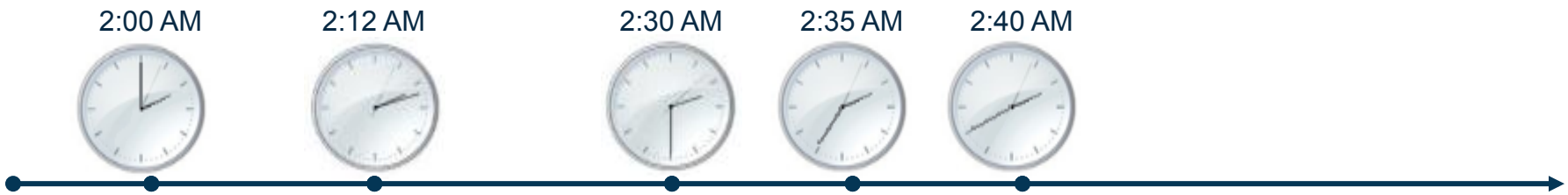
DATA ENCRYPTION



Each site has a unique write key.
Both sites can read/update all tapes.

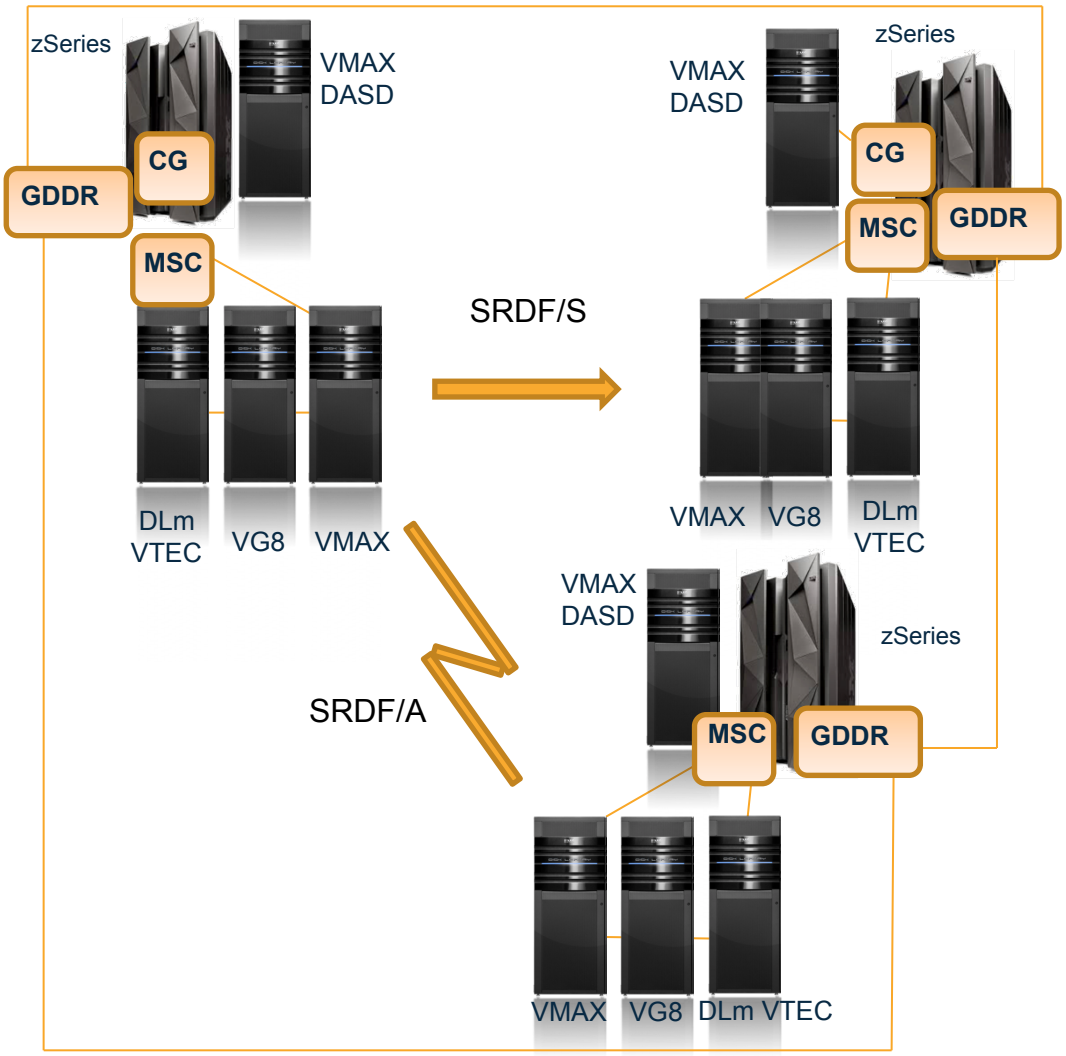
- DLm supports two types of Data Encryption
  - D@RE performed by the Storage
  - Static Key Encryption in the VTEs
- 1 to 7 keys are defined in a VTE key store
- A single key is assigned to each drive performing encryption
- All volumes written to drive are encrypted with same “static” key.
- On read/update any key in the key store will be used to decrypt the VOLSER

Guaranteed Replication



- GR guarantees a tape volume (VOLSER) has been replicated to a DR site before unloading the tape at the primary tape
- GR is triggered by double tape marks written to the tape.
- This insures the mainframe application will be notified if the replication does not complete.
- GR is configured on individual tape drives.
- All VOLSERs written to a GR drive will be guaranteed.
- VOLSER write times will be longer when using GR.
- Caution should be used with applications that write double tape marks other than at end of tape.

VMAX Universal Data Consistency™



- Universal Data Consistency insures data consistency between VMAX DASD and Tape
- DLm VMAX storage connects to VG8 and to Mainframe (using FICON).
- ConGroups and MSC allow VMAX DASD and DLM Storage to be defined in the same consistency group
- GDDR communicates over IP with DLMDR software running in VTEs
- GDDR controls DLM VMAX failover via FICON interfaces

Want to win an APPLE TV?

Engage with #EMCer and #WinPrize during this session and one lucky winner will walk away from the room with the prize

How to win** – Tweet a picture or comment on Twitter using these two hashtags **#EMCer #WinPrize

Questions?


EMC²