EMC DISK LIBRARY FOR MAINFRAME

Jim Stout

Global Mainframe Architecture Team

EMC Corporation

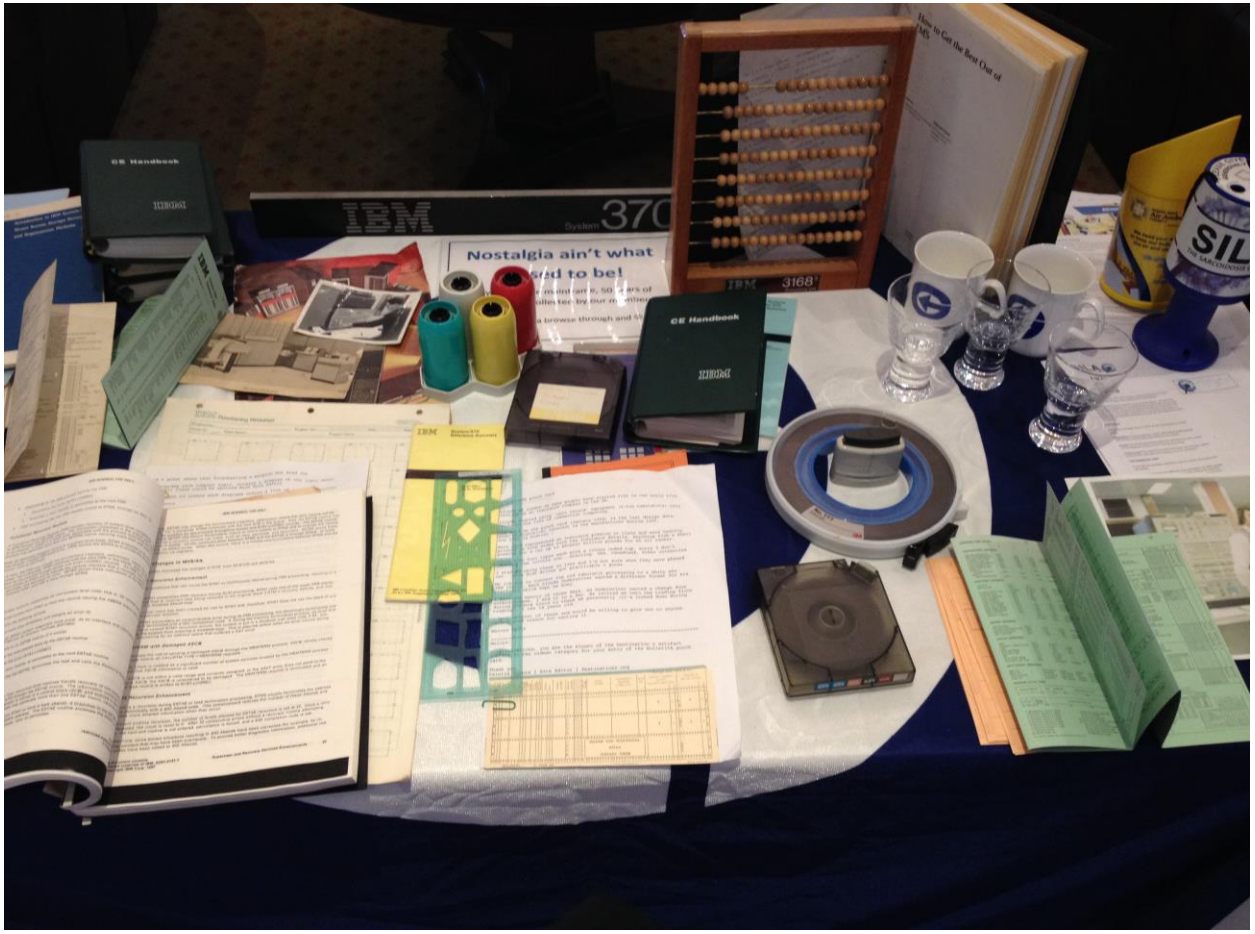
Session 17171

March 4, 2015

#SHAREorg

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

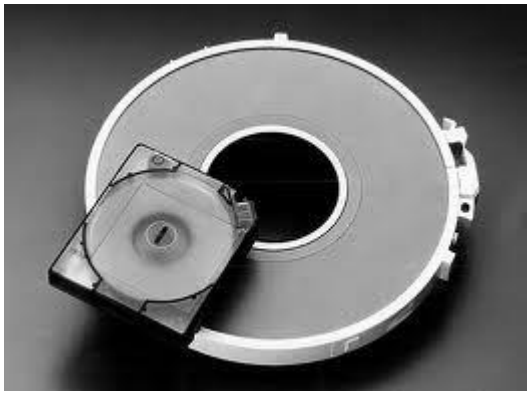
Guide SHARE Europe 2014



In The Beginning (Circa 1980) Economics Established The Policy

- DASD Street Price ~$10/MB
- Tape Prices:
  - $14/Tape
  - $0.07/MB
- Sequential Files To Tape
- Tape To DASD Ratio ~10:1

The Rapid Evolution Of Mainframe Tape



- In 1984, IBM introduced the 3480 tape drive
- Smaller form factor
- Greater durability
- Faster data access
- Same capacity as a 3420
- ~200 MB/cartridge
- IDRC compression introduced in 1986
- Capacity increased to ~600MB/cartridge

The Rapid Evolution Of Mainframe Tape

- In 1991, IBM introduced the 3490
- Same cartridge as the 3480
- Changed from 18trk to 36trk
- Double the capacity - ~1.2GB/cartridge
- Extended length media doubled capacity again
- ~2.4GB/cartridge

Typical Mainframe Tape Workloads

Backup

- 3390 DASD volume
- z/OS leading backup applications
  - FDR—Innovation Data Processing
  - DSS—IBM
- DB2 Image Copies

Space Mgmt

- IBM's HSM
- CA/Disk (DMS/OS)
- different
  classes
- Meant to conserve DASD usage
- Can use significant CPU cycles

Data Archive

- Fixed content data
  - Check images, etc.
  - Variable data
  - ASG
  - IBM ImagePlus

Work Tape

- Short retention—temporary datasets
- High read/write requirements
- Example: Syncsort work files
- Large sequential files

All With Different Reference Patterns

Typical Mainframe Tape Solution

IBM z/OS mainframe

Backup

Space Mgmt

Data Archive

Work Tape

Physical tape

Virtual tape

Deduplication

You would need two or three different tape platforms (and code bases) to satisfy the needs of all workloads

EMC Mainframe Tape Solution

IBM z/OS mainframe

Backup

Space Mgmt

sical tape

EMC DLm8100

Data Archive

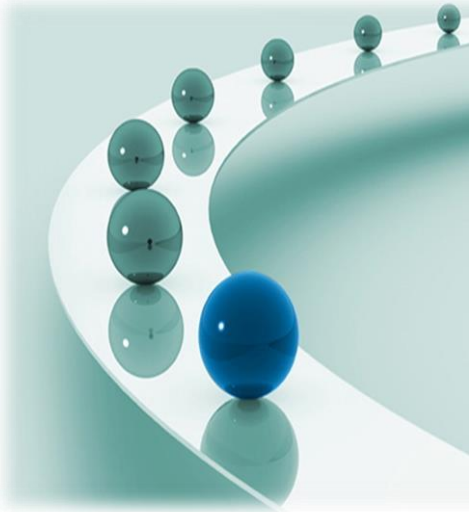
Work Tape

Deduplica

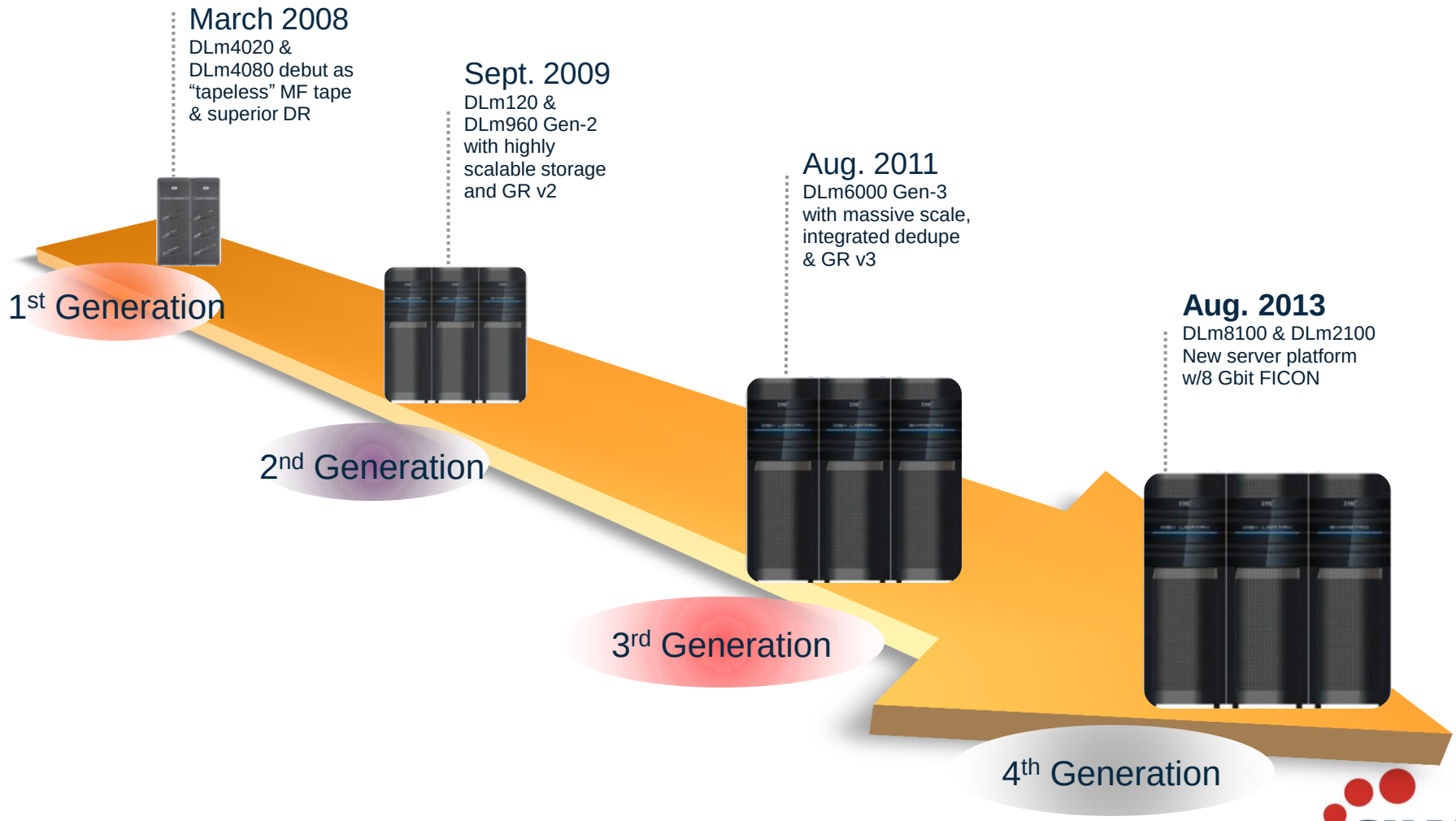
EMC satisfies the needs of all tape workloads in a single, manageable solution

Why DLm For Mainframe Tape?

- Eliminate physical tape media
  - DLm scalability enables customers to go completely “tapeless”
  - Risk mitigation to avoid litigation and public embarrassment due to lost or stolen media
- Consistent performance
  - Single tier of storage eliminates the need to de-stage and re-stage tape volume to and from tape
  - Instant tape mounts coupled with high-speed locate for disk-like response time regardless of the age of the tape dataset
- Breakthrough disaster recovery
  - Read / Write Point-In-Time copies enable fully destructive disaster recovery testing without compromising production replication
- Superior availability
  - Microcode updates and corrective maintenance activities are performed concurrently without the need to interrupt tape operations
- Integrated data de-duplication capability
  - Inline data de-duplication improves storage efficiencies and reduces data replication bandwidth requirements



Evolution of DLm For Mainframe Tape



What is DLm Gen 4 ?

- 4th Generation Virtual Tape Engine
  - Higher Performing Intel Platform
  - Integrated ACP functionality
  - 8Gb FICON connectivity vs. 4Gb
  - Enhanced Compression Adapter
  - (2) 10Gig & (4) 1 Gig E ports
  - Full 256 Device Count support
  - 4096 Logical Paths
  - Single VTE across DLm offering
    - 1 & 2 channel version
- 60 Port 10Gig Ethernet Switch

Platform Overview

- 4th Generation successor to current DLm
  - DLm1000, DLm2000, DLm6000, DLm8000
- Model Structure
  - DLm8100 (Enterprise)
    - Consolidates DLm6000 & DLm8000
  - DLm2100 (Mid-market)
    - Consolidates DLm1000 & DLm2000
- Release Approach
  - VTEs comprise base DLm8100/2100
  - Supports current DD, VNX & VMAX storage
- Value Add
  - Dedup & Traditional storage in a single system
  - Modular/HA architecture
  - Scalability/Performance
    - FICON Connectivity, Device Count, Logical Path Count
  - Synchronous replication

DLm Family

	DLm8100/x	DLm8100/d/v	DLm2100/v	DLm2100/d
Number of VTEs	2 - 8	2 - 8	1 or 2	1 or 2
Connectivity	FICON	FICON	FICON	FICON
Number of channels to host	4-16	4-16	2 or 4	1 or 2
Number of virtual tape drives	Up to 2,048	Up to 2,048	Up to 512	Up to 256
Maximum capacity (usable non-dedupe)	56 TB -1.792 PB	40 TB to 2.3 PB	13 TB – 221 TB	Depends on DD
Performance	Up to 6.4 GB/s	Up to 6.4 GB/s	Up to 1 GB/s	800/1600 MB/s
Deduplication storage	N/A	Yes – DD990 (2)	N/A	Yes – current DD
Deduplication capacity	N/A	Up to 1.1 PB	N/A	Up to 570 TB
Number of Cabinets	3-13	2-16	1-2	N/A
Replication	Yes (SRDF)	Yes	Yes	Yes

DLm2100

- DLm2100 w/Data Domain Storage
  - 1 or 2 VTEs
  - 1 or 2 channels per VTE
  - 256 Devices per VTE
  - Support for DD models 2200 through 990

- DLm2100 w/VNX Storage
  - Single Rack Solution
  - Optional 2nd cabinet for storage expansion
  - 1 or 2 VTEs
  - 2 channels per VTE
  - 256 Devices per VTE
  - Support for VNX 5400
    (13TB – 221TB)

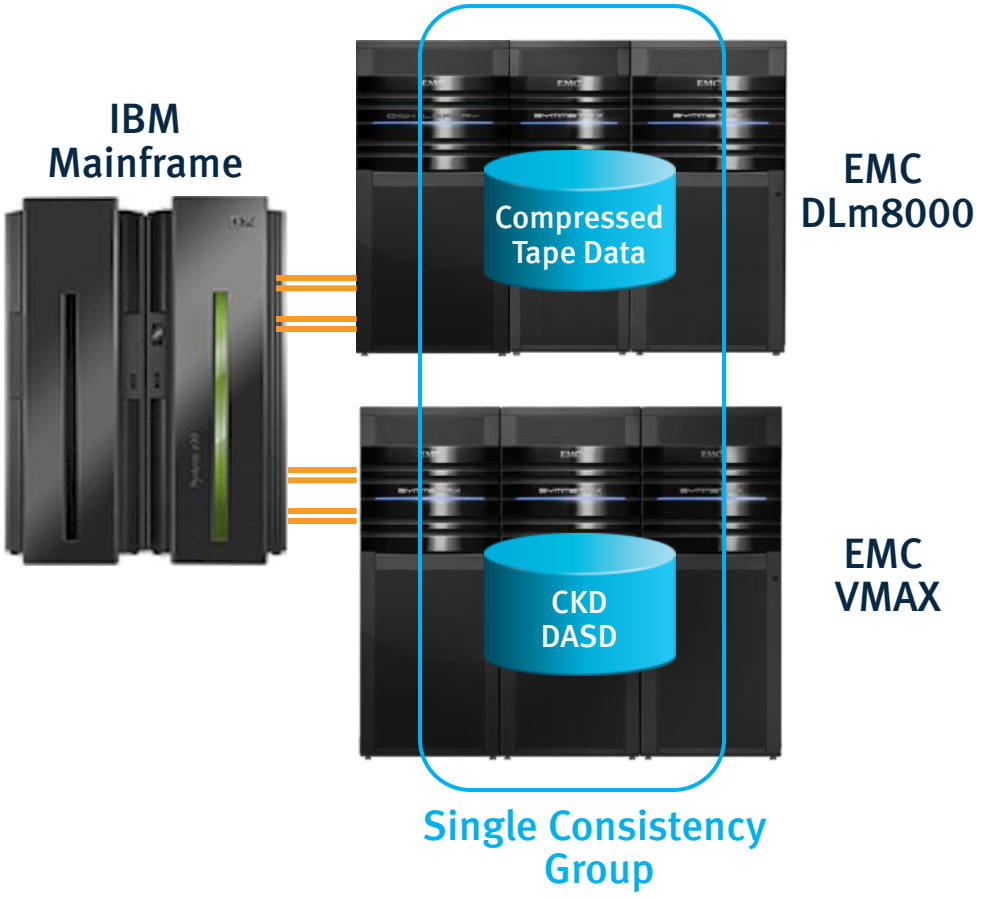
DLm8100

- DLm8100 VTEC
  - 2 - 8 VTEs
  - 4 - 16 channels
  - 512 - 2048 Devices
- Supported Storage Combinations
  - 1 or 2 VNX 7600s (up to 2560 TB)
  - 1 or 2 DD990s (up to 1140 TB)
  - 1 VNX 7600 & 1 DD990
  - 1 VMAX 20K or 40K
- Replacement for DLm6000

EMC DLm8000
Unprecedented Scale, Resiliency

- High-availability (HA) DLm architecture
- 56TB – 1,792TB per VMAX
- 1600MB – 6.4GB/second throughput
  - Up to 8 VTE's
- 512 – 2,048 virtual devices
- 4 –16 FICON attachments
- SRDF/S and SRDF/A replication for tape
- GDDR for automated recovery
- Universal Data Consistency™ DASD & tape
- Transparent to mainframe operations
- 3-13 cabinets

Universal Data Consistency™ for Tape and DASD (DLm8000)



- VMAX storage arrays for both Tape and CKD DASD
- Remote Replication using SRDF/A or SRDF/S
- Local Replication using TimeFinder
- Full Failover Automation using GDDR
  - Planned Failover
  - Unplanned Failover
  - Planned D/R Test
- Single replication methodology
- Tape and DASD Consistent to one another!

Gen 4 Offerings

Primary Storage	Dedupe Storage	Combined Primary and Dedupe Storage

DLm 8100 Enterprise

DLm8100
w/ VNX7600

DLm8100
w/ VMAX

DLm8100
w/DD990

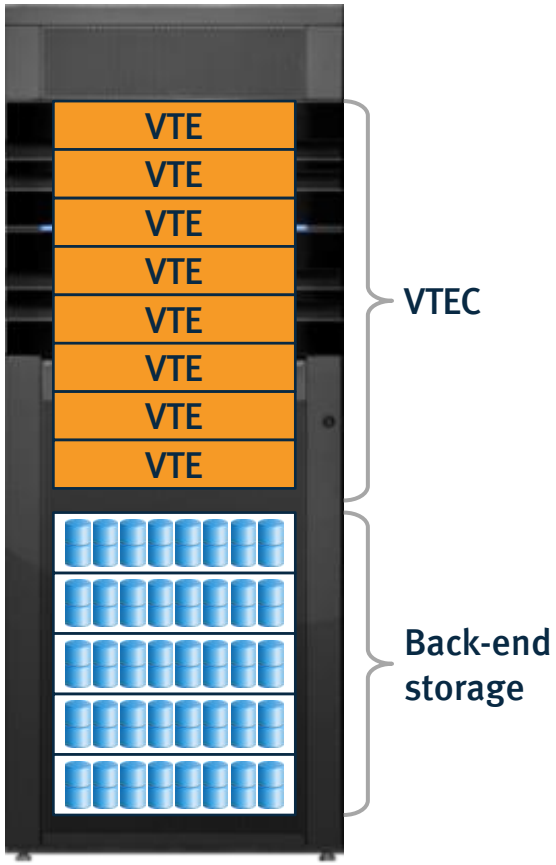
DLm8100
w/ VNX7600 &
DD990

DLm2100 Small to Medium

DLm2100
w/VNX5400

DLm2100
w/DD 600/800/900

DLm Components



EMC Disk Library for mainframe

Virtual tape emulation controller (VTEC)

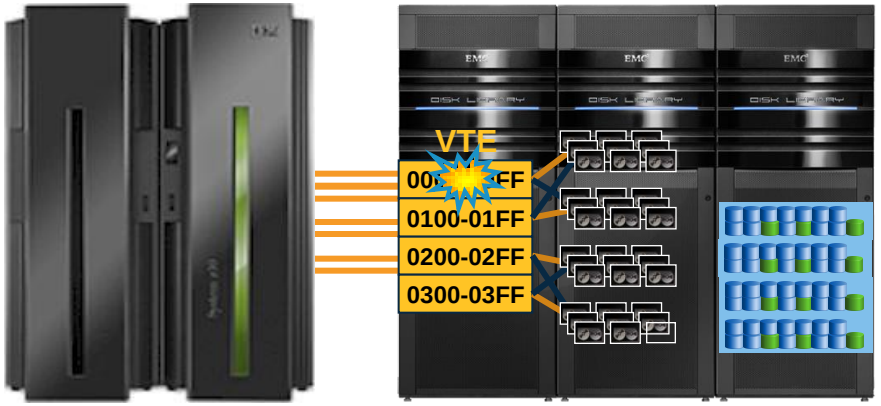
- Emulates IBM 3480/3490/3590 tape drives
  - 256 tape drives per VTE
  - 4096 Logical Paths
- 8 Gb FICON connectivity
- Virtual cartridge size up to 16 TB
- Disk consumption is based on data written
- Supports deduplication and/or hardware compression

Back-end storage

- Leverages traditional NFS storage – shared storage
- Supports de-duplication storage
  - DLm2100 w/DD and DLm8100 w/DD
  - Or both storage types concurrently (DLm8100 only)
- Stores all tape images as files
- Shares all tape volumes among all VTEs

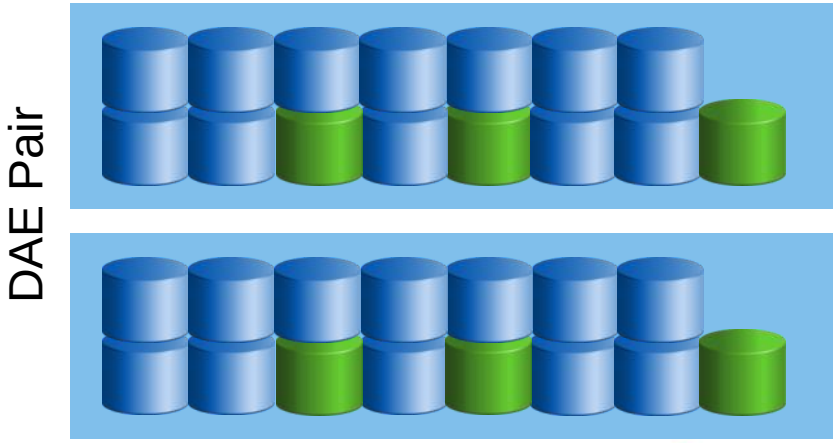
High Availability Architecture

- All VTEs can see all tape volumes
- If a VTE fails…
  - Job will abend with tape error (same as regular tapes)
  - Tape volumes are still available through alternate VTE
  - VTE does not contain metadata
- Disk drives are RAID 6 protected
  - 8 + 2 protection – the best protection
  - Global Hot spare drives are located in the Vault Disk Array Enclosure (DAE)
- Storage controllers configured as N+1
  - One stand-by to take over in case a storage controller fails



All tape volumes are available

EMC Disk Library for mainframe



Disk Library for Mainframe Performance vs. Two-Tiered Virtual Tape

Physical tape

Mount ? Sec → Seek ? Sec → Stage

Time to first byte: ~120 seconds

Job completion

Disk Library for mainframe

Near Instant Mount

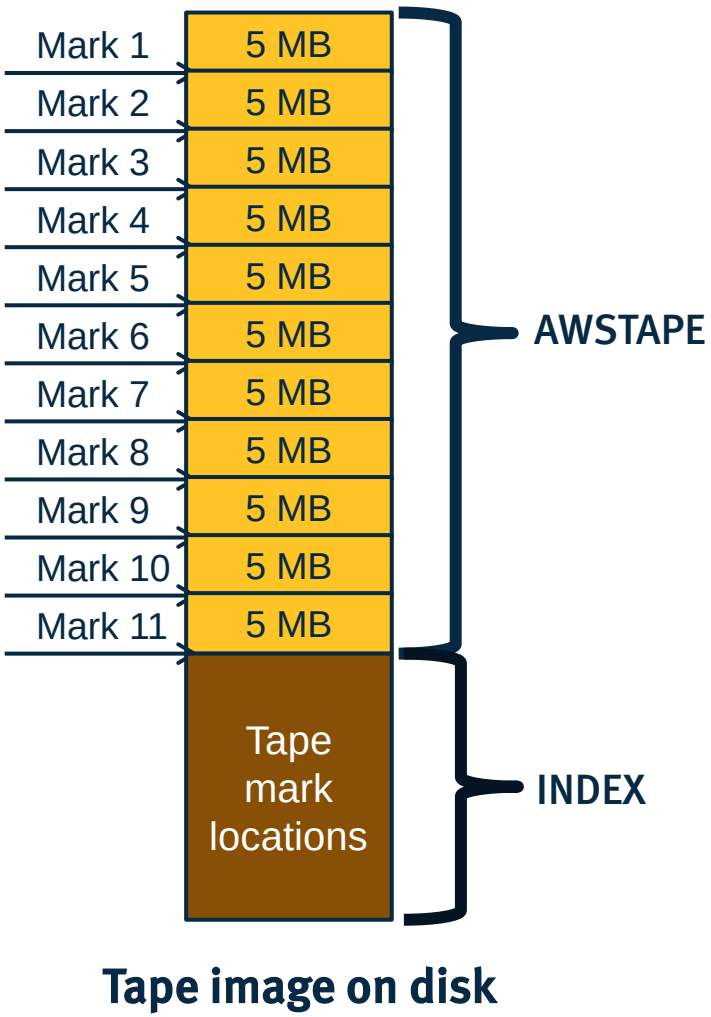
Job completion

Disk Library for mainframe

Provides…

- Faster time to first byte
  - Sub-second mount time on average
- Faster batch processing
  - Over 800 MB/s per VTE
  - Up to eight VTEs

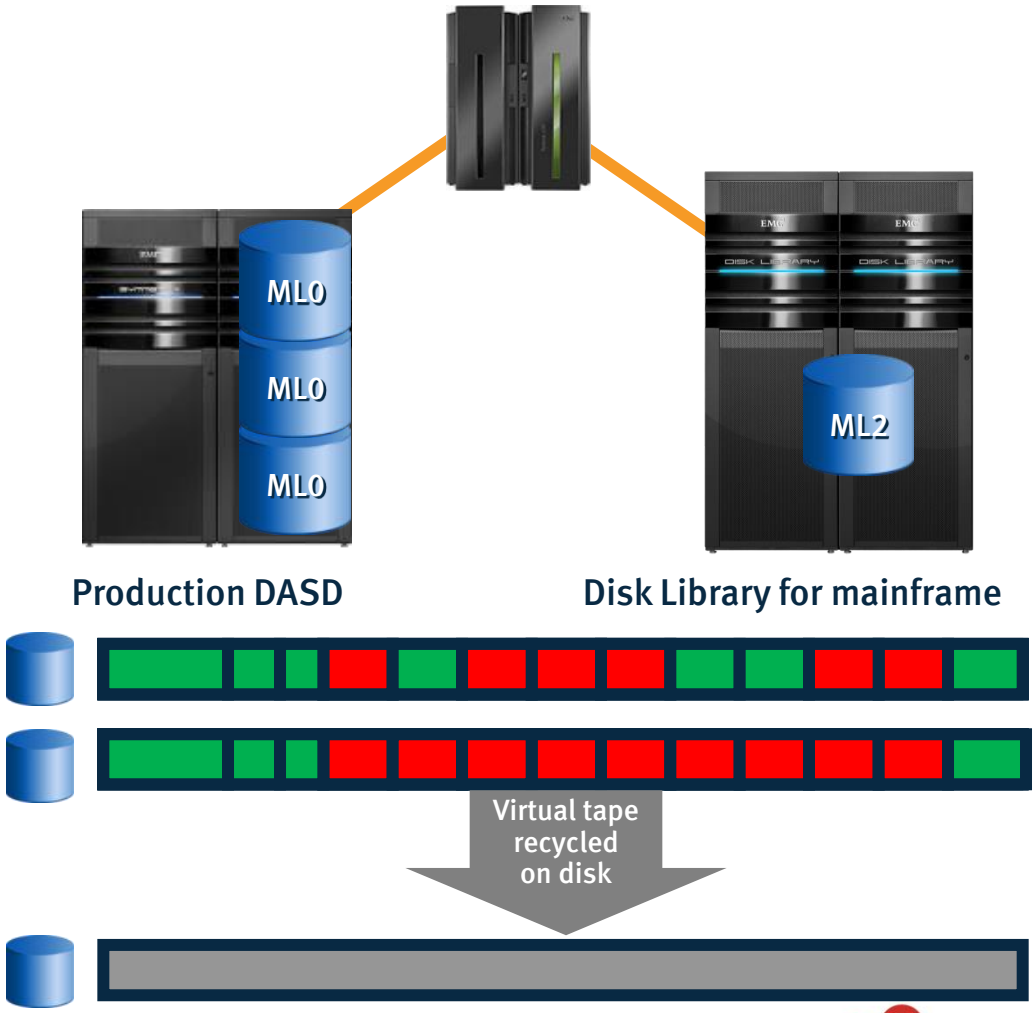
Improved Response Time



- When Disk Library for mainframe writes to tape, it builds an index for fast locate
  - Index is stored at the end of the volume's disk file
  - Contains a pointer to:
    - Each tape mark location
    - Each 5 MB block boundary
  - On read, the VTE can go directly to any tape mark
- Fast locate mechanism provides additional performance benefits
  - Response times drop from 15–40 seconds on physical tapes, to less than one second on average
  - Improves service level agreements
  - Immediate major benefit to fixed content applications such as:
    - Archive / Report Distribution: SAR, Mobius, OAM, others
    - Space Management: HSM, CA/Disk, FDR/ABR

Reduce HSM CPU Cycles

- Reduce/eliminate DFHSM ML1
  - Move directly from ML0 to ML2
  - Save the mainframe CPU compression cycles
  - Recoup the ML1 DASD Pool
  - ML2 data is compressed and kept on disk in the DLm
  - Recalls at disk speed with indexed high-speed locate
- DFHSM recycle time optimized
  - DFHSM will continue to perform tape recycling
    - Recycling at disk speed
    - User -settable virtual cartridge size
    - No contention for tape drives
  - Hours of savings potential



How DLm Is Genned To The z/OS Host

IBM z/OS mainframe

EMC DLm

VNX storage
Batch
HSM
Fixed content

Data Domain storage
Backup
Archive

- The mainframe host views Disk Library for mainframe as tape drives
- Data for different tape workloads allocation is directed to the appropriate storage type
- Drive allocation:
  - Simple Esoteric
  - Sun/STK HSC/SMC
  - SMS MTL
- Scratch Sub-pooling
- Each VTE can access any tape dataset in the logical tape library
- Each tape VOLSER is kept on disk as a file
  - No meta data

Data Domain Systems

Data Domain Software Options

- DD Boost
- DD Encryption
- DD Extended Retention
- DD Replicator
- DD Retention Lock
- DD Virtual Tape Library
- DD Management Center

Small Enterprise/ROBO

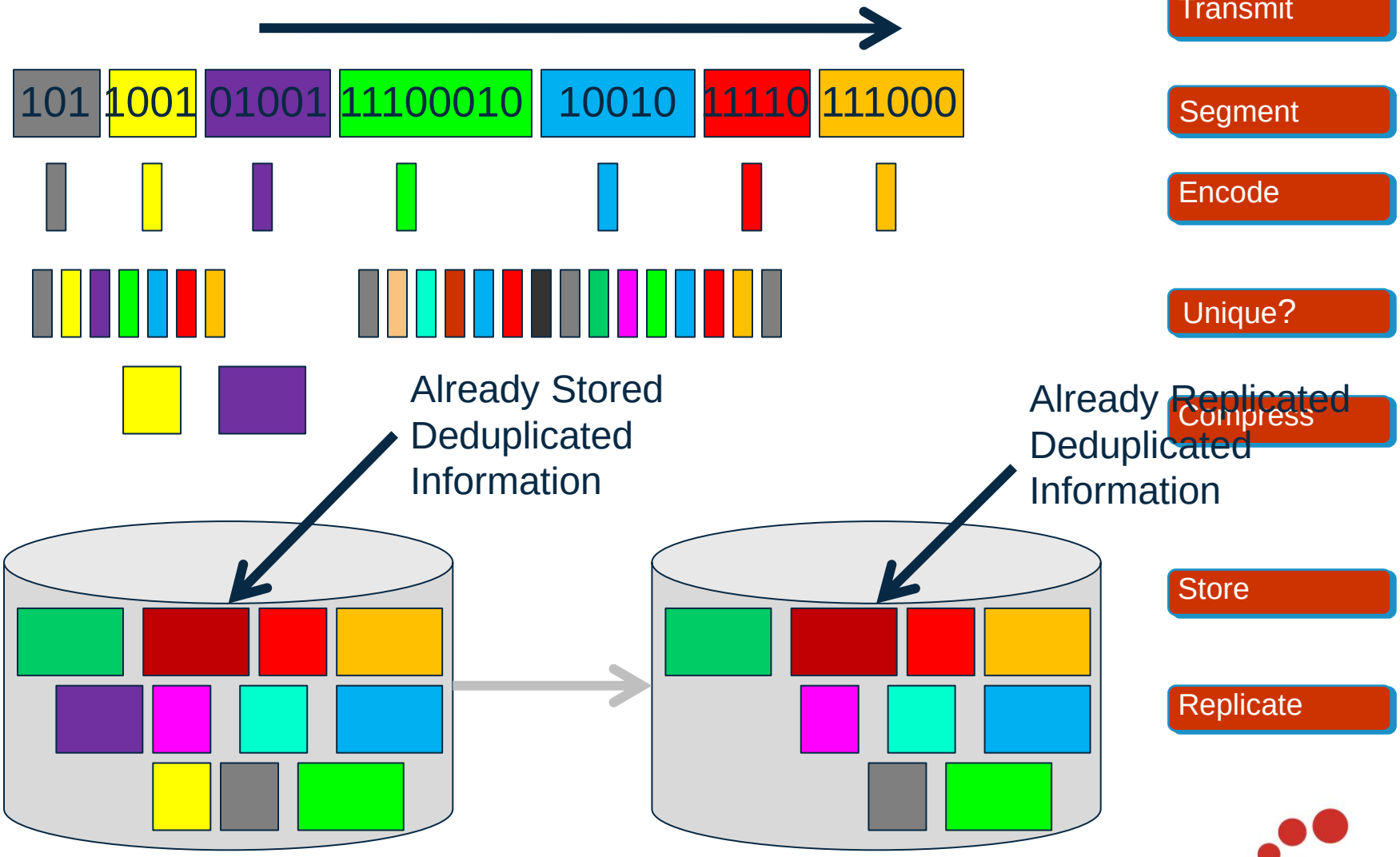
Midsize Enterprise

Large Enterprise

	DD160	DD2200	DD2500	DD4200	DD4500	DD7200	DD990
Speed (DD Boost)	1.1 TB/hr	4.7 TB/hr	13.4 TB/hr	22.0 TB/hr	22.0 TB/hr	26.0 TB/hr	31.0 TB/hr
Speed (other)	667 GB/hr	3.5 TB/hr	5.3 TB/hr	10.2 TB/hr	10.2 TB/hr	11.9 TB/hr	15.0 TB/hr
Logical capacity	40–195 TB	172-860 TB	1.3–6.6 PB	1.8-9.4 PB 5.6-28.4 PB[1]	2.8-14.2 PB 11.4-57.0 PB[1]	4.2-21.4 PB 17.1-85.6[1]	5.7–28.5 PB Up to 100 PB[1]
Usable capacity	Up to 3.98 TB	Up to 17.2 TB	Up to 133 TB	Up to 189 TB Up to 569 TB[1]	Up to 285 TB Up to 1.1 PB[1]	Up to 428 TB Up to 1.7 PB[1]	Up to 570 TB Up to 2.0 PB[1]

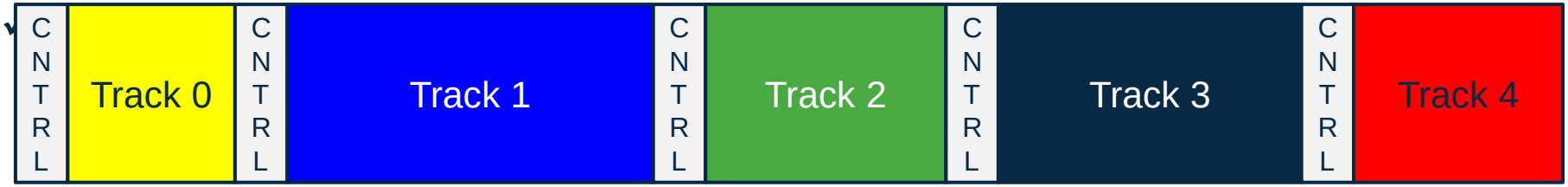
[1] With DD Extended Retention software option

Deduplication: The Process

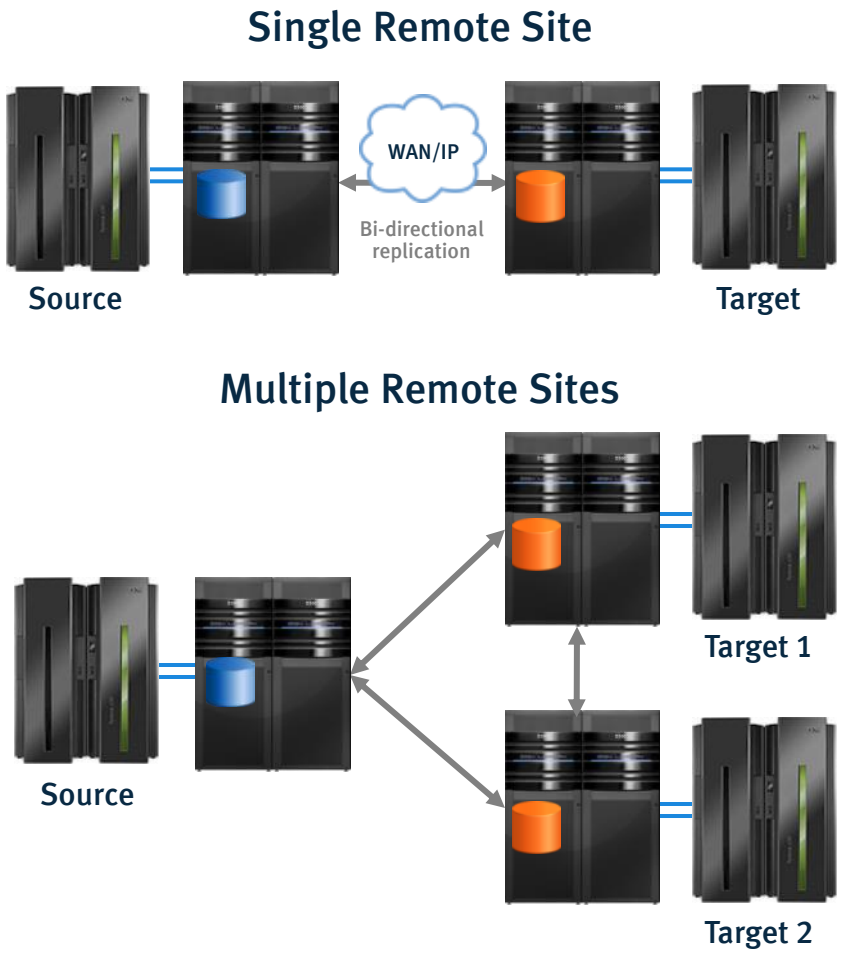


AMDD - Assisted Mainframe Data De-Dupe

- ✓ Virtuent detects the "signature" of backup applications (e.g. FDR, DFDSS)
- ✓ When a VOLSER arrives in the VTE, the first few blocks of data are analyzed
- ✓ If the data is NOT from one of the AMDD supported applications it will be processed as it is, without using AMDD. The data is written to the back-end storage.
- ✓ If the data is from one of the AMDD supported dataset, then the data will be pre-processed by AMDD. (AMDD Isolates Variable Control Data).
- ✓ Once re-organized, the data is forwarded to the back-end storage, as normal.

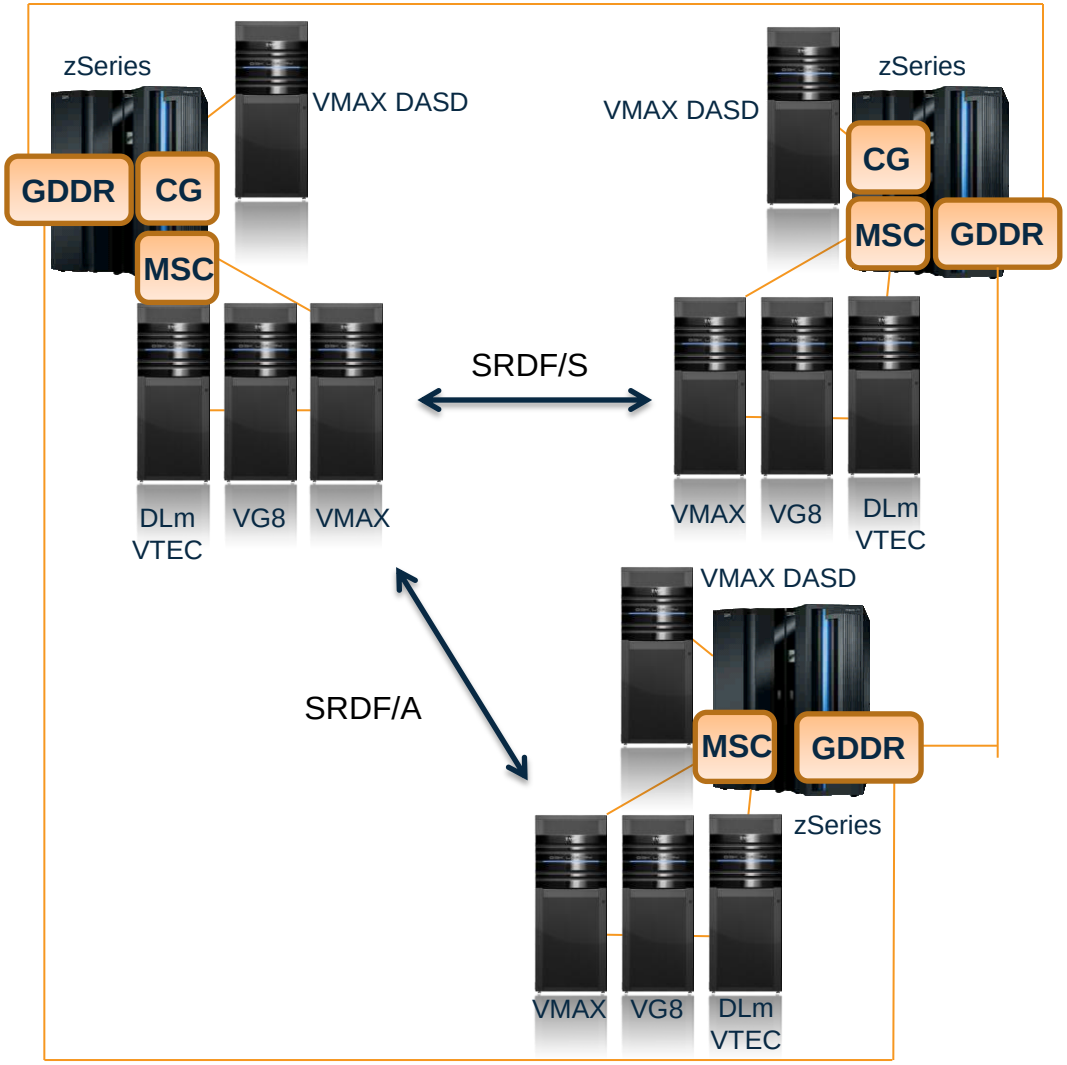


Remote Replication



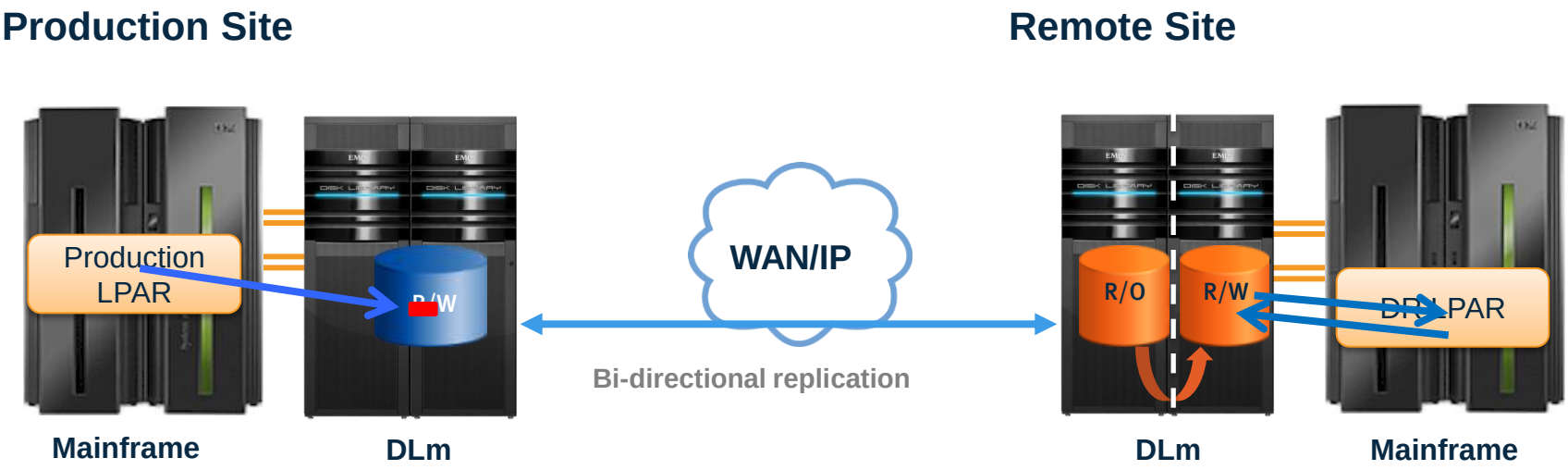
- Deduplication yields less data
  - Less bandwidth needed for replication
- Reduce current tape RPO/RTO
  - From days down to hours or minutes
- No performance penalty on mainframe host
- Back-end storage IP replication
- Flexible replication configurations
  - Prioritize data for replication
  - Replicate to one or two sites

DLm Vision: SRDF, GDDR, Universal Data Consistency™ (DLm8000)



- DLm8000 provides 3 Site Star using VMAX 20K or 40K; providing Ultra Resilient DASD / tape environment
- Simple 2 Site SRDF/S environment supported w/o GDDR or Con Groups
- Follow-On release will expand support to include 4 Site SQAR configurations

Testing the Disaster Recovery Environment



Read-only mounts

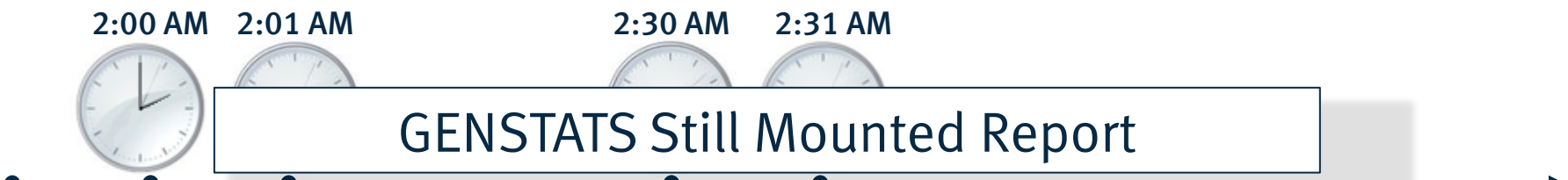
- Disk arrays allow instant “read-only” copies
- Confirm that tapes can be mounted and all required data can be accessed
- No incremental storage capacity required
- Issues / Testing Capabilities Limited

Remote replication is uninterrupted during testing

Snapshots

- Disk arrays allow creation of “read-write” snapshot
- Confirm operation at the disaster recovery site
- Some incremental storage capacity required

Tape Status Management

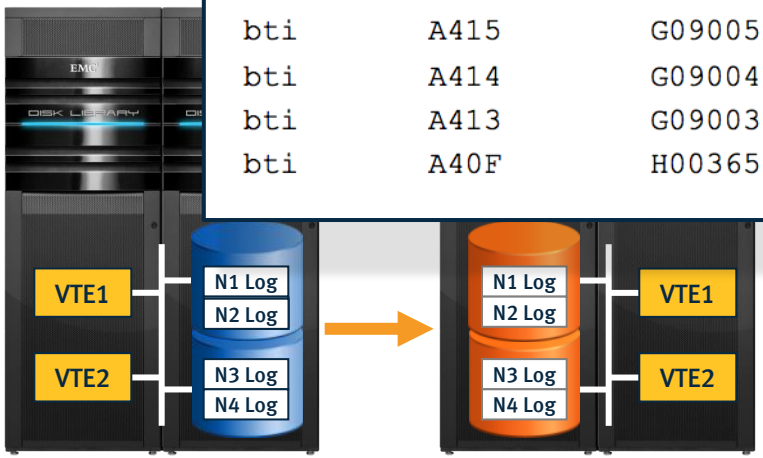



```
VOLSERS STILL MOUNTED    :

NODENAME   DEVICE      VOLSER   LAST MOUNTED          PATH
bti        A405        GF0825   2008/11/10 10:23:18   tapelib/GF
bti        A404        GF0826   2008/11/10 10:23:18   tapelib/GF
bti        A419        G09009   2008/11/07 19:26:43   tapelib/G0
bti        A415        G09005   2008/11/07 19:26:43   tapelib/G0
bti        A414        G09004   2008/11/07 19:26:43   tapelib/G0
bti        A413        G09003   2008/11/07 19:26:43   tapelib/G0
bti        A40F        H00365   2008/11/10 10:23:18   tapelib2/H0
```

Local S

/EE
EE00188 m
EE00188 un



nd maintains
s
cific VOLSER

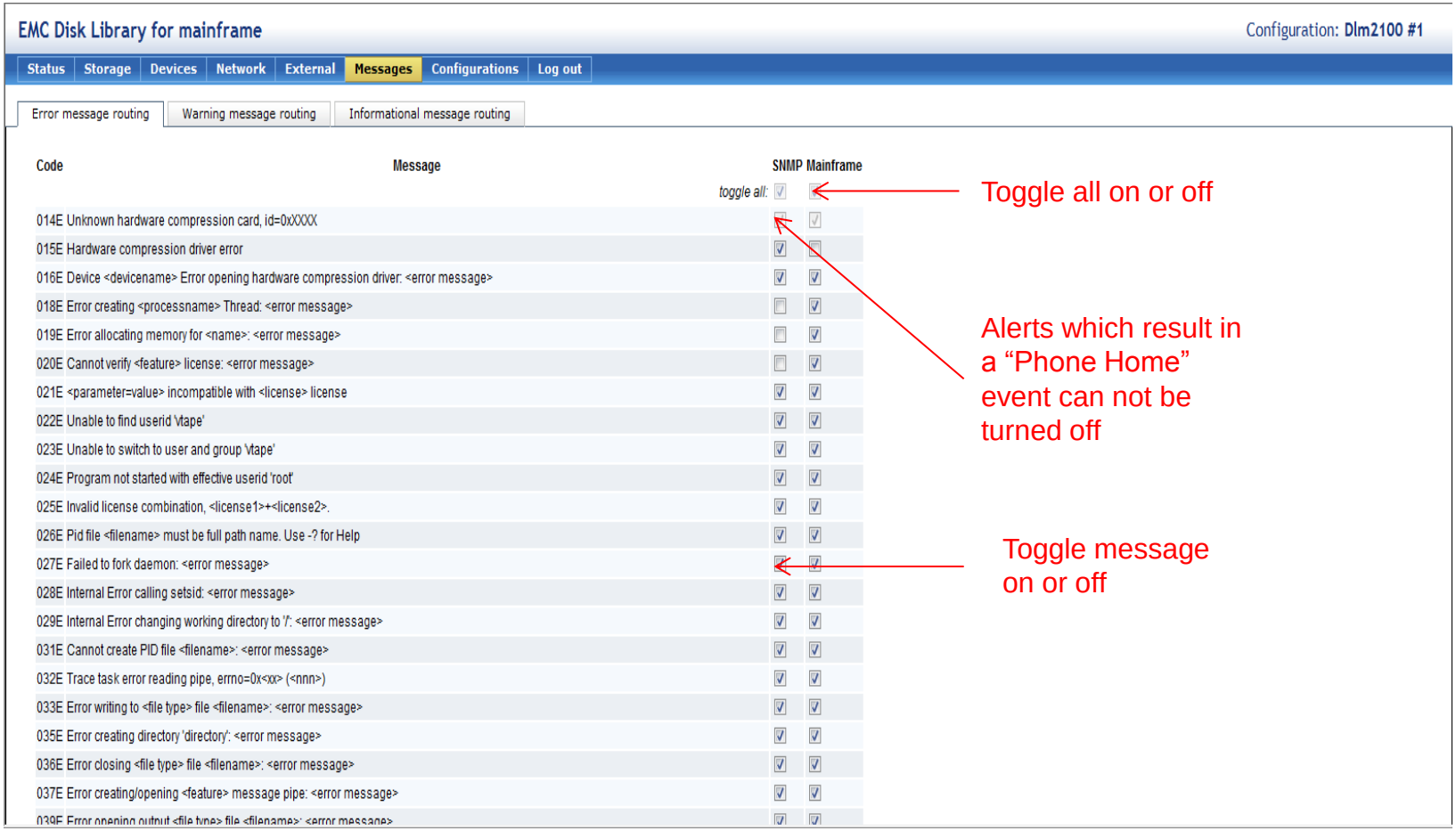
nted, unmounted),
nt log file,
c VOLSER
e from any VTE

Lm storage

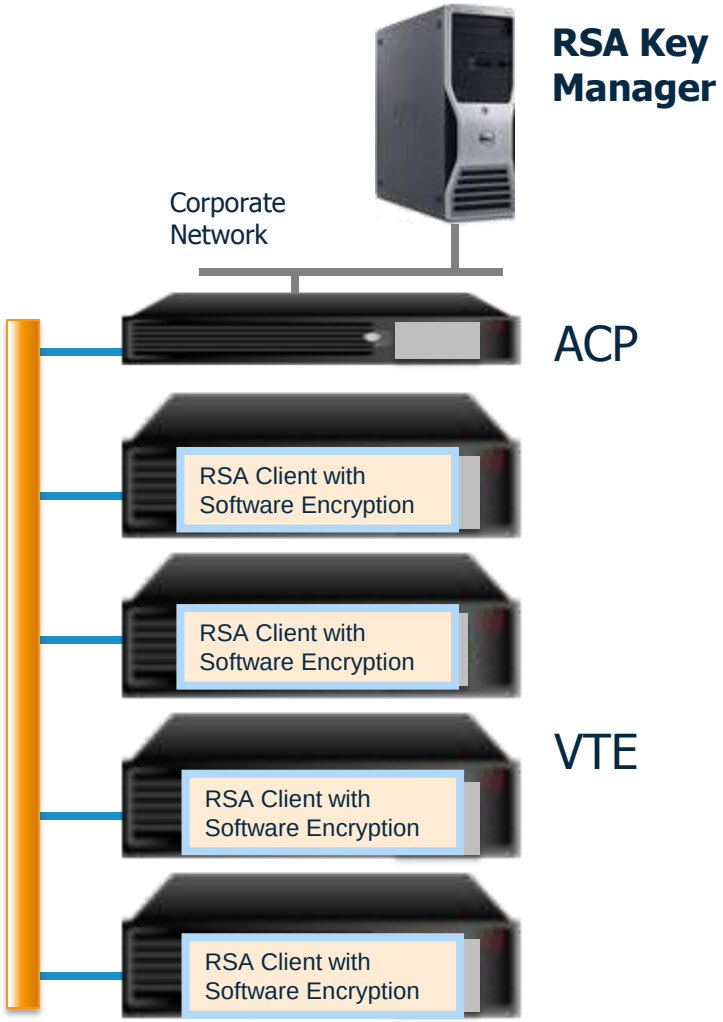
ed to the
remote site along with the tape volumes

- Provides status of all tape volumes inside the DLm at any given time, both local and remote DLm systems
  - Can be used as part of disaster recovery operations

DLm Phone Home

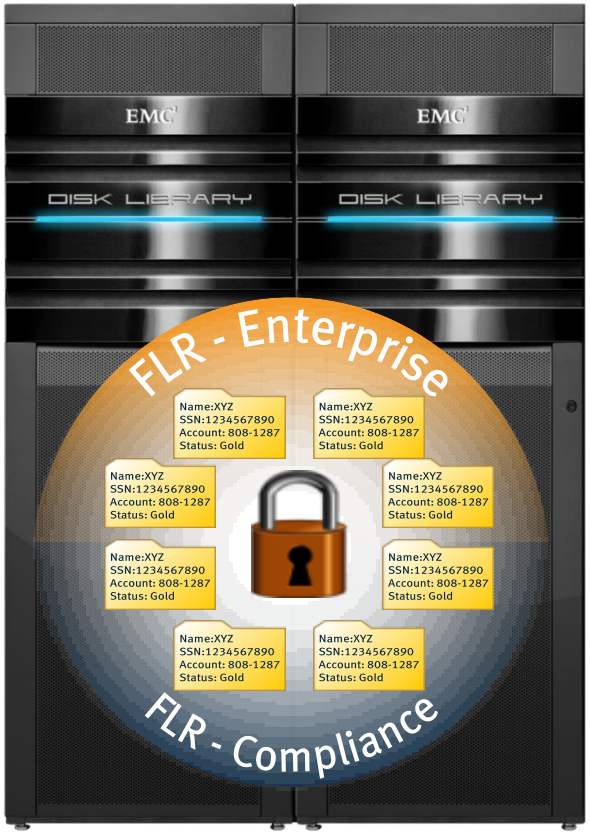


Virtuent 7 RSA Key Management



- VTEs will always include the RSA client
- Encryption requires a separately licensed RSA Key Manager
- Individual drives will be configured for encryption (as they are today)
- When a new tape is written to a drive configured for encryption a call to the RSA Key Manager is made
- The Key Manager returns 2 keys; a clear key and an encryption key
- RSA client uses the encryption key to encrypt the data and stores the clear key on the volume
- Support for VNX Storage Only

Virtuent 7
VNX File Level Retention (WORM)



File Level Retention Enterprise Option (FLR-E)

- Provides for retention periods per file
- Enables adherence to good business practices
  - **Tamper proof clock**
  - **Activity Logging**

File Level Retention Compliance Option (FLR-C)

- Meets SEC Rule 17a-4(f) compliance requirements
  - Prevents file system deletions with locked files
  - "Hard" default retention periods
  - **Data verification** to validate committed content
- Retention periods cannot be modified
- File systems can only be deleted when retention period has expired
- Third-party compliance validation paper

Virtuent 7
Guaranteed Replication

- Replication of identified VOLSERs to a second location completes before ending status is returned to host
- Intelligent stacking of queue to replicate all VOLSERs in identified filesystem at once
- Significant performance improvement over GR v2
- Minimizes transmission traffic
- Similar to IBM's RUN Mode

EMC and IBM for Mainframe Support

- IBM partnership
  - License key technologies with automatic renewal
  - Member of IBM System z Early Ship Program (ESP)
- First and only vendor with GDPS lab
- Customer Support Agreement
  - Outlines details of EMC/IBM cooperation
  - Provides escalation procedures
  - Covers strategic product families from both companies
- Base Interoperability Agreement
  - “EMC and IBM have a program of interoperability testing that allows the companies to evaluate and document the compatibility of EMC Products and IBM System z servers and to evaluate and resolve joint customer support issues.”
- “Co-opetition” model
  - Cooperate on applications and operating platforms
  - Compete on storage and storage software

Why EMC Virtual Tape For Mainframe?

- Eliminates all issues related to traditional tape handling
  - Eliminates manual intervention, physical movement of tape cartridges, robotic issues, and single points of failure
- Works seamlessly with existing applications
  - No Job Control Language (JCL) changes are required and works with your existing tape management system
- Significantly improves performance
  - Reallocates all of the data to disk and uses smart I/O buffering, allowing potentially significant reductions in batch windows
- Extends disaster recovery capabilities to the tape workload
  - Utilizes array-based replication process over IP to seamlessly move tapes offsite
- Enhanced data protection with encryption
  - For data at rest or during transmission
- Easily scales as the workload increases
  - Add throughput performance and storage capacity as needed

Thank You!

Enjoy The Rest Of The Conference!