White Paper

Deploying a Mainframe Virtual Tape Solution: Planning and Design Considerations for EMC DLm4080

Technology Concepts and Business Considerations

Abstract

This white paper describes planning and design considerations for implementing EMC® DLm4080 in a mainframe production environment. It includes topology, tape analysis, and design considerations as well as risks and limitations associated with the configuration. It describes the considerations that must be taken into account as part of the planning and design process, prior to implementing the DLm4080 design.

April 2008

Part Number H4395

Table of Contents

Executive summary

The EMC® Disk Library for mainframe series is a new product for IBM z/OS mainframe environments. It provides a tape emulation appliance for data centers looking to replace outdated physical tape and mainframe virtual tape solutions.

A properly sized and configured EMC Disk Library for mainframe solution can provide both the capacity and performance required for successful tape operations in today's mainframe environment. By examining the current configuration, sizing the environment, and applying some best practices to the configuration, maximum performance and availability can be achieved. This white paper discusses the planning and design considerations for deploying an optimal DLm4080 configuration.

Introduction

EMC's DLm4080 offers a number of advantages over traditional tape and virtual tape systems. Whether the application is backup/restore operations to scratch volumes or private volumes that supplement DASD, HSM recalls, or any other application, the advantages are significant. The most obvious advantage is faster, more reliable transactions. Other advantages include the seamless integration into existing enterprise mainframe environments. DLm emulates disk as tape, and the mainframe host views the DLm just as another IBM physical tape library, which allows organizations to use all of their existing policies and procedures as little to no reconfiguration is required to bring the solution online.

When planning a DLm deployment the first things that need to be determined are the types of tape devices that are currently configured within the mainframe infrastructure. The host device configuration will dictate the type of virtual devices that will be configured at the DLm. Knowing the type of device is only part of the planning effort for deploying a DLm. Time should be taken to planning the configuration for optimal performance. Typical deployments configure the DLm with the same number of tape devices and tape cartridges as currently configured. If performance was a consideration for moving away from physical tape to virtual tape the configuration should be reconfigured for an environment that is built for availability with the best possible performance.

To properly determine the appropriate number of tape drives and cartridges it is important to re-evaluate the current configuration and its current utilization. It should be expected that the configuration will change as the DLm allows for hardware to be seamlessly added into the configuration with only the cost of storage capacity, providing there are network and processor resources available. The DLm automatically handles the provisioning and allocation of the virtual tape drives and cartridges.

Audience

This white paper is intended for mainframe tape and storage administrators who are interested in deploying an EMC Disk Library for mainframe solution in a z/OS environment.

DLm4080 topology

The initial release of the DLm is model 4080. It offers configurations of either eight FICON channels or 12 ESCON channels, emulates up to 1,024 drives (with four virtual tape engines), and has a maximum capacity of 190 TB for storing tape data.

The Virtual Tape Emulator Controller (VTEC) is a virtual tape subsystem that connects to an IBM mainframe channel and provides drive and tape emulation utilizing a large-scale storage array. The VTEC consists of individual engines that run the software to emulate IBM 3480/3490/3590 tape drives for mainframe lpars to access.

Each VTEC includes two or four virtual tape engines (VTE), each capable of emulating 16 to 256 tape drives. All the VTEs in the VTEC connect to the mainframe either through ESCON or FICON channels. Each VTE is also connected to the back-end storage using Gigabit Ethernet links.

When the tape management software on the mainframe creates a new tape volume, it will be stored as a file in the back-end storage system. The tape file is identified and correlated to a volume by the tape VOLSER. The VOLSER is acquired from the Load Display command sent from the mainframe whenever it wants a tape volume to be mounted for reading or writing. Tape volumes are presented to the mainframe as being mounted on an appropriate emulated drive.

While each VTE operates independently from the others, multiple Network File System exports from the back-end storage are shared across all VTEs. As all the VTEs have access to all file system mounts, any emulated tape drive can access all the tape volumes stored in the DLm4080.

The VTEs are connected to the back-end storage via dual internal VTEC Gigabit Ethernet switches. The back-end storage consists of 1 TB physical disks installed in individual disk-array enclosures consisting of 15 drives per DAE. Within each set, 14 drives are configured using RAID 6, to provide the equivalent of 12 data drives and two drives for data protection, while the remaining drive is available as a hot spare. The DLm4080 is expandable in increments of 15 drives. All of the hot spares act as a group to protect all of the drives in the entire array. A fully populated EMC DLm4080 is configured with 190 TB of capacity.

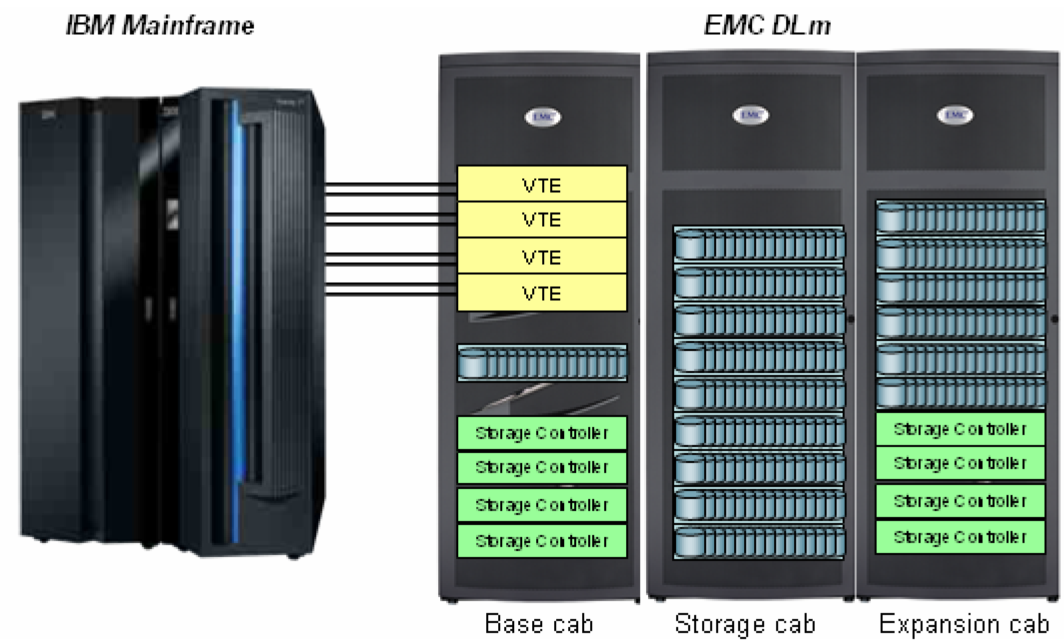

Figure 1. Maximum capacity DLm4080 configuration

Mainframe tape analysis

Planning an EMC Disk Library for mainframe installation starts with understanding the environment in which the solution will be deployed.

Characterize existing tape infrastructure

Characterizing the tape environment starts with capturing the number and type of tape drives as well as understanding the uses of tape in the environment. Armed with this information the utilization and capacity of the existing environment can be determined. Host data can also be used and is required for use when EMC performs a Mainframe Tape Analysis. Understanding the existing environment is required to properly configure the DLm solution that will be seamlessly integrated into the existing production environment.

Physical tape drives

A big part of the mainframe tape infrastructure is the physical media to which the data is written. Most mainframe environments have a variety of tape configurations. They place data on different physical tape drives and media as each device type offers different performance and capacities.

The DLm4080 supports tape drives of IBM 3480, 3490, and 3590. They are some of the most common tape drives found in mainframe data centers today.

3480

Tape drives built on the IBM 3480 tape standard were manufactured by a number of vendors from 1984 to 2004. 3480 drives still have a presence in enterprise data centers today. They are limited in that they do not do compression and have a stated capacity of 200 MB.

3490 and 3490E

The IBM 3490 is basically a 3480 with compression. The media and number of tracks on the drive are the same as the 3480. The compression ratio achieved depends entirely on the data content, but greater than 2:1 is considered typical.

Despite the name similarities, the tape devices 3490 and 3490E are quite different. The 3490E is recorded with 36 tracks across the tape, not 18 as in the 3490. It has a higher bit density and uses thinner and longer media. Like the 3490, it uses data compression for a typical capacity of 800 MB per tape. Most 3490E drives can read 3480 and 3490 tapes, but can't write them, because of the thinner recording heads for the 36 tracks

3590 and 3590E

IBM's 3590 tape drive is similar to the 3490. The half-inch tape is contained in a cartridge the same size as a 3480, but the tape itself is very different. The 3590B drive writes 128 tracks, 32 tracks at a time, in a serpentine pattern, and has a capacity of 10 GB or 20 GB uncompressed. There is also a 3590E drive that is similar to the 3590B and uses the same media, but writes 256 tracks, for double the capacity of the B drive.

Tape connectivity

In the early days of mainframe tape environments, host connectivity to the physical tape was provided over Block Multiplex Channels (BMC) with a transfer speed of 4.5 MB/s or ESCON with a transfer speed of 17 MB/s. Today, connecting channels are either ESCON or FICON; FICON channels provide significantly higher data bandwidth than ESCON and support much greater distances.

The EMC DLm4080 supports three – ESCON and two – FICON boards per Virtual Tape Emulation (VTE) node or 6/12 ESCON or 4/8 FICON per VTEC. The DLm FICON interfaces may be either with single-mode or multi-mode fiber-optic cable support. Single-mode fiber-optic cable has a micron rating of 9/125. Multi-mode fiber-optic cable has a micron rating of either 50/125 or 62.5/125.

The host connectivity options should be discussed early when planning DLm4080 deployments. Existing tape environments running ESCON that are looking to improve performance can implement a DLm with FICON but require available FICON channels to the lpars currently accessing tape. If there are plans to move to FICON it is recommended to implement VTEs with FICON as a migration from ESCON to FICON can be a complex multi-staged process that could involve DLm outages.

Current capacity

Configuring the DLm for the same number of tape drives and cartridges used in production is the simplest approach to planning configuration. This might result in oversizing the DLm as physical tapes are only 30 percent utilized.

To truly understand tape storage capacity it helps to start with identifying the lpars and applications that are dependent on tape. Are the lpars using SCRATCH pools for tape jobs or is there a requirement for PRIVATE jobs to supplement physical DASD storage? If the environment is extensive the tape usage may be pulled from the SMF21 Record output.

Once the environments tape usage has been revealed then the policies for number of copies kept, number of copies sent offsite, and the length of time the copies remain online in the physical library are determined. The capacity sizing should account for growth as well as a contingency for peak production for miscellaneous batch jobs that might spill over to tape.

DLm4080 planning

When planning the deployment of EMC Disk Library for mainframe it is important to get the actual workload throughput requirements for storage capacities with and without compression as well as the number of Channel Connects required. This information can be determined from the tape system information on each lpar managing tape resources.

Data collection

For proper planning the customer volume information from the Tape Management Catalog and the SMF21 Record information is required. IBM System Management Facility (SMF) is a component of IBM's z/OS for mainframe systems that provides full "instrumentation" of all baseline activities running on that mainframe operating system, including I/O, network activity, software usage, error conditions, and processor utilization, as well as tape. SMF21 Records produce Tape Volume (3480, 3490, and 3590) and Error statistics. These records should be collected to provide a detailed view of the actual tape workload seen in the environment. SMF records are written to a Data Set; typical SMF Data Sets are very large but the SMF21 Data Set is small and can be easily transferred via FTP.

EMC Assessment for Mainframe Tape Environments

The EMC Assessment for Mainframe Tape Environments is a consultative service that begins with interview sessions to document, understand, and analyze the current z/OS attached tape environment. Upon completion of the engagement a report documenting the findings and recommendation based on existing infrastructure and specific company business priorities is provided.

To streamline the tape assessment process EMC has Tape Tools that are used to measure existing tape workload and provide required input for sizing tape replacement solutions. The tools are available to EMC TS professionals with mainframe access and produce an output document that is a comprehensive tape utilization report.

It is highly recommended to have EMC perform a formal tape assessment to reduce risk with the deployment, accelerate batch processing, achieve higher tape reliability, and reduce downtime by improving application and system availability.

DLm4080 design

The EMC DLm4080 solution provides a tape on disk environment for mainframe customers. The product consists of two or four virtual tape engines (VTE) and EMC storage. The appliance provides the ability to emulate mainframe tape devices. The VTEs connect to EMC storage via Gigabit Ethernet interconnects to present a virtualized storage space.

The benefits of a properly designed DLm solution are to determine the right number of VTEs and the number and capacity of the back-end storage. This is based on the tape volume capacity and maximum I/O bandwidth required. The I/O bandwidth requirements will help determine the number of VTEs and type of host connection (FICON or ESCON).

Calculating the number of VTEs

To determine the required number of VTEs starts with understanding the total amount of usable storage required. To accurately predict the total amount of storage it is important to calculate a "performance" compression ratio. A "typical" compression ratio is between 2.5:1 and 3:1; if it is not known use a conservative number of 2:1.

The maximum observed tape bandwidth from the assessment report would be very helpful. Without the actual report the bandwidth numbers will be approximations and could introduce risk with the deployment.

This information gives us the worst case for required throughput numbers from the existing tape infrastructure. The formulas used to determine the number of VTE and storage controllers is as follows:

Number of VTEs = Tape bandwidth/150 MB

Estimating storage capacity

Determining the required storage capacity also requires understanding the data compression ratio. But this compression ratio is considered the "capacity" compression ratio. For this data compression ratio a best practice is to underprovision storage capacity by using a larger compression than that observed in the customer environment. Use a 4:1 compression ratio if customer data is not available.

Using the capacity compression ratio with the total utilized space derived from the customer's environment or from the tape assessment, the formula for estimating the storage capacity is:

Required storage = Total utilized space (TB) * compression ratio + 10% for SAVVOL

Calculating "performance" compression ratio

An example of determining the "performance" compression ratio is if the customer observes a compression ratio of 3:1. A performance ratio of 2:1 would overprovision capacity. This overprovisioning ensures there is room for expansion and will account for unanticipated data bursts.

If a performance ratio of 3:1 matches the capacity of the estimated requirement there is no margin for error if the estimate is wrong. This may introduce performance degradation.

If a performance ratio of 4:1 is introduced it will underprovision storage capacity where the customer could see performance degradation. This can be addressed by adding additional VTEs or storage controllers to the DLm.

A best practice for calculating the performance compression ratio is to overprovision bandwidth capacity by using a smaller compression than that observed in the customer environment. Use a 2:1 compression ratio if customer data is not available.

Calculating a "capacity" compression ratio

An example of determining the "capacity" compression ratio is if the customer observes a compression ratio of 3:1. A ratio of 2:1 overprovisions storage capacity and could potentially create stranded storage. To claim that stranded storage, additional VOLSERs would need to be added to the configuration. Adding additional VOLSERs is an offline operation that requires an outage.

A capacity ratio of 3:1 matches the capacity to estimate and could also lead to stranded storage. There is also no margin for error if the estimate is wrong.

A capacity ratio of 4:1 underprovisions storage capacity but most customers do not use the full VOLSER range. EMC includes utilities to easily reclaim space when the utilization passes a user-defined threshold. Unlike adding additional VOLSERs, adding physical storage can be added online without introducing a DLm outage.

A best practice for calculating a capacity compression ratio is to underprovision the storage capacity by using a larger compression than that observed in the customer environment. Use a 4:1 compression ratio if customer data is not available

DLm and back-end storage

The DLm is designed to work with networked (NFS) storage. It provides two Gigabit Ethernet ports on each DLm connected to the back-end storage through two gigabit switches. Once the storage is attached it must be configured to export the NFS shares for use by each VTE.

Typically, the library definition in all VTEs within a DLm will be identical to ensure that any tape drive on any VTE can access the tape volumes in the library. Configuring one VTE with four NFS shares and a second VTE with only three would mean the second VTE will not have access to any volume written to the fourth share on the first VTE.

Estimating the number of network shares

For each VTE, the tape library is controlled by a file system directory /tapelib stored on the VTE system disks. Each subdirectory must be a two-character name where the first two characters represent the tape VOLSERs that will be stored on that share. The two characters must use only valid VOLSER characters (A–Z and 0–9).

It is recommended to have no more than 16 network shares mounted and exported under tapelib that become part of the library. Each share can have a maximum of 10,000 VOLSER to initialize per NFS mount. For optimal performance create a minimum of four NFS shares to be used across each VTE.

Since each share can have a maximum of 10,000 volumes the number of shares is influenced by both the number and size of each of the individual VOLSERs.

Estimating the size of each NFS share

For the initial release of the DLm4080 the maximum capacity supported is 192 TB. The capacity that a storage controller can handle is 32 TB, which means the maximum size of a network share is 32 TB. For optimal performance network shares should be configured between 4-6 TB in size. A greater number of smaller shares often works best for locating tape VOLSERs.

The number of VOLSERs is used to determine the file system size. The number of VOLSERs can be determined using the number of datasets and the number of network shares. The number of VOLSERs can not exceed 10,000 per NFS share.

Number of VOLSERs = Total number datasets / Number of shares

The network share file system size can be derived from the total utilized space and the total number of datasets. This information can be easily extracted from the tape assessment.

File system size = Number of VOLSERs * (Total utilized space / Total number of datasets)

Configuring virtual tapes in z/OS

The DLm4080 supports three alternatives for tape device configurations. The preferred approach is using System Managed Storage (SMS) with the Manual Tape Library (MTL) on the z/OS mainframe. But there are other options available for environments that do not use or will not allow SMS managed storage.

Manual Tape Library

The preferred approach is using SMS with the Manual Tape Library (MTL). IBM introduced the concept of an MTL with APAR OW45271. This APAR allows standalone tape drives and their associated volumes to be SMS-managed by treating a group of such drives as a logical tape library. Allocations to such a logical library are managed by SMS just as any automated tape library data server (ATLDS) would be with the exception that mount messages will be routed to a tape operator console rather than the ATLDS robotics. The IBM document *SC35-0427, "PISA Guide for Tape Libraries"* describes the issues regarding MTL support.

Note that once the DLm is configured, it is treated much the same as a real library. That is, cartridges must be "entered" into the library before they can be used. To simplify this, the DLMLIB utility may be used to enter cartridges into the MTL.

MTL-related IBM maintenance

The PTFs for the following APARs should be applied when using DLm in an MTL environment:

- APAR OA03749
  - MTL more than one device fails to vary online
- APAR OA06698
  - Replacement tape drives get MSG IEA437I in an MTL environment
- APAR OA07945
  - Mount hangs / times out using MTL with OEM Automated Library
- APAR OA08963
  - Tape volume capacity incorrect for OAM object support users
- APAR OA10482
  - MTL scratch volume mount error

Always check with IBM for any new MTL APARs.

Real Tape

DLm can emulate 3480, 3490, or 3590 tape drives. Real 3480, 3490, or 3590 device configuration is used when the device does not exist to the host or for environments that do not use SMS managed storage. StorageTek HSC is a common configuration that requires the Real Tape virtual tape type to access the EMC DLm4080 virtual tape devices. Environments looking to replace the existing StorageTek environment with EMC Disk Library for mainframe should consider introducing SMS with MDL and having the two co-exist until StorageTek could be aged out.

However if you have real 3480, 3490, and 3590 tape drives configured in your system do not attempt to define MDL devices in this manner. Configuring the devices as a device type that is already present will result in misallocation errors where z/OS will request a real 3480/3490/3590 cartridge on a device or request a tape-on-disk volume on a real 3480/3490/3590.

EMC UIM

As an alternative to defining Real Tape 3480s, 3490s, or 3590s or using a MTL, EMC provides a user Unit Information Module (UIM) that allows DLm tape devices to be configured to HCD using a unique device type. Using the EMC UIM prevents the operating system from allocating DLm virtual tape drives to jobs requesting a mount of a real tape cartridge.

The UIM provides for four control unit types of V3480, V3481, V3482 and V3483; they support device types V3480, V3481, V3482, and V3483, respectively. The generic names for these devices are also V3480, V3481, V3482, and V3483. If you have already defined a generic name of V348*x*, please contact EMC Customer Support. Multiple virtual device types are defined to support multiple DLm systems or a single DLm with multiple virtual tape libraries configured.

Considerations, limitations, and restrictions

EMC Disk Library for mainframe (DLm) is a complete disk library system for IBM mainframe environments. EMC Disk Library offers compelling total cost of ownership, performance, and operational benefits without requiring difficult operational changes for backup and tape resident data sets.

DLm4080, the first release in the EMC Disk Library for mainframe 4000 series, is a first generation tape-on-disk system for IBM or IBM-compatible mainframes. DLm provides a redundant, highly available system offering four to eight Fiber Connection (FICON) or six to 12 Enterprise Systems Connection (ESCON) channels for each DLm and up to 1,024 tape devices of type 3480, 3490, or 3590.

Host components and release versions

System components	Description	Notes and release versions
Mainframe host environment		
Operating systems	z/OS, OS/390, and MVS environments	
Tape emulation	3480/3490/3590 tape drives	A DLm with four VTEs can concurrently emulate a maximum of 1,024 tape drives
Utilities software	DLMSCR. Sends VOLSER scratch requests to DLm. DLMCMD. Allows the mainframe console to send DLm commands. DLMLIB. Defines scratch volumes to a Manual Tape Library (MTL).	DLMZOS.XMI release 4.0

DLm internal disk drives

The Access Control Point (ACP) and Virtual Tape Emulator (VTE) units in a DLm each contain two front-access disk drives attached to an internal SCSI adapter. These drives are used exclusively by the code running within the ACP and VTE. No other use of these drives is supported. No user data or "state" data is kept on these drives. The drives function in a RAID 1 mirrored pair arrangement for high availability. In the event one of these drives fails, contact EMC Customer Support to schedule a replacement.

Disaster recovery

The EMC Disk Library for mainframe uses asynchronous remote replication to copy network share information to a remote site for disaster recovery purposes. For disaster recovery solutions the tape catalog must be available on the remote system. Mainframe tape environments have their tapes catalog, media labeled, and stored in a tape library. Backup software like CA-1 maintains a separate catalog, TMC, of all registered tapes. The tape catalogs are required for data recovery. Since every environment has site-specific processes for tape catalog management the individual procedures followed today for offsite tape storage should be followed with DLm remote replication.

Returning scratch volumes

When a scratch tape is mounted and then unmounted without being accessed by the host, it is not returned to scratch status when it is unmounted. A scratch tape is removed from scratch status when it is mounted even if it was never accessed by the host. It is recommended to run the DLMSCR utility after the initial testing and before going into production to return any test scratch tapes.

Conclusion

EMC DLm4080 for mainframe has performance gains over physical tape. It offers better overall utilization than physical tape and other existing virtual tape solutions. The DLm 4080 leverages proven EMC technology to provide the best in performance, availability, and reliability.

The DLm works in just about any existing z/OS environment that currently deploys tape devices for either supplementing DASD or for batch jobs providing a secondary copy of production data. The DLm supports Manual Tape Libraries integrated with SMS as well as environments that do not use SMS. A properly configured DLm solution starts with planning and a detailed design.

With proper planning and an optimized design a solution can be configured that provides increased reliability and availability over other physical tape media. Advantages such as improved restores from tape resources, faster access to tape resources, and no host reconfiguration provide a compelling case for deploying an EMC Disk Library for mainframe 4080 solution.

References

The following list includes reference materials used to create this paper:

- *EMC Disk Library for Mainframe DLm4000 Series User Guide*
- *EMC Disk Library for Mainframe DLm4000 Series Release Notes*
- *EMC Disk Library and Symmetrix DMX Series Unpacking Guide*
- *EMC Disk Library for Mainframe DLm4000 Series Physical Planning Guide*