Introduction

System z data centers have long used tape devices for many reasons, certainly predating the origins of z/OS. During all that time, tape and DASD data have always existed in separate but equal domains. That is tape and DASD data were routinely managed apart; control data sets for tape data and ICF catalogs for DASD data. Moreover, this divergence extended to tape and DASD data replication services. As a result, tape data were often difficult to use during automated Sysplex recovery.

Consequently, one of the reasons EMC® introduced their new Disk Library for mainframe 8000 (DLm8000) virtual tape system was to join tape and DASD data closer together, at least for replication and recovery. With the new DLm8000, virtual tape data can now be replicated using the same capabilities as DASD data and more importantly, employ similar recovery tools. Thus, with the new DLm8000, applications that use both virtual tape and DASD should be easier and swifter to recover.

Enterprise business continuity/disaster recovery concerns

Business continuity and disaster recovery (BC/DR) have been crucial considerations for many executives ever since 9/11. That tragedy inspired System z data center staff and their regulators to rethink their DR alternatives. One outcome of this analysis was the need for three-site solutions that could recover from disasters within a single locality as well as region-wide catastrophes. These multi-site recovery environments required replication and restart services that could support multiple failover scenarios.

In addition, one continuing challenge for mainframe recovery planning was the lack of any coordination between virtual tape and DASD replication, which was further aggravated under three-site recovery. Today, System z data centers use virtual tape for many purposes, e.g., backup, migration, checkpoint, etc. As such, tape and DASD data must both be replicated to alternate sites to support application restart.
EMC’s New **DLm8000 for System z Data**

However, standard DASD replication supports multiple, sophisticated mirroring options while virtual tape replication is only done after end-of-volume (EoV) processing. In z/OS, tape EoV replication occurs minutes after job step termination or dynamic tape data set un-allocation. In contrast, DASD data set replication can be done on a write IO basis. As a result, virtual tape and DASD data were often replicated at different times and ended up being hard to correlate during disaster restart activities. This was especially problematic for tape management catalogs residing on DASD that controlled the current state of virtual tape data.

So when disaster struck, operations had to expend time and effort determining which was most current, the virtual tape or the DASD data replicated to the secondary site. This generated two distinct cases during recovery restart, namely:

- **Replicated DASD data was more current than tape data** causing IT staff to back out or ignore recent DASD resident control data set updates that didn’t match current virtual tape data available. This situation was very likely given tape’s EoV replication and DASD’s more frequent, IO based replication.
- **Replicated virtual tape data was more current than DASD data** driving operations to throwaway the most recent virtual tape data to coincide with the most up-to-date DASD data on hand or alternatively, to fabricate control data set catalog entries. This scenario was less plausible but had to be considered.

If replication breakdowns were occurring, recovery restart could be even more complex. For example, if virtual tape data replication had failed because of networking problems, remote storage issues, etc., it was probable that DASD replication would continue to operate and be unaffected by these outages. Similarly, if DASD data replication failed for some reason, it was distinctly possible that virtual tape data replication would continue, unimpeded and unaware of any DASD replication issues.

Failures such as these led to ever more extreme versions of the two cases discussed above. But here virtual tape and DASD replicated data could be disconnected by hours or even longer time frames. The central, overriding problem could not be overcome, i.e., virtual tape and DASD data replication could not be automatically coordinated and therefore, were replicated independently.

**The new DLm8000 system**

EMC introduced their new DLm8000 storage to resolve these and other issues with the use of virtual tape in System z enterprise data centers. EMC’s solution to this separation of DASD and virtual tape replication was to use VMAX storage behind the DLm8000. With the new DLm8000, mainframe customers can now take advantage
EMC’s New **DLM8000 for System z Data**

of VMAX sophistication to replicate virtual tape data similar to EMC based DASD data. At last, joining together virtual tape and DASD data under the same replication engine, under the control of a single management interface with identical replication capabilities.

In addition to the VMAX backend storage, the new DLM8000 also offers better throughput, more virtual tape drives and higher storage capacity. These other new features are all discussed later. However, the next section will focus on VMAX replication and recovery services for virtual tape and how they can improve System z disaster restart for system services and applications.

**DLM8000 advanced replication and recovery**

Replication for DLM8000’s backend storage operates just the same as any primary storage VMAX system. That is, all VMAX storage systems use Symmetrix Remote Data Facility (SRDF®) for remote replication. SRDF provides a number of advanced capabilities but they all depend on two underlying services, namely

- **SRDF/S** is Symmetrix **synchronous** replication. With SRDF/S, data written to the local VMAX storage is copied to the remote VMAX system before the IO operation can complete at the host channel.
- **SRDF/A** is Symmetrix **asynchronous** replication. With SRDF/A, the IO operation causing data to be written to the local VMAX system is completed at the host channel before the data can be copied to the remote VMAX storage.

Consequently, SRDF/S has distance limitations, i.e., it can only be used for sites that are less than 200 Km apart. Also, as SRDF/S copies data while the IO executes, writes will take longer to complete. As a result, SRDF/S is mainly used for short distances, e.g., across a data center hallway, on different floors of a building or across a small region.

In contrast, SRDF/A has no distance restrictions and as the copy takes place after the write IO completes, has little to no impact on performance. As such, for SRDF/A a remote VMAX system could easily be 1000 Km away from the local site and replication works just fine. Thus, SRDF/A is often used to support out-of-region recovery sites.

In addition, System z VMAX storage supports **SRDF/CG or consistency groups**, for replication uniformity between multiple VMAX devices, even across devices residing in different VMAX storage systems. SRDF/CG provides application points of consistency for easier disaster restart. As such, virtual tape volumes on DLM8000 and DASD data on primary VMAX storage devices can now form a single consistency group, which provides some unique features, specifically,
EMC’s New **DLM8000 for System z Data**

- Any disruption that halts one VMAX replication activity automatically causes all other replication within the same consistency group to cease, thus maintaining failover congruity across all DASD devices and virtual tape volumes within the group.
- Operations staff can now start and/or stop replication activity across both virtual volumes and DASD devices within a consistency group at the same time, with one command.

Prior to DLM8000, consistency groups were reserved for when multiple DASD device data sets needed to be replicated together, keeping all of them in agreement with one another. For instance, databases and its journal data often reside on separate devices both for higher availability and better performance. But database and log data must be replicated in synch with one another, as having a more current copy of one without the other will lead to lost transactions. Hence, a single consistency group is used to connect the two devices that hold this data. Thereby, a current copy of both the database and log data is always available at the replication site whenever disaster strikes.

But now with the new DLM8000 backend VMAX storage these consistency capabilities also support virtual tape. As such, SRDF/CG can now link multiple virtual tape volumes and/or multiple DASD devices to be replicated together.

In addition, with the DLM8000, virtual tape data can now be incorporated into sophisticated VMAX three-site DR solutions. Currently, mainframe VMAX storage supports **SRDF/STAR**, a three-site solution that uses a combination of SRDF/S and SRDF/A to replicate synchronously within region and asynchronously outside a region. VMAX consistency groups function just as well in SRDF/STAR configurations.

There are more VMAX SRDF capabilities than those described above, such as **SRDF/Concurrent, SRDF/Cascaded** and **SRDF/Extended Distance Protection**, which use a combination of SRDF/S and SRDF/A to extend synchronous level protection over asynchronous distances (>200 Km). All these can also be used with DLM8000 virtual tape volumes, the same as any VMAX DASD device.

Equally important, for z/OS environments VMAX and DLM8000 SRDF can now be controlled from the **EMC z/OS Storage Manager (EZSM)**. All VMAX SRDF replication services and other VMAX advanced storage features can be managed from this host based, EZSM ISPF panels. Alternatively, the **Symmetrix Management Console (SMC)** provides an array-based solution to manage DLM8000 and VMAX storage replication services. SMC is a web based GUI management system, supporting all VMAX storage systems in a data center. With either solution, VMAX and DLM8000 replication can now be managed from the same, single pane of glass.
EMC’s New **DLm8000 for System z Data**

Finally, for faster Sysplex recovery EMC offers their **Geographically Dispersed Disaster Restart (GDDR)** software that supplies within region and out-of-region restart automation. Using EMC GDDR with VMAX primary and secondary storage, z/OS recovery can be fully automated, eliminating all unnecessary operator intervention. With VMAX SRDF and EMC GDDR, System z data centers can easily support both planned and unplanned downtime for two- or three-site recovery environments. EMC GDDR automatically recovers your EMC DLm8000 virtual tape and VMAX DASD storage, z/OS host systems and critical applications supplying rapid System z recovery restart.

**DLm8000 replication use cases**

There are a number of specific examples of these new virtual tape SRDF capabilities that can be very advantageous for mainframe environments. Specifically, DLm8000 replication can be a vital resource for z/OS service recovery such as HSM, DB2 journal archives, checkpoint/restart and many other tape intensive data processing activities.

**HSM (DFSMShsm) data**

For example, with HSM migration, primary DASD data sets such as the Migration Control Data Set (MCDS), Backup Control Data Set (BCDS) and Offline Control Data Set (OCDS) record location information as data moves up and down its storage hierarchy. Previously, when HSM migrated data to virtual tape the OCDS and its tape table of content (TTOC) records could easily become out-of-synch with virtual tape replication to a remote site. However, with the new DLm8000 consistency groups, this no longer has to happen, as the virtual volume and OCDS DASD replication could now be maintained in-synch, to the last write operation, if necessary.

**DB2 active and archive journals**

In addition, DB2 database transaction activity is journaled to active VSAM data set logs that then, over time are closed and archived to virtual tape. By creating a consistency group that links both DB2 primary and secondary storage, using the same replication services for both, operations can now guarantee that all journal and DB2 database data at the remote site corresponds to one another.

**Checkpoint/restart jobs**

Similarly, z/OS applications have always been able to exploit checkpoint/restart services to insure that long running batch jobs complete in a timely fashion. Large checkpoint data sets are often written to virtual tape and thus, can be inconsistently replicated with any primary storage updated by these jobs. With the new DLm8000, both the checkpoint virtual tape and an applications DASD data can now be replicated together, utilizing a single consistency group that traverses both.
Oil and gas seismic analysis

Moreover, many oil and gas seismic processing applications will read sensor data and record the current status of their analysis to both virtual tape as well as primary storage. Once again, with consistency groups that bridge both virtual tape data under DLm8000 and primary DASD data, VMAX can provide consistent and reliable replication of all this data to remote sites.

Disk Library for mainframe product family

EMC’s portfolio of Disk Library for mainframe (DLm) systems can support from the largest, most critical System z enterprise customer needs to the smallest, entry-level data centers. EMC’s DLm storage includes the DLm8000, the DLm6000 with concurrent support for both deduplication and disk storage, the low-end DLm2000 and for mixed environments, and the DLm1000 utilizing Data Domain deduplication storage.

DLm8000 for DR/performance critical data

In addition to the use of VMAX discussed above, the new DLm8000 hardware supports from two to eight virtual tape engines (VTE) in its VTEC bay. Each VTE emulates up to 256 drives and for the DLm8000, this can provide a maximum of 2048 virtual devices. A DLm8000 can be configured with from four to sixteen FICON channel attachments for increased flexibility.

Also, from ~56 TB to up to ~1.8 PB of user capacity is available on the DLm8000 backend storage. The VMAX can be scaled up from two to eight engines to provide the highest storage performance. Note, the DLm8000 dedicates this VMAX storage system to servicing virtual tape activity and it cannot be shared for any other purposes.

With more VTEs and VMAX high performance storage, the new DLm8000 can now ingest data at up to 2.7 GB per second. This is 35% faster than the previous, best performing Disk Library for mainframe system.

DLm6000 for deduplicated backup and traditional virtual tape data

The DLm6000 supports from two to six VTEs housed in a VTEC for up to 1,536 virtual devices used for both backup and other performance demanding data streams. Further, it can be configured with from four to twelve FICON attachments.
In contrast to the DLm8000, the DLm6000 utilizes up to a pair of backend storage systems that consists of an EMC VNX7500 and/or EMC Data Domain DD890 appliance. This dual storage capability provides increased flexibility as the DLm6000 can have up to one VNX7500 and one DD890 appliance or two of either. As such, the DLm6000 can be adapted for use by different System z activities, e.g., backup workloads, traditional tape batch activities or both at the same time.

With Data Domain storage behind it, the DLm6000 can deduplicate System z backup data. Consequently, with up to 570 TB of raw storage capacity, the DLm6000 could potentially store up to 5.7 PB of deduplicated mainframe backup information. Alternatively, a DLm6000 using VNX storage can ingest virtual tape data at 2.0 GB per second, providing excellent performance for the more traditional mainframe tape processing workloads.

**DLm2000 for entry-level data centers**

For entry-level System z users, EMC offers their DLm2000 system. The DLm2000 can be configured with one to two virtual tape engines (VTEs), emulating up to 512 virtual devices and can attach from two to four FICON channels.

The DLm2000 uses a VNX5300 that supports from 13 to 143 TB of usable backend storage. However, a data compression feature is available on the DLm2000 and when used with the maximum configuration, can effectively multiply this capacity by two to three times.

**DLm1000 and Data Domain storage for mixed system environments**

The DLm1000 differs from other members of the DLm family in that it also supports distributed systems use of its backend storage. It offers both a mainframe tape emulation front-end supporting up to 128 virtual devices as well as distributed systems direct access to the backend EMC Data Domain appliance. The DLm1000 can be configured to support from one to two FICON attachments and direct access to the Data Domain appliance is available using multiple Ethernet and/or Fibre Channel attachments.

Any Data Domain appliance can be used as the DLm1000’s backend storage and it can ingest z/OS data at a rate of up to 300 MB per second. As a result, the DLm1000 can be ideal for smaller data centers that need both System z virtual tape data as well as distributed systems deduplicated backup storage.

**Summary**

The EMC DLm8000 has expanded the already comprehensive EMC Disk Library for mainframe portfolio by enlarging storage capacity, increasing ingest performance and adding virtual devices. With the introduction of the DLm8000, EMC has now
EMC’s New **DLm8000 for System z Data**

assembled a complete family of products that encompasses all uses of virtual tape in System z data centers.

More importantly, with the new DLm8000, EMC has introduced cutting-edge replication and recovery innovations to System z virtual tape services. These new technologies allow both secondary and primary storage to be replicated together, within a single consistency group, vastly simplifying mainframe recovery. Not only is the full power of SRDF available for DLm8000 data but virtual tape and DASD data can now be jointly recovered under EMC GDDR z/OS recovery automation.

As a result, the new DLm8000 has transformed virtual tape data recovery from an isolated, cumbersome activity to a storage-wide, coordinated and simplified process. With the DLM8000, EMC has added a new dimension to System z storage, never before possible that will revitalize the use of virtual tape in mainframe data centers.

---

Silverton Consulting, Inc. is a Storage, Strategy & Systems consulting services company, based in the USA offering products and services to the data storage community.