White Paper

Considering the EMC DLm8000 for OCC Compliance and Operational Recoverability

By Jason Buffington, Senior Analyst, and Monya Keane, Research Analyst

February 2013

This ESG White Paper was commissioned by EMC and is distributed under license from ESG.

© 2013 by The Enterprise Strategy Group, Inc. All Rights Reserved.
# Contents

Introduction .................................................................................................................................................. 3  
Understanding the Costs of Compliance ... and Non-compliance ................................................................. 4  
The Difference Between Compliance and Readiness ..................................................................................... 4  

The Office of the Comptroller of the Currency: Regulations and Ramifications ........................................... 5  
What Is the OCC, and What Does It Have to Do with IT? ........................................................................... 5  
OCC Regulations and Recommendations .................................................................................................... 5  
OCC Ramifications ....................................................................................................................................... 6  

Overview of EMC DLm8000 .......................................................................................................................... 6  
Product Characteristics ................................................................................................................................. 6  

Looking at This Technology Through the Eyes of the OCC ........................................................................ 8  
High Availability Design = Continuous Availability .................................................................................. 8  
Replication = Assured Data Survivability ........................................................................................................ 8  
Consistency Groups = Faster Recovery .......................................................................................................... 9  
Geographically Dispersed Disaster Restart (GDDR) = Disaster Recovery and Restart ....................................... 9  

Understanding BC/DR in a Regulatory Context .......................................................................................... 10  
Understanding Tape in Regulatory Context .................................................................................................. 10  

The Bigger Truth ........................................................................................................................................... 11
**Introduction**

Say “regulatory compliance” to an IT manager, and you might hear an exasperated sigh in response. That’s because it is often the IT team that is tasked with ensuring that the organization adheres to numerous and often stringent data protection-related rules and regulations. Most of those compliance mandates don’t come packaged with explicit instructions or even guidelines on how IT can achieve compliance, either; they often dictate a result without specifying methods to attain it. In fact, risk reduction and regulatory compliance initiatives have a significant effect on IT investment in all industries (see Figure 1).

**Figure 1. 2012 Business Initiatives with the Greatest Impact on IT Spending Decisions**

Which of the following business initiatives do you believe will have the greatest impact on your organization’s IT spending decisions over the next 12-18 months? (Percent of respondents, N=614, three responses accepted)

- Cost reduction initiatives: 43%
- Business process improvement initiatives: 30%
- Security/risk management initiatives: 29%
- Regulatory compliance: 24%
- Improved internal collaboration capabilities: 18%
- Business growth via mergers, acquisitions, or organic expansion: 17%
- Improved business intelligence and delivery of real-time business information: 16%
- Increased use of social networking technology for marketing, customer outreach, market research, etc.: 14%
- Green initiatives related to energy efficiency and/or reducing company-wide environmental impact: 13%
- International expansion: 13%
- Research and development innovation/improvement: 12%


Risk management (item 3) and regulatory compliance (item 4) seem to go hand in hand. But it is notable that besides the top item, cost reduction, business process improvement is often the most tangible and rewarding benefit of any business continuity (BC) or disaster recovery (DR) plan.

---

Understanding the Costs of Compliance ... and Non-compliance

Soaring regulatory compliance costs can put a dent in an IT budget. But it is a pay-now-or-pay-later situation. Compliance-related expenditures are a “necessary evil” because the cost of non-compliance is often much higher. In some cases, regulatory non-compliance results in penalties, fines, pricey settlements, loss of reputation and revenue, plunging market capitalization ... and occasionally, handcuffs. Non-compliance problems have cost some companies billions of dollars.

Worse yet, many IT and business managers don’t know exactly what risks their business faces because they work in such a frequently changing regulatory landscape. In some cases, they may not even have adequate staff resources to oversee the compliance effort. It’s a real challenge for some IT groups, and clearly, the stakes are high.

The Difference Between Compliance and Readiness

When ESG looked at the top planned data protection investments in 2012 (see Figure 2), it found that the top two investments centered on “improving disaster recovery capabilities” and “meeting compliance requirements.”

Figure 2. Top Areas of Planned Investments in Data Protection for 2012

Why measure actual DR capabilities separately from compliance? Because they are very different.

Passing a regulatory audit doesn’t necessarily ensure that a business can recover operationally following a significant services interruption. Conversely, being able to maintain continuity of business operations may not ensure that the company can pass a regulatory audit. Compliance with regulations related to data protection is sometimes based mostly on subjective assessments by the auditor.

Source: ESG Research Report, Trends in Data Protection Modernization, August 2012. Please note that Figure 2 reflects the results of an enterprise-level data protection survey encompassing mainframe and open systems environments alike.
Regulations are not only about technology. They certainly don’t mandate use of one specific data protection scheme, one high-availability approach, one disaster recovery method, or one specific vendor’s product that “makes you compliant” when you deploy it. Implementing a particular technology will rarely automatically make you compliant with an entire regulation; it will merely enable you to check one or more boxes related to one tactical objective tied to one section of a much broader rule.

That being said, technology usage is a key component of both operational recoverability and achieving regulatory compliance, and that means the burden still falls on IT.

For example, it is the IT department at a regional bank that would have to ensure consistency of tape data between a production and a recovery site to avoid exposing the bank to fines. It is that bank’s IT team that must ensure predictable recovery time objectives and recovery point objectives (RTOs/RPOs) in an extraordinarily transaction-heavy environment. And it is that bank’s IT team that will be partly responsible for reestablishing timely, organization-wide operations following a disaster.

The Office of the Comptroller of the Currency: Regulations and Ramifications

A bank’s IT team is a good example for showcasing regulatory-related data protection pressures. Financial institutions in the U.S. are among the most heavily regulated industries in the world. The data that they are processing is quite sensitive (credit information, bank account numbers, Social Security numbers), not to mention being voluminous, frequently changing, and hard to replace should data loss occur. This situation is compounded by the fact that a large U.S. financial institution’s inability to process data (literally money) would affect not only that institution, but also other large institutions (such as those buying/selling stock or mortgage funds), downstream U.S. banks, and even the global economy. Thus, banks are subject to rigorous oversight.

What Is the OCC, and What Does It Have to Do with IT?

The Office of the Comptroller of the Currency (OCC), part of the Department of the Treasury, is the primary regulator of banks and federal savings associations in the United States. The OCC regulates and supervises about 2,000 national banks and 50 federal branches of foreign banks in the U.S., accounting for more than three-quarters of the total assets of all U.S. commercial banks.

The OCC has been around since 1863: It’s had plenty of time to become very good at imposing regulations on financial institutions. Its charter allows it to issue/enforce rules and regulations concerning banking practices and to take supervisory action against banks that don’t conform to laws and regulations or that otherwise engage in unsound banking practices. At the OCC, the Chief Counsel’s Office advises bank examiners (regulatory enforcement auditors) and ensures compliance by national banks and federal savings associations through administrative action (i.e., it hands down fines and penalties).

The OCC’s scope of influence extends even beyond supervising traditional banks. It is also empowered to impose regulations and penalties on big trading companies, on banks dealing directly with those traders, and on financial institutions dealing with commodities.

OCC Regulations and Recommendations

After the terrorist attacks of 9/11, the OCC, along with fellow agencies the SEC and the U.S. Federal Reserve, published the Interagency Paper on Sound Practices to Strengthen the Resilience of the U.S. Financial System. Among the many best practices that the paper describes, it emphasizes that financial institutions should maintain sufficient geographically dispersed resources to meet recovery and resumption objectives following a business-interruption event. (Some financial firms in lower Manhattan couldn’t resume operations after 9/11 using their secondary facilities because the sites were too close to the primary facility. The attack impaired both locales.)
OCC Ramifications

That interagency paper did not specify particular technologies that banks should install in their data centers. But it did make a few technology-oriented recommendations, including that:

- “The business continuity planning process should take into consideration improvements in technology and business processes supporting back-up arrangements and the need to ensure greater resilience in the event of a wide-scale disruption.”

- “Firms that use synchronous back-up facilities or whose back-up sites depend primarily on the same labor pool as the primary site should address the risk that a wide-scale disruption could impact either or both of the sites and their labor pools. As part of their ongoing planning process, firms with such back-up arrangements should strive to develop even more distant data back-up and operational resources that prove sufficient to recover clearing and settlement activities within the business day on which the disruption occurs.”

- “Plans should provide for back-up facilities that are well outside of the current synchronous range that can meet within-the-business-day recovery targets.”

Those excerpts are intriguing. The “well outside of the current synchronous range” mandate is particularly interesting, seemingly indicating that banks should be setting up secondary sites to replicate and protect data asynchronously. And the statement that facilities must be able to “meet within-the-business-day recovery targets” statement is equally interesting, implying that tape recovery would be inadequate, making a disk-based protection model a must-have.

So, is the only way to achieve the OCC’s recommendations to use an asynchronous disk-based protection? It used to be the only way. But IT mainframe owns may now have a new option—following the announcement of a specialized virtual tape library/subsystem (VTL) from EMC called the DLm8000.

Overview of EMC DLm8000

EMC describes the latest offering in its Disk Library for mainframe (DLm) family as the first mainframe disk library from EMC to deliver synchronous replication. The DLm8000, unveiled by EMC in August 2012 and made generally available in Q412, utilizes the vendor’s VMAX storage arrays and, more pertinently, it can leverage EMC’s SRDF synchronous replication technology—thus ensuring consistency of tape data at production and disaster recovery sites that may be physically far apart. According to EMC, large enterprises that are z/OS mainframe users will experience highly granular and predictable recoveries by using this new disk library, enjoying accelerated RTOs and more flexible RPOs.

Notably, the DLm8000 can leverage EMC VMAX storage—supporting both synchronous replication (SRDF/S) between two sites that are geographically close, and supporting asynchronous replication (SRDF/A) for an out-of-region recovery site. EMC says it has observed that, in many cases, large financial institutions employ both replication methods in what is known as an SRDF/STAR configuration. The approach results in a near-zero RTO to the local site and a longer RTO for the out-of-region site.

Product Characteristics

In its most basic definition, an EMC Disk Library for mainframe is a virtual tape library (VTL). Various VTL approaches have been available for years, but the earlier solutions focused heavily on tape-stacking efficiencies. In the beginning, a few IT scenarios (data protection as well as traditional tape jobs centered on batch, sorting, HSM, etc.) demanded better performance and manageability than traditional tape could offer. To solve that problem, storage vendors began providing disk arrays that “appeared” to be tape devices but could perform like the underlying disk solutions they actually were. Of course, mainframe computers absolutely understand how to work with tape systems, and these VTLs were able to meet performance demands without the IT team needing to change anything else in the environment.
EMC and Bus-Tech, the Bedford, Massachusetts-based VTL vendor that it acquired in 2010, assert that they were the first vendors in the IT industry to offer tape-on-disk for mainframe environments. EMC can lay legitimate claim to driving innovation and raising the standards for mainframe tape with the Disk Library for mainframe product family, which it initially unveiled in March 2008. The Disk Library for mainframe wasn’t focused on tape stacking; it was focused on performance and throughput of tape jobs.

Fast forward: Computing power and the volume of data being managed by IT both continue to grow rapidly, and the need for faster performance and manageability continues to be met by higher-performing disk solutions. Thus, it should come as no surprise that EMC, a major provider of enterprise disk systems, would continue to innovate around VTL systems, addressing ever-increasing storage demands through disk that can appear to be tape.

Two solution components are present within the Disk Library for mainframe family of storage products (see below). A Disk Library for mainframe is basically composed of one or more controllers plus one or more storage arrays—which large, mainframe-using organizations probably have installed in their data centers already. Incidentally, what differentiates this product family in the market today is that EMC leverages a consistent front end to the mainframe while leveraging various EMC storage offerings to meet a variety of scale, throughput, and price/performance needs articulated by EMC’s customers. The technology includes:

- **Virtual Tape Engine (VTE) controllers**—VTEs present themselves and their underlying storage capacity as tape controllers. Based on technology that EMC picked up in the Bus-Tech acquisition, VTE controllers enable EMC to offer a range of underlying storage to systems that need it.

- **Disk Library for mainframe storage arrays**—EMC provides storage solutions through several parts of its portfolio, including select offerings from its VMAX and VNX lines when Disk Library for mainframe systems are used as primary storage, and offerings from the EMC Data Domain deduplication family when Disk Library for mainframe systems are used for data protection and recovery.

With the combination of VTEs and the underlying storage components, a variety of Disk Library for mainframe offerings exist:

<table>
<thead>
<tr>
<th>Table 1. Available EMC DLm Configurations and Use Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>DLm1000</strong></td>
</tr>
<tr>
<td>A one-VTE engine: meant for deduplication scenarios</td>
</tr>
<tr>
<td>Storage: EMC Data Domain</td>
</tr>
<tr>
<td>Description: a one-engine, Data Domain-only gateway well suited for SMBs. It is compatible with open systems also.</td>
</tr>
</tbody>
</table>

*Source: Enterprise Strategy Group, 2013.*
Looking at This Technology Through the Eyes of the OCC

Looking at the highly regulated and well-defined banking industry, we can consider how a few DLm8000 capabilities can apply to regulatory mandates such as those from the OCC and SEC.

High Availability Design = Continuous Availability

The modularity of EMC Disk Library for mainframe solutions (see Table 1) is not just about right-sizing according to customer size or use case; the component-level structure enables customers to grow through the Disk Library for mainframe family without interruption—including across versions of controllers and across storage array families. After all, if an organization is going to embrace the use of VTLs for their performance, manageability, and reliability characteristics, the organization will have little to no tolerance for downtime caused by system upgrades and capacity changes. The modular architecture provides HA, scale, and RAS in a single platform—and this HA design provides continuous availability.

Because the VTEs are independent, they can be upgraded (or suffer individual outages) without affecting the Disk Library for mainframe overall. And because the storage is modular to the Disk Library for mainframe, systems can be upgraded or replaced while the overall configuration is preserved. Similarly, EMC designed the system to accept software updates in a non-disruptive manner.

For example, EMC reports supporting customers who have been using this product since EMC announced its first generation (based on the EMC Celerra NS-4080). Often, those EMC customers would purchase a second-generation Disk Library for mainframe (i.e., a DLm960), and because the generations of storage are based on the same backend architecture, the customers have been able to migrate their data off older Celerra systems, onto newer-generation Celerra systems, then to the current DLm6000, which uses VNX. Again, those migrations were possible because the systems were all based on the same technology. This fast migration represents a radical departure from the old-fashioned, CPU “overhead-heavy” tape-centric migration approach, which required the reading of all data on the tapes through the mainframe. Disk Library for mainframe leverages these storage-to-storage replication capabilities to eliminate going through the host system, resulting in a lower TCO.

Replication = Assured Data Survivability

One of the key features of the DLm8000 and its underlying VMAX storage is an ability to use EMC’s synchronous (SRDF) and asynchronous (SRDF/A) replication mechanisms. As such, multiple systems can replicate data for business continuity and disaster recovery purposes. For example, a large financial institution might synchronously replicate its data from Manhattan to New Jersey to provide continuous availability at a systems or site level, while also replicating asynchronously to Philadelphia, Boston, or Phoenix for long-distance disaster recovery.

- In the language of the OCC, “Plans should provide for back-up facilities that are well outside of the current synchronous range that can meet within-the-business-day recovery targets.”
- In the language of EMC customers, SRDF has long enabled nearby facilities to provide continuous operation, but regional outages due to power or weather had a strong potential to affect both of the sites. Adding SRDF/A capabilities enables Disk Library for mainframe users to ensure that is data is stored “well outside of the current synchronous range” and ensures “greater resilience in the event of a wide-scale disruption.”

And while fast (SRDF/S) and far (SRDF/A) copies are ideal for multi-site protection, EMC Disk Library for mainframe also utilizes EMC TimeFinder technology for multiple versioning and near-instantaneous restoration from within a single storage array, as a means of assuring viable data usability.
Consistency Groups = Faster Recovery

The ability to replicate data synchronously or asynchronously is vitally important, but it is just one piece of the data protection scheme. What happens if the tape catalog residing on DASD and replicated to the DR site is not in sync with the data residing on tape at the DR site? In a recovery situation, what would be the impact of that scenario? Well, for example, if the DFSMShsm catalog says certain data is at the secondary site but it’s not, then data loss results.

Disaster recovery plans need to take into account how to get tape and DASD data synchronized to achieve exact point-in-time recovery capability such that tape and DASD data are consistent. The DLm8000 leverages the VMAX technology stack to ensure that tape data is consistent with DASD data and that both replicate to the DR site(s) with that same level of consistency.

Essentially, using consistency groups (EMC refers to them as ConGroups) with SRDF replication technology ensures that critical data is replicated safely with each I/O, and it ensures that recovery, if/when needed, is imbued with “ultra” resiliency to meet the OCC’s “...within-the-business-day recovery targets” directive in the shortest possible time. The VMAX/DLm ConGroups help users avoid wasting time getting tape and DASD consistent to the same point in time during critical recovery periods when time is money and every minute lost could also mean millions of dollars lost. EMC regards such “universal data consistency” as a key benefit of the DLm8000 when coupled with VMAX DASD for “ultra” resiliency.

Geographically Dispersed Disaster Restart (GDDR) = Disaster Recovery and Restart

Whether considering regulatory mandates related to being able to recover from a remote site, or business requirements related to ensuring systems availability, “real” DR is always part of the equation. Said another way, it isn’t enough for an organization to prove that its data is in another location and ensured to survivable. The organization also must possess the technology and the processes to ensure timely access to that data and resumption of services.

DR plans based on manual steps or stagnant assumptions are almost always going to fail. An automated approach is needed to start up the surviving data storage arrays and bring up the servers so that applications can, in turn, be restarted. EMC’s approach to achieving those goals, found in Disk Library for mainframe and other EMC storage offerings, centers on its Geographically Dispersed Disaster Restart (GDDR) functionality. GDDR starts with a foundation of assured data recoverability via SRDF/S, SRDF/A, and TimeFinder, then it builds up the awareness and relationships between the underlying storage and the servers and applications. This capability is key to both an operational DR plan and an audit to ensure DR capabilities such as OCC’s “sufficient to recover ... within the business day on which the disruption occurs” advice.

EMC purports that GDDR technology is used by the world’s most prominent financial sector companies. And, because this technology extends its management interface to the DLm8000, those customers now have available true disaster recovery and disaster restart, which not only covers the disk environment but also extends across tape images. That breadth of coverage improves the resiliency of the solution and lowers the recovery time objective because tape operations can be resumed quickly following a disaster declaration—the tape images and the system catalog information are kept permanently synchronized.
Understanding BC/DR in a Regulatory Context

The difference between “operational recoverability” and the ability to pass a “regulatory audit” cannot be underestimated.

- **Regulatory compliance** is based on auditors’ subjective assessments of a range of technologies and how they apply to business practices mandated by committees. Case in point, the primary mandate of the SEC prior to the post-9/11 guidance with the OCC was SEC Rule 17A-3 and SEC Rule 17A-4, which dealt with the preservation and retention of information. When folks talk about data retention, many lament as if it’s a new phenomenon. Those SEC rules apply because the SEC considers electronic information to be equivalent to paper records, which in many ways, it is. Rules 17A-3 and 17A-4 were first written in 1934 with paper records in mind, and the most recent interpretations were published in 2003.³

- **Operational recoverability** is based on an organization’s IT and business stakeholders collaborating on promoting a shared understanding of what their organization needs to continue doing business. This effort encompasses IT and non-IT policies and procedures that are based on the specifics of the given business but built upon contemporary technologies and methodologies that take into account the larger world that the company lives in.

Having great technology that enables operational recoverability may not pass an audit. Why? Because the auditor may not fully understand what all those racks of blinky lights do. So, having great operational recoverability in a heavily regulated industry such as finance requires well-articulated and demonstrable data protection, recovery, and resiliency capabilities that will be trusted and approved of by auditors and compliance officers alike.

Similarly, having data outside certain distance limits or for extended periods of time may not ensure that a company can recover their business services (only their data). Why? Because most regulations think about data universally, regardless of company size, regardless of whether or not they have distributed infrastructures, and without consideration of interdependent applications, systems, and teams. The key is to understand how to satisfy regulatory requirements while authoring, implementing, and sustaining “real” business continuity and disaster recovery processes.

Understanding Tape in Regulatory Context

For many, “tape” and “long-term retention” are synonymous. But a certain class of organization moves data at a higher volume and at higher speeds than what most tape solutions can sustain—a situation evident in mainframe environments. In those environments, “tape” isn’t just about backup; tape is the interoperable method of managing logical containers of production data, data iterations, and archive data tiers in large volumes. All of the production data needs to be protected and ensured to be available. But unless you’re physically shipping large containers of tapes with mind-numbing frequency, that goal may not be possible with traditional tape.

To address the need, especially in a highly regulated industry in which data literally has the financial value of the money “held within its bits,” disk solutions that appear to their hosts to be tape are necessary and, in fact, often preferable. To their hosts, they appear to be very fast and scalable tape solutions. To their administrators, they have the flexibility, resiliency, and manageability characteristics of the storage systems IT professionals are used to managing.

Combining these realities will rightfully point organizations to solutions such as EMC Disk Library for mainframe, a family of systems that offer different solutions for different areas of focus. Backup-intensive organizations may find it appealing that the DLm1000/DLm6000 leverage Data Domain. To support mixed workloads of batch, archive, space management, and backup, EMC offers the DLm2000/DLm6000. And for ultra-resiliency, there’s the DLm8000.

---

The Bigger Truth

Every organization has its own responsibility to plan how its IT and non-IT processes will ensure that the company is resilient. Some organizations, especially those affecting the public “common good” or, in this case, the global economy, have additional regulations that attempt to ensure their organizations’ resiliency on behalf of all of us.

Regulations on their own are not bad—as long as you understand that compliance alone is not enough. No CIO ever saved his job following an unrecoverable disaster just by showing his board of directors a “passed” audit certificate. Along with regulatory compliance, an organization needs a real BC/DR plan that includes IT and non-IT processes and policies that understand the business itself, as well as how to ensure that the business continues.

In considering the vulnerability of the U.S. financial system in the days after 9/11, the key regulatory bodies of the SEC, the Treasury, and the OCC gathered business and technology experts to create mandates to strengthen the infrastructure the U.S. and its global partners rely on to move money around the world properly. The mandates include suggestions tied to recoverability times, survivable data distances, etc., but it is in many ways up to individual financial institutions to choose the technologies that they will use to satisfy those mandates and demonstrate their resiliency capabilities—to auditors and to the company’s management alike.

For many of those financial systems, IT still relies heavily on mainframe systems because of their processing power. Those mainframe systems utilize tape, not so much as a media type but as a logical container for production data, archival data, and data being protected as part of the organization’s effort to adhere to regulations and achieve overall “ultra-resiliency” for the data. For those organizations, the volume (capacity) and demand (speed) of required data leads those IT teams to disk libraries that appear to be tape to their host systems, while providing the scale, resiliency, performance, and manageability that enterprise disk storage provides.

In considering enterprise disk storage that offers scale, resiliency, and manageability, it should come as no surprise to see the Disk Library for mainframe offerings taking advantage of the range of EMC storage platforms from VMAX to VNX to Data Domain in its mainframe storage solutions. In recognition of the brand and experience that EMC brings to the table, auditors, compliance officers, and general IT administrators should have friendlier conversations relating to data preservation and assurance for even the most daunting of regulated industries—and for mandates such as those recommendations by the OCC.