

WHITE PAPER

Delivering Tomorrow's Backup and Recovery Infrastructure

Sponsored by: EMC

Eric Sheppard

January 2012

EXECUTIVE SUMMARY

The data protection market has changed considerably over the past decade. During this time, the market witnessed a fundamental shift away from relying solely on tape for backup and recovery to using disk-based backup solutions to address challenges including backup performance, reliability, and recovery time objectives. The early days of this market shift saw a heavy reliance on tape emulation to ensure disk storage systems could be introduced into an organization's data protection scheme without the need to replace or disrupt the long-standing backup policies and schemes. Tape emulation served to introduce disk as a backup target and continues to be used in many enterprise environments. However, with the availability of disk-based backup solutions with deduplication and remote replication, the use of disk as the primary backup target has become more prevalent.

The proliferation of virtualized infrastructure, the growth in remote and branch office data, the changing roles of storage and backup administration, and the scale of today's data sets in parallel with a need for high availability have introduced a new set of requirements for the software responsible for data protection and recovery. As a result, the next generation of data protection and recovery software functionality continues to evolve. While functional needs have increased and more firms seek to use disk for operational and disaster recovery, the role of tape has also changed. Tape, once the de facto standard for backup, is now being used for less frequent monthly, quarterly, or annual backups as well as for longer-term backup-driven archive purposes.

This technology and infrastructure evolution, coupled with the challenges of legacy backup solutions, is driving firms to rearchitect their current backup approaches. To accelerate the evolution from legacy to next-generation protection approaches, technologies must be more tightly integrated than ever before. The supplier-led integration of these disparate technologies can address the limitations of previous backup approaches while also offering simplified implementation, integrated management, resource optimization, and faster recovery.

This paper highlights that firms evaluating next-generation data protection solutions must expect a greater degree of integration between the technology components in today's data protection path. Firms that have made an investment in tape infrastructure may want to continue to leverage it, but for different needs. While the use of tape will likely continue to evolve, firms need a fast, efficient, and cost-effective approach for operational and disaster recovery. In addition, today's data volumes,

virtualized infrastructure, and changing management models require software that provides functionality to address these environments while supporting both disk and tape as a backup target. The value of this integration is around extensibility and investment protection while offering a paradigm shift in the way data is protected and recovered, enabling an IT organization to meet today's performance, reliability, and recovery challenges while also supporting a fundamental shift in how and by whom data protection and restoration functions are performed in the future.

SITUATION OVERVIEW

Close to a decade ago, organizations began to see data sets reach volumes that were too big to be successfully backed up by the existing LAN and tape technology. Further exacerbating the problem was the fact that global operations meant backup windows were actually shrinking to very short time frames. Not surprisingly, organizations began to turn to gigabit Ethernet, multigigabit Fibre Channel storage area networks, and disk-based backup targets to ensure their data protection environments were able to cope with the modern challenges of that time.

While such changes introduced new technologies into the data protection scheme, the process of data protection remained largely unchanged. That is to say, backup software was still backing up production data from a physical server over a network (albeit now at faster speeds) to a physical target whose sole purpose was to very simply accept and store the copies of production data. These new technologies introduced higher costs and more complexity into the data protection process.

Virtual tape libraries (VTLs) were the first sign of a then young purpose-built backup appliance (PBBA) market. These devices were primarily installed to replace tape-based backup by emulating tape and thus not requiring IT organizations to also replace their long-standing backup applications. This was the first sustained trend of disk-based appliances becoming the primary backup target in place of tape.

While a broad move toward disk-based data protection has been helpful, backup and recovery challenges remain. Indeed, growth of newly created data has actually accelerated, while the reasons for creating copies have multiplied. Business analytics, application resiliencies, compliance with government regulation of digital assets, and increased corporate reliance on distributed operations (e.g., remote offices/branch offices [ROBOs]) are now placing tremendous pressure on IT personnel tasked with managing traditional backup and recovery schemes within an environment of accelerating data growth. When the relatively new complexities brought on by the rapid adoption of virtual infrastructure (VI) are added to the mix, one can quickly see why organizations of all sizes are working to update their backup and recovery architectures.

The data protection market has since progressed even further. Today, the need to reduce the complexity of the backup and recovery process remains a top priority, but an increasing number of customers are also demanding that disk-based data protection solutions address poor utilization rates through storage optimization (e.g., data deduplication) while requiring that new solutions extend protection to ROBOs, virtual environments, and endpoints. Simply stated, today's data protection

offerings must go beyond traditional backup software to include advanced efficiency features while also enabling customers to stage their backup modernization progression in a way that targets their most pressing needs.

Looking forward, market stakeholders and observers should expect a greater degree of integration between the individual components of many organizations' data protection and recovery strategies. That said, the modernization of backup and recovery won't happen all at once. Organizations will continue to search for a set of offerings that are flexible enough to address targeted backup and recovery needs while also ensuring that additional elements of their infrastructure can be accommodated at a later date.

DATA PROTECTION AND RECOVERY: THE VALUE PROPOSITION OF TIGHT INTEGRATION

Within most datacenters today, disk storage, backup and recovery software, and tape constitute the fundamental technology of a firm's data protection and recovery strategy. However, the challenges with legacy backup approaches in concert with accelerated requirements to increase data protection performance, reliability, and recovery needs are driving firms to rearchitect their current backup approaches.

Backup and recovery technology, which has largely been developed independently, must be more tightly integrated. The supplier-led integration of these disparate technologies can address the limitations of previous backup approaches and provide a foundational platform for the future. The benefits a firm will realize in deploying a data protection solution that has integrated these once disparate technology domains include:

- ☒ **Performance optimization.** Integration of file format-aware functions such as replication and/or backup instrumented through host-side software with target-side storage efficiency algorithms can drive backup performance gains by reducing the amount of data scanned, indexed, or moved. Integrated solutions must provide greater optimization of functions provided by any component in the larger data protection system.
- ☒ **Ease of implementation.** Backup configurations are complex, involving integration with the application, the network, and the storage. Tight integration and delivery of software and system solutions that have been packaged as standard configurations can reduce this complexity of setup and reduce implementation time.
- ☒ **Function optimization.** Increasingly, deduplication is a must-have function in backup environments. An integrated solution that offers and is aware of both client-side and target-side deduplication optimization routines, as well as content type, can be tuned to perform this function at the right level of the infrastructure to optimize performance with available resources and deduplication ratios with content type.

- ☒ **Improved recovery.** Restore and recovery needs vary. An integrated solution is well suited to respond to a range of recovery requests, from file, to image, to systemwide or sitewide recovery. Additionally, integrated systems can best respond to a combination of replication and disaster recovery policies. An integrated solution must provide fast recovery from both localized and sitewide issues, with recovery options aligned with application priority.
- ☒ **Support for new backup management models.** Backup was once performed by dedicated administrators, and tape was managed by operators. Virtualization and the use of disk as the preferred recovery mechanism are fueling the use of backup and recovery functions by VI administrators, application owners, and cloud architects. An integrated solution must offer a single management schema whereby role-based functions can be performed, leveraging components of the data protection stack.
- ☒ **Extensibility and investment protection.** Changes in data protection have not happened and will not happen overnight. Tape still has a role in today's datacenters, and today's cost-sensitive storage organizations must have a strategy that allows them to extend the value of existing investments. An integrated solution approach must address and support disk, tape, and software integration.
- ☒ **Ability to offload overhead from applications.** Software functions that run on application servers themselves can often introduce overhead on these systems, consuming valuable CPU and memory cycles that might otherwise have been allocated to the application itself. Offloading specific functions from running on the application server to a separate and dedicated system (or to the storage layer) can have operational benefits and avoid unnecessary performance degradation.
- ☒ **Reduced maintenance.** Standalone backup software typically has underlying dependencies, and vendors release support matrices that identify what versions of which software have been tested with specific operating systems, patches, microcode levels, and so on. When end users seek to upgrade a given piece of software, intentions can be delayed due to incompatibility issues. Maintaining a "supported" environment takes valuable cycles that can be minimized with the deployment of an appliance. An integrated deployment is a tested, simplified, turnkey approach that will not have the pitfalls a standalone set of software and hardware would have.

EMC'S BACKUP AND RECOVERY PORTFOLIO: AN EXAMPLE OF TIGHT INTEGRATION

Over the past 10 years, EMC has made significant investments in the backup and recovery market. In this time and through a combination of both organic and inorganic growth, the company has become a formidable and material player in the data protection and recovery market. Today, EMC's backup and recovery portfolio is one of the strongest in the market.

EMC Backup and Recovery Solution Success

EMC realized that its customers needed an integrated, disk-based solution to free themselves from the constraints of their legacy backup and recovery processes. As such, it began to invest in appliance-based backup and recovery technology to tackle the problems of data deduplication, ROBO backup, and virtual machine (VM) protection.

EMC's investments in the deduplication software and storage market through its acquisitions of Avamar and Data Domain, coupled with its acquisitions of Legato, WysDM, and Bus-Tech, have given EMC a backup and recovery portfolio that is expansive and flexible enough to meet the diverse needs of the mainframe and open systems data protection market. EMC has made continued investments within its backup and recovery business to ensure its product portfolio is broad enough to span an exhaustive list of use cases, yet interoperable enough to scale effectively and efficiently.

EMC Backup and Recovery Technology Integration

What makes the EMC portfolio particularly strong is not just the series of products included but also the integration of the products, offering considerable gains in backup performance, optimization, management, and recovery.

EMC has integrated many of its core data protection technologies to offer customers value above and beyond the value of each of the solutions loosely integrated. EMC's portfolio integration offers advantages in backup performance, storage optimization, centralized management, and improved recovery. Examples include:

- ☒ **NetWorker and Avamar.** Over four years ago, EMC integrated Avamar's client-side deduplication with NetWorker. This integration allows NetWorker customers to benefit from deduplication or traditional backups, or both. This functionality is available through a single, integrated NetWorker client and software deployment and management model.
- ☒ **Avamar Data Store.** EMC's Avamar Data Store tightly integrates Avamar deduplication software with servers to offer a backup appliance for protection of ROBOs, desktops/laptops, NAS servers, and VMware environments.
- ☒ **NetWorker and Data Domain.** EMC has made great strides in integrating NetWorker with Data Domain appliances to enable customers to back up to deduplication-enabled systems as well as tape. NetWorker integrates with Data Domain Boost. With Boost, parts of the deduplication process are distributed to the NetWorker storage node or application clients, enabling each to send only unique data segments to the Data Domain target. This increases network throughput, reduces the amount of data transferred, and decreases CPU utilization on NetWorker components. Additionally, NetWorker can be used to manage setup, alerting, reporting, and replication with and between Data Domain systems.

- ☒ **Avamar and Data Domain.** More recently, EMC integrated Avamar with its Data Domain Boost software for application-specific backup, such as Oracle, SQL, Exchange, SharePoint, or VMware images. Because of the Data Domain Boost integration, the Avamar clients can send unique data segments directly to the Data Domain system. This allows the company to provide its customers with greater scope and scale for backup and data reduction while also continuing to minimize backup windows and network traffic. Additionally, Avamar can be used to provide a single point of management among Data Domain and Avamar appliances.
- ☒ **Disk Library for mainframe and Data Domain.** The DLm6000 is a unique virtual tape library solution in that it supports both primary and backup use cases. Recognizing the tremendous value of deduplication for mainframe backup, EMC integrated Data Domain storage with the DLm6000, resulting in increased backup application performance, more efficient replication, and more cost-effective onsite retention of backup data.
- ☒ **Data Protection Advisor.** All EMC backup and recovery products are integrated with EMC Data Protection Advisor to help manage service levels, reduce complexity, and eliminate manual efforts. Data Protection Advisor provides data protection management to automate, monitor, analyze, alert, and report across backup, replication, and virtual environments.

Showing considerable foresight, EMC has amassed a set of backup and recovery offerings that bring significant customer value through flexible, integrated backup and recovery solutions that are able to scale to meet its customers' future needs. During this time, the company has also emerged as a leader in the PBBA market. Generally speaking, purpose-built backup appliances are disk-based storage solutions built specifically to serve as a target for backups and replicated backups. These solutions include functions such as deduplication, compression, encryption, and remote replication. Over time, IDC expects to see these systems expand to serve as a target for other secondary storage workloads.

While EMC is not the only PBBA supplier in the market, many competitors have been slow to admit the strong customer value that PBBA solutions provide. This has left EMC ample time to effectively demonstrate a market need to go beyond a single or pure-play set of products to effectively manage and protect ever-growing data requirements. In some cases, the PBBA model could prove to be very disruptive to business models of any competitors that do not possess the core competencies needed to develop and maintain a hybrid hardware and software solution. Not surprisingly, many other suppliers now view the PBBA market as key and strategic to their long-term success.

Today, EMC's backup and recovery solutions provide measurable customer benefits:

- ☒ **Reduced backup volumes.** Customers' backup volumes are decreased significantly using Avamar in concert with Data Domain systems.
- ☒ **Enhanced backup performance and fidelity.** Typically, customers achieve a backup success rate of less than 100%. After an EMC deployment, backup failures are reduced to nearly zero.

- ☒ **Improved network performance.** Network resources are significantly improved using data deduplication with Avamar and Data Domain solutions.
- ☒ **Improved recovery performance.** Restoration of data takes a matter of minutes as opposed to hours, days, or weeks when relying on traditional tape-based methodologies.
- ☒ **Improved administration.** Significant cost reductions are realized with backup administrator resources redeployed to other areas.
- ☒ **Measurable cost savings.** Significant capex/opex improvements are realized, and tape media consumption is reduced to zero. In the long term, more cost savings are realized with the elimination of tape infrastructure and media vaulting.
- ☒ **Centralized administration and maintenance.** Customers can take advantage of key tools such as Data Protection Advisor to manage and automate data protection processes.
- ☒ **Protection of physical and virtual infrastructure.** EMC's backup and recovery solutions provide customers with protection of physical and virtual servers as well as consolidated management.
- ☒ **Improved security and reliability.** Customers do not have to rely on the tape-based model of supporting offsite repositories of tape, which reduces the threat of possible loss or compromise of physical tape.

CHALLENGES

Increasingly, IT solutions in the datacenter are being delivered in integrated form factors for many foundational IT services and functions. The most significant new integrated solutions are optimized for specific workloads such as data protection.

The growing interest in these new integrated data protection solutions is having a significant impact on the storage market. A decade ago, numerous ISVs and storage software pure-plays were participating in the data protection market. Today, the backup markets at large have only a handful of major material players. Natural market behaviors of consolidation, maturation, and convergence contributed to this trend; however, what was once a broad range of distinct standalone software is being replaced by bundled data management products, often provided in the form of an integrated solution. IDC expects even tighter integration of data protection functions within integrated solutions in the coming year.

We believe the biggest challenge for EMC in expanding its presence in the backup and recovery market will be displacing legacy products, practices, and vendors. The data protection market is notoriously slow to change due to high investment costs and long product life cycles. This creates inertia among customers and limited desire to change incumbent backup and recovery infrastructure. But here is where EMC's broad portfolio can help. EMC has done an exemplary job of providing its customers with integrated PBBA solutions to address targeted backup and recovery pain points rather than try to force a far-reaching data protection change. Still, EMC needs to pursue a broader group of customers outside its own installed base to mitigate

against apathy or inertia. This is critical to EMC's continued dominance in the PBBA market. Also, EMC needs to educate potential customers about its sizable investments and integration efforts now and in the future.

CONCLUSION

EMC has made judicious investments and acquisitions to build out key features for customers dealing with burgeoning storage and management requirements for backup and recovery. EMC's approach has been to offer its customers a robust and well-integrated set of hardware and software solutions to meet their data protection needs, be they targeted or far reaching. Furthermore, IDC believes that EMC offers its customers a strong value proposition by allowing them to modernize all facets of their backup and recovery architecture, but in a fashion that suits their budgets and established processes.

Copyright Notice

External Publication of IDC Information and Data — Any IDC information that is to be used in advertising, press releases, or promotional materials requires prior written approval from the appropriate IDC Vice President or Country Manager. A draft of the proposed document should accompany any such request. IDC reserves the right to deny approval of external usage for any reason.

Copyright 2012 IDC. Reproduction without written permission is completely forbidden.