

## INSIGHT

---

### Network Appliance Enables Role-Driven Data Protection with Protection Manager Product

---

Laura DuBois

---

#### IDC OPINION

---

Increasingly, disk storage is being placed in the data protection path in open systems environments, either as an interim resting place for data en route to tape or as a final destination. Firms are using disk storage as a complement to traditional tape processes to speed operational recovery, increase reliability, and improve backup performance for increasing volumes of data requiring protection. A myriad of different data movement approaches are in place today to copy data to disk storage — including backup to disk, virtual tape to array, and network- and host-based storage replication. Storage replication (including snapshots, mirrors, and replicas) continues to be a proven yet fast-growing approach to creating copies of primary data available for recovery in the event of a logical or physical failure. As firms create more frequent replicas, across more and more locations, systems, and applications, the following trends surface:

- ☒ **Replication growth is driving a need for improved management.** The use of replication solutions for data protection and business processes is increasing. It is being used across centralized datacenters as well as IT staff-constrained, distributed branch locations. Moreover, the number of replica copies is growing to provide more granular recovery points. This growth has driven a need for improved replication management and policy-driven automation.
- ☒ **Replication management takes on data management functions.** As replicas and snapshots continue to proliferate, a richer set of policies, similar to those found in traditional data management products, around managing replicas must be available. For snapshots or replica data protection copies that terminate on disk, a series of policies can be applied to control the retention, expiration, scheduling, numbers, locations, and methods of creation of these data protection or replica copies.
- ☒ **Role-driven storage and protection facilitates self-service capabilities.** Through application API integration, the controls for the replication processes are expanding beyond the storage domain to include the application administrator domain. Allowing application owners to control the frequency of snaps and the schedule of snaps, copies, and restores gives them greater flexibility, reduced administration time, and improved productivity. The application administrators know the application and are therefore closer to its policies and information.

## IN THIS INSIGHT

This IDC Insight looks at the replication market and Network Appliance's (NetApp's) participation in this market in the context of protection and recovery. It goes on to examine the need for improved levels of policy-driven management of replicas and analyzes the Network Appliance Protection Manager product in the context of user demand, changing storage usage models, and underlying storage replication products.

## SITUATION OVERVIEW

Storage replication has seen double-digit revenue growth over the past five years, in part as a result of the demand for data protection and recovery from logical or physical errors as well as supporting business processes such as test/development, data warehousing, and decision support. In the context of data protection, storage replication solutions have offered enabling technology for business continuity and disaster recovery. Different from other data protection approaches such as traditional policy-driven backups, replication formats have been native to the source data getting replicated and focused on making the replication or restore process efficient although the policies to control replication frequency, granularity, and pre- and postprocesses have been largely manual or scripted. However, this is changing. Improved levels of replication management are becoming available to offer policy controls and automation of protection tasks. These policies, traditionally found in backup applications, are moving down to the replication layer as a result of the following dynamics:

- ☒ **Distributed data growth.** Data growth outside larger datacenters in satellite and branch remote has fueled the need for replication out in distributed enterprise locations. However, limited technical staff at remote and branch locations has driven a need for centralized management of remote replication and recovery processes.
- ☒ **Cross-system replication.** The growth in the use of replication solutions for data protection and business processes across homogeneous and heterogeneous systems has driven a need for more centralized replication policy management and automation.
- ☒ **Limited storage management resources.** The 52% growth rate year over year in primary storage is at odds with available storage administrators to manage this increasing capacity. Limited administrative time to develop and update scripts or perform manual process is driving an increased need for storage provisioning and replication automation.
- ☒ **Compressing recovery requirements.** The need for more frequent recovery points to minimize data loss as well as fast recovery in the event of failure is driving the creation of snapshots as frequently as hourly or, in some cases, up to every 15 minutes. Growth in the number of snapshots requires some automated means of controlling the replica creation, usage, and expiration processes.

- ☒ **Role-driven management.** Most large firms are trying to align business objectives and policies with IT implementations. The closer the administrator is to the information that lives within an application, the more likely the administrator is to understand the business use of the information and its business policies. In many cases, the application or database administrator may be in the best position to know when a replica needs to be created or when a restore should be initiated.

---

## **Network Appliance Major Driver in Replication Growth**

Network Appliance's performance in storage replication has been a major driver in the growth of the overall storage replication software market. Network Appliance's storage replication revenue grew from \$349 million in 2005 to \$525 million in 2006, representing a 50% growth rate and well exceeding overall market growth rates. With this growth, the company holds the number 2 market share position and gained five market share points from 2005 to 2006. Preliminary estimates for 2007 performance suggest the company is poised for another year of strong double-digit growth with its replication offerings.

Network Appliance offers three primary options — SnapMirror, SnapVault, and ReplicatorX — for replication-based data protection:

- ☒ SnapMirror is replication software intended for disaster recovery solutions. The mirror is an exact replica of data on the primary storage that can be mounted read/write to recover from failure. If a backup is deleted on the source, it will go away on the mirror at the next replication.
- ☒ SnapVault, in contrast, is replication software intended for disk-to-disk backup. It retains all backup copies as they appeared at the time they were created on primary storage for a user-specified period of time. Secondary storage used by SnapVault cannot be mounted read/write. Backups must be recovered from secondary storage to the original or an alternative primary storage system to restart.
- ☒ ReplicatorX is data replication software for heterogeneous storage environments. While SnapMirror and SnapVault work within the DataONTAP architecture and thus replicate data between or within NetApp systems, ReplicatorX provides network-based replication between dissimilar storage systems. (This product came from Network Appliance's acquisition of Topio in 2006.)

Note: Other replication offerings include SnapRestore, which provides data restoration for near-instantaneous restores; SyncMirror, which provides synchronous data replication for disaster recovery protection; LockVault, which supports backups for regulatory compliance and archival storage requirements; and FlexClone, which enables instant dataset clones with minimal storage overhead.

Factors that have contributed to the growth of Network Appliance's storage replication business include industry-leading attach rates of replication with storage system sales; the use of a single replication and storage operating system (DataONTAP)

schema across low-end to high-end NetApp platforms; and cost-effective, easy-to-deploy remote one-to-many replication for distributed branch office locations. Network Appliance also differentiates its replication offerings based on cloning technologies for DR testing purposes that clone without any space utilization or performance penalties, thin provisioning technology, and deduplication to support network-efficient replication.

NetApp replication product attach rates differ based on vertical, application, customer environment, and product, but approximately 50% of NetApp's storage platforms ship with one or more of its replication software products. Consistent with an overall market trend, the higher-end storage systems typically have higher attach rates. Network Appliance reports that SnapMirror is its most popular shipping replication software, followed by SnapVault and SnapRestore. Disaster recovery, business continuity, disk-based backup, test and development, and quality assurance are the most prevalent use cases for data replication, with database (predominantly Oracle), mission-critical business applications, and file server consolidation also driving data replication software sales. Increasingly, users want to leverage NetApp snapshots to serve multiple use cases as a means of minimizing storage capacity consumption and costs.

As the replication market continues to grow, Network Appliance has recognized the need for increased levels of automation in the management of distributed replication processes — in particular, as replication continues to increase in use for both protection and recovery. This increased level of workflow and management automation translates to policy creation, monitoring and enforcement, and the support for role-driven data protection and replication. As a result, the company has developed and released a product called Network Appliance Protection Manager. Protection Manager is designed to address three fundamental challenges:

- ☒ **Complexity.** Protection Manager provides a level of abstraction that masks the complexity of configuring and managing underlying storage tasks such as provisioning, snapshot creation, and physical data movement. Protection Manager creates an abstraction layer to underlying NetApp software that allows server and storage admin to think about protection in terms of creating backups and mirrors.
- ☒ **Manual processes.** Protection Manager frees the storage administrator from manually (or in a scripted fashion) tracking, monitoring, and ensuring mirroring relationships are maintained for thousands of LUNs and volumes, across hundreds of NetApp systems. Instead, Protection Manager automates these tasks.
- ☒ **Vulnerabilities.** Protection tasks are critical to a business' ability to recover. Traditional protection processes are fraught with error, obviating a valid recovery. Protection Manager provides a centralized, policy-driven management of distributed, heterogeneous replicas and monitoring and tracking of replica policies and events to ensure compliance with business rules and isolation of protection vulnerabilities. Protection Manager will detect unprotected or orphan volumes and assign a protection policy to them.

---

## **Overview of Network Appliance Protection Manager**

Protection Manager is a software option for Network Appliance SnapMirror and SnapVault replication environments, enabling policy management and automated configuration of data protection. Released in February 2007, the product already has 200 customers running it in production. At its core, Protection Manager is a policy engine and interface that works across distributed and disparate SnapMirror and SnapVault replication environments to provide dataset and resource pool creation, policy configuration, and automated execution of configured Protection Manager policies. While it works across both SnapMirror and SnapVault, the attach rate of Protection Manager with SnapVault has been higher.

---

## **Protection Manager Architecture**

The Protection Manager software acts as a client to the Operations Manager (Data Fabric Manager [DFM] Server). The Protection Manager client can run on an existing or new Operations Manager server running a Windows, Solaris, or Linux operating system. The Protection Manager client ships with an embedded Sybase database that is used for storing Protection Manager configuration and log data. A single Protection Manager instance can manage up to 250 Network Appliance local or WAN-connected storage systems. The Protection Manager interface is a Web-based Java UI that works with all standard Web browsers. The Protection Manager Web-based UI provides an abstraction to underlying storage tasks while offering a unified view into NetApp SnapMirror and SnapVault protection environments. Administrators use the Protection Manager interface not only to create datasets and policies but also to improve manageability by tracking protection events, status, unprotected or orphan data, and configured datasets and resource pools.

Central to the operation of Protection Manager are three concepts: policies, datasets, and resource groups, which are described in the sections that follow.

### ***Protection Manager Datasets and Resource Pools***

Protection Manager enables creation of datasets. Datasets are logical containers that visually represent data that is mapped to physical storage. For example, a dataset might be created to hold backup data or replicated data. Physical storage is provisioned from available resource pools. A resource pool is analogous to a destination tape pool or other aggregated storage entity that can be shared. As part of the Protection Manager setup process, an administrator designates which storage resource pools to use to hold remote replicas of the dataset. Distributed secondary storage resources can be unified and consolidated into a resource pool. A unified resource pool can then be allocated to several datasets. Protection Manager will provision destination volumes and qtrees from those resource pools to receive backup copies of the dataset. Protection Manager then creates SnapVault and/or SnapMirror relationships from each volume and qtree in the dataset to the newly provisioned destination storage in the resource pool.

### ***Protection Manager Policies***

Protection Manager allows for the creation of policies for data protection in a wizard-driven graphical process and then calls SnapMirror or SnapVault for the execution of the replication process. A policy is a rule that describes how to protect the data. A policy might indicate "make copies every week and keep them for at least a year" or "retain undeletable copies for seven years." Administrators define the policy within Protection Manager, which then automates the execution of the policy. Administrators can apply a policy to a single volume or LUN, or to a user-defined group called a dataset. For a large number of LUNs that all support the same application, administrators can group them together in a dataset and apply the policy to the dataset as a whole. The product comes with nine canned or predefined policies, although administrators can also create their own. Protection Manager policy configuration and execution allows for the definition of:

- ☒ Frequency of replica creation
- ☒ Scheduling of replication tasks
- ☒ Numbers of replicas to retain and for how long
- ☒ Storage location on which to retain replicas
- ☒ Replication method (backup snapshot or mirror)

Protection Manager provides an abstraction layer for underlying Network Appliance software functions for execution of tasks such as storage provisioning, snapshot creation, and physical data movement and allows users to think about protection in terms of creating backups and mirrors. Protection Manager frees the storage administrator from tracking, monitoring and ensuring mirroring relationships are maintained for thousands of LUNs and volumes, across hundreds of NetApp systems. Instead, management is done by a smaller number of Protection Manager–created datasets, each with the appropriate policy.

### ***Manageability, Monitoring, and Reporting***

Additionally, Protection Manager will detect new volumes (as long as they are on the same subnet) that are currently not protected, thus ensuring no orphan data exists. Once these volumes are detected, they are presented as "unprotected data" in the Protection Manager UI. Administrators can then protect these volumes according to predefined policies. Additionally, Protection Manager monitors the whole replication process, watching the capacity and performance against policy, and ensures that protection policies are not out of compliance and recovery is in a state of vulnerability.

## FUTURE OUTLOOK

As replicas continue to proliferate, and an average firm creates anywhere from three to nine copies of its data across different systems, locations, and datacenters, increasingly, it will be the management of and tracking for these data copies that will be paramount. With information running today's businesses against a landscape of legal and competitive pressures, firms need to ensure their systems and data are managed according to policy. NetApp Protection Manager addresses these market dynamics. Going forward, NetApp should consider the following:

- ☒ Incorporating support for a broader set of NetApp replication products, including ReplicatorX, to support a heterogeneous replication environment
- ☒ Improving distribution of Protection Manager with the company's network of partners (resellers, distributors, and system integrators) that sell storage hardware products
- ☒ Integrating Protection Manager with the recently acquired Onaro portfolio of products, including integration with Onaro's VMware Insight Manager
- ☒ Working with standards bodies such as SNIA to continue to develop industry-standard interfaces for the management of heterogeneous replication products such as those from other storage suppliers to expand the value of the Protection Manager product
- ☒ Integrating Protection Manager with server, application, and virtualization admin interfaces, which will allow application owners to control the frequency of snaps and the schedule of snaps, copies, and restores and give them greater flexibility, reduced administration time, and improved productivity (The application administrators know the application and are therefore closer to its policies and information.)

---

## Copyright Notice

This IDC research document was published as part of an IDC continuous intelligence service, providing written research, analyst interactions, telebriefings, and conferences. Visit [www.idc.com](http://www.idc.com) to learn more about IDC subscription and consulting services. To view a list of IDC offices worldwide, visit [www.idc.com/offices](http://www.idc.com/offices). Please contact the IDC Hotline at 800.343.4952, ext. 7988 (or +1.508.988.7988) or [sales@idc.com](mailto:sales@idc.com) for information on applying the price of this document toward the purchase of an IDC service or for information on additional copies or Web rights.

Copyright 2008 IDC. Reproduction is forbidden unless authorized. All rights reserved.