

INSIGHT

NetApp Enables Role-Driven Data Protection with Protection Manager

Laura DuBois

IDC OPINION

Firms are using disk storage as a complement to or outright replacement of traditional tape processes to speed operational recovery, increase reliability, and improve backup performance for increasing volumes of data requiring protection. Storage replication (including snapshots, mirrors, and replicas) continues to be a proven approach to creating copies of primary data, which are 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 replication management.** The use of replication for data protection and business processes is increasing. Replication 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 even 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 facilitate self-service capabilities.** Through application integration, the controls for the replication processes are expanding beyond the storage domain. Allowing application administrators to control the frequency and scheduling of snaps, copies, and restores gives them greater flexibility, reduced administration time, and improved productivity. It also builds self-sufficiency. Application administrators know the application and are therefore closer to business needs, various IT policies, and the information itself.
- ☒ **Data and storage management are being integrated.** The old model of application or database owners calling on storage teams for capacity or data management services does not scale. Some application or virtual machine administrators are starting to fulfill basic self-service tasks like creating a backup before an upgrade or initiating an out-of-place restore for test and development. As a result, controls provided in the storage layer and leveraged by storage administrators may be released to data owners themselves. However, this is a balancing act between controlling the process while relinquishing the management.

IN THIS INSIGHT

This IDC Insight looks at the replication market and NetApp's participation in it 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 NetApp Protection Manager product in the context of user demand, changing storage usage models, and underlying storage replication products. This Insight also includes highlights from the latest release of the Protection Manager product.

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. They have been 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, which are traditionally found in backup applications, are moving to the replication layer as a result of the following dynamics:

- ☒ **Distributed data growth.** Data growth outside larger datacenters in satellite and branch remote offices 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 the need for more centralized replication policy management and automation as a means of maintaining control and consistency. Consumers are demanding that tools work with multiple brands of storage.
- ☒ **Limited storage management resources.** The 52%+ growth rate year over year in storage capacity is at odds with storage administrators available to manage this increasing capacity. Limited administrative time to develop and update scripts, or perform manual processes 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 ensure 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 administrators are to the information that lives within an application, the more likely they will 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. Giving these administrators certain authorized roles leads to scheduling freedom as well as independence and self-sufficiency while still retaining overall IT responsibility.

NetApp as a Major Driver in Replication Growth

NetApp's performance in storage replication has been a major driver in the growth of the overall storage replication software market. NetApp offers three primary options for replication-based data protection: SnapMirror, SnapVault, and Open Systems SnapVault (OSSV):

- ☒ SnapMirror is an asynchronous replication software intended for disaster recovery solutions. The mirror is an exact one-for-one replica of data on the primary storage that can be mounted read/write to recover from failure. If data is deleted on the source, it will go away on the mirror at the next replication.
- ☒ SnapVault, in contrast, is replication function intended for disk-to-disk backup. It enables block-level incrementals — copying only the data blocks that have changed since the last backup, not entire files. Unlike SnapMirror, 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 in order to restart.
- ☒ Open Systems SnapVault leverages the block-level incremental backup technology in SnapVault to protect Windows, Linux, and VMware systems running on mixed storage.

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

The factors that have contributed to the growth of NetApp'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. NetApp 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.

Overview of NetApp Protection Manager

As the replication market continues to grow, NetApp has correctly recognized the need for increased levels of automation in the management of distributed replication processes. 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 NetApp Protection Manager.

NetApp Protection Manager is a software option for NetApp SnapMirror, SnapVault, and Open Systems SnapVault environments, enabling policy management and automated configuration of data protection and disaster recovery. The product has close to 1,000 unique customers, most of which have several copies of the software installed and running in production.

At its core, Protection Manager is a policy engine and interface that works across distributed and disparate SnapMirror, SnapVault, and OSSV replication environments to provide data set and resource pool creation, policy configuration, and automated execution of configured protection and disaster recovery policies. Protection Manager is designed to address the following 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 administrators to think about protection in familiar 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.
- ☒ **Recovery risk.** Protection Manager not only manages existing protection policies but can also automatically discover unprotected or orphan data and apply a protection policy to it and avoid recovery risk. Integration with Provisioning Manager enables provisioning of secondary storage, which can be deduplication enabled.
- ☒ **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.

- ☒ **Lack of visibility.** Protection environments often include many systems, tools, people, and backup methods. NetApp Protection Manager consolidates into a single framework, a holistic view of a NetApp disk-based data protection environment. The user interface (UI) for Protection Manager provides mapping to hosts and volumes and detects existing data protection relationships for import, making it easy to deploy in existing environments.

Protection Manager Feature/Functions

Protection Manager provides an abstraction layer for underlying NetApp software functions for execution of tasks such as storage provisioning, snapshot creation, and physical data movement and allows the user 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 data sets, each with the appropriate policy. Protection Manager provides the following:

- ☒ **Discovery.** Detects new volumes not protected and presents as "unprotected data" in the Protection Manager UI, where predefined policy can pick them up for protection
- ☒ **Policy creation.** Creates policies for data protection in a wizard-driven graphical process and then calls lower level NetApp tools for execution of the replication process
- ☒ **Monitoring.** Monitors the whole replication process, watching the capacity and performance against policy, and ensures that protection policies are not out of compliance
- ☒ **Visualization.** Provides discovery and mapping views (The user interface [UI] progresses from top to bottom and from general to the specific. Drilldowns are particularly effective wherein the operator can click on elements for more information or for showing lower level members in the hierarchy.)
- ☒ **Reporting.** Offers status and health reporting such as a "data transfer report" to identify transfer amount, performance metrics, and duration of transfer for replication processes, eliminating the need to parse through disparate log files
- ☒ **SnapMirror, SnapVault, and OSSV support.** Supports all forms of NetApp replication
- ☒ **Virtual machine support.** Support through Open Systems SnapVault includes VMware ESX, Microsoft Hyper-V, and Citrix XEN
- ☒ **Application integration.** Integrates with SharePoint, SQL Server, Microsoft Exchange, Oracle, and SAP via NetApp SnapManager
- ☒ **Data ONTAP deduplication.** Interoperates with data ONTAP deduplication, which is not charged for separately, to reduce storage footprint and cost
- ☒ **DR task automation.** Automates tasks, leverages templates, and provides ongoing monitoring with subsequent reporting to those in authority, as well as leverages policy-driven actions rather than haphazard work lists

- ☒ **DR readiness.** Monitors resources for changes that could compromise a disaster recovery and proactively communicates them to administrators for remediation
- ☒ **One-button failover.** Provides continued data access to users, even in the event of a disaster, and configures SnapMirror relationships that support backup and failover processes and assignment export protocols to the DR data set so users can access data in the remote system using the same protocols used to access the original

Baylor College of Medicine Uses NetApp Protection Manager

Located in Houston, Texas, the Baylor College of Medicine (BCM) (www.bcm.com) is recognized as a premier academic health science center known for its excellence in education, research, and patient care. In clinical and research settings like the Baylor College of Medicine, IT systems tend to proliferate organically, with different groups making their own choices when it comes to servers and storage. Although there was a centralized system designed to back up the majority of information, there was also a distributed infrastructure with nearly 1,000 servers and data spread across distributed disks — or sitting on DVDs in filing cabinets — all outside the purview of the centralized backup systems. Given this infrastructure, Baylor IT faced a significant challenge in managing its data and ensuring complete backups — and even more difficulty restoring data on demand.

Baylor turned to NetApp to help centralize management, and perform consistent, standardized backups based on established policies to minimize data at risk. The solution includes NetApp V-Series controllers and FAS storage systems, NetApp Operations Manager, Protection Manager, FileStorage Resource Manager (FSRM) SnapMirror, and SnapVault software. Together, these products provide an integrated solution to address the management, backup, and restore challenges facing BCM.

Today, Baylor uses Operations Manager and Protection Manager software to:

- ☒ **Automate and standardize data protection and recovery.** With the introduction of Protection Manager, Baylor has standardized and automated its Snapshot backup and restore processes and minimized reliance on ad hoc scripts.
- ☒ **Automate procedures based on set policies.** Baylor manages hundreds of existing SnapVault and SnapMirror relationships using Protection Manager policies to automate backup and replication.
- ☒ **Safeguard data based on key strategic priorities.** Protection Manager allows administrators to group data with similar protection requirements into data sets and apply preset policies to simplify and standardize protection of data related to particular clinical, scientific, or research departments.
- ☒ **Manage and monitor replication and restore tasks.** Protection Manager enables the organization to manage, monitor, and receive alerts on all SnapVault and SnapMirror replications and to manage restore tasks across eight NetApp systems with over 198TB of SnapVault data.

- ☒ **Enable storage as a service (StaaS) and align with growing organizational needs.** BCM uses Operations Manager to determine metrics such as resource utilization capacity consumption, and growth rates to determine chargebacks across departments and streamline annual budgeting.

According to the NAS and SAN Administrator at Baylor, "With Protection Manager, our IT team can examine a variety of metrics such as resource pool utilization, replication status, and unprotected data. It helps us maximize our storage resources and substantially reduces our administrative overhead."

Protection Manager Architecture

Central to the operation of Protection Manager are three concepts: data sets, policies, and resource pools:

- ☒ Data sets are collections of user data managed as a single unit, plus all the replicas of that data.
- ☒ Policies are sets of rules that specify the intended management of data set members.
- ☒ Resource pools are collections of unused physical storage (such as storage systems or aggregates) that new volumes or LUNs can be provisioned from to contain data. Once a storage system is assigned to a resource pool, all aggregates on that storage system become available for provisioning.

Protection Manager software acts as a client to NetApp Operations Manager software. Protection Manager can run on Windows, Unix, or Linux operating systems and has been qualified for use in various virtual environments such as VMware, Citrix, and Microsoft. Protection Manager configuration information and data such as logs, counters that determine growth rates, capacity, and alerting mechanisms are stored on the Protection Manager database server.

A single Protection Manager instance can manage up to 250 local or WAN-connected NetApp storage systems. The Protection Manager interface is Java based and works with all standard Web browsers. The Protection Manager console offers performance metrics and volume usage information. It also provides at-a-glance information on what is not protected, as well as the current status of backups. Administrators use the Protection Manager management console to not only create datasets and policies but also improve manageability by tracking protection events, status, unprotected or orphaned data and configured datasets and resource pools. The most important functionality of the Protection Manager dashboard for administrators is its ability to provide full transparency into the overall storage environment. In addition, Protection Manager provides a unified view into NetApp SnapMirror and SnapVault reporting to streamline data management.

The Protection Manager server component crawls the network on demand or according to a predefined schedule to discover new information. If a discovered node is a child node of a protected parent node, then the child node assumes the protection policy of the parent. For example, if a customer puts a whole storage appliance into a data set, then new volumes are automatically protected using the same policy.

Protection Manager Data Sets and Resource Pools

With Protection Manager, physical storage is provisioned from available physical resource pools that are contained within data sets. The resource pool and data set are part of the Protection Manager setup process. During the setup process, an administrator designates which storage resource pools to use to hold remote replicas of the data set. A unified resource pool can then be allocated to several data sets and have multiple policies affecting each data set individually. Protection Manager will provision destination volumes and qtrees from those resource pools to receive backup copies of the data set. Resource pools are thinly provisioned via NetApp's operating system, Data ONTAP. When consumption of those pools reaches user-specified thresholds, alerts can be sent via SNAP/email, PDAs, cell phones, BlackBerrys, and so forth as chosen by the user. Because users can select their thresholds, each warning mechanism can be considered site specific.

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 might indicate "make copies every week and keep them for at least a year" or "retain undeletable copies for seven years." It is also possible to dictate via these policies whether these Snapshot replication technologies replicate to a tertiary node or create local Snapshot copies or replicate to two different replication destinations. 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 data set. For a large number of LUNs that all support the same application, administrators can group them together in a data set and apply the policy to the data set as a whole. Using the policies to contain multiple data sets also simplifies administrative efforts, allowing administrators to focus on monitoring a handful of policies instead of each relationship individually.

Protection Manager comes with nine 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 duration
- Storage location on which to retain replicas
- Replication method (backup via Snapshot or SnapMirror)

Some Popular Topologies and Use Cases Covered by Policies

- ☒ **Local Snapshot copies:** Local Snapshot copies are locally retained point-in-time images of data. NetApp Snapshot is unique in that it incurs minimal performance overhead and can be safely created on a running system.
- ☒ **Open Systems SnapVault backup:** When installed on a system, the OSSV agent enables administrators to back up file data from that system to a NetApp filer.
- ☒ **Backup:** Backup copies data to tape, disk, or other media. With SnapVault, NetApp's core backup technology, only data blocks that have changed since the last backup are copied, not entire files. This enables backups to run more frequently and use less capacity.
- ☒ **Mirror:** Replicate data to another location:
 - ☐ **Backup then mirror:** Back up data to one location and replicate it to another
 - ☐ **Mirror then backup:** Replicate data in one location and then back it up in another
 - ☐ **Mirror and backup:** Replicate data and back it up simultaneously
 - ☐ **Mirror to two different locations:** Replicate the same data to two different locations

FUTURE OUTLOOK

As replicas continue to proliferate, most average firms create anywhere from three to nine copies of data across different systems, locations, and datacenters. Increasingly, managing and tracking these data copies will become even more paramount. With information running today's businesses against a landscape of legal and competitive pressures, firms need policy-based automation to ensure comprehensive data protection and reduce data risk.

In today's economic climate, Protection Manager can prove to be particularly rewarding because it improves operational efficiency and allows fewer headcount to manage ongoing replication activities. It will autodiscover and autoprotect new data and autoprovision storage targets. Since it leverages NetApp's thin provisioning technology, it can help reduce capital expenditures as well. Going forward, NetApp should consider integrating Protection Manager with its broader manageability portfolio. An example would be integration of Protection Manager with the SANScreen and VMware Insight Manager offerings.

Summary

In light of changing dynamics in the adjacent data protection and recovery and replication markets, NetApp is well positioned to capitalize on the growing use of replication technology in place of legacy backup applications. Its Protection Manager product offers higher level services to make this shift a practical reality in user environments seeking improved reliability, recovery, and compliance.

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 to learn more about IDC subscription and consulting services. To view a list of IDC offices worldwide, visit www.idc.com/offices. Please contact the IDC Hotline at 800.343.4952, ext. 7988 (or +1.508.988.7988) or 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 2010 IDC. Reproduction is forbidden unless authorized. All rights reserved.