

BACKUP ACCELERATION WITH EMC NETWORKER AND SNAPSHOT MANAGEMENT

A Detailed Review

Abstract

IT departments are under pressure to deliver higher levels of backup and recovery services than ever before to keep pace with round-the-clock business operations. By combining management of heterogeneous replication and snapshot technologies with backup operations, EMC® NetWorker® provides a unified “command-and-control” platform to deliver unparalleled performance and reduced complexity for mission-critical application and data protection.

February 2011

Copyright © 2007, 2008, 2009, 2011 EMC Corporation. All Rights Reserved.

EMC believes the information in this publication is accurate of its publication date. The information is subject to change without notice.

The information in this publication is provided “as is.” EMC Corporation makes no representations or warranties of any kind with respect to the information in this publication, and specifically disclaims implied warranties of merchantability or fitness for a particular purpose.

Use, copying, and distribution of any EMC software described in this publication requires an applicable software license.

For the most up-to-date listing of EMC product names, see EMC Corporation Trademarks on EMC.com.

All other trademarks used herein are the property of their respective owners.

Part Number H2807.3

Table of Contents

Executive summary	4
Audience.....	4
Snapshots and replication	4
Split-mirrors.....	5
Copy-on-write.....	5
Continuous data protection.....	6
EMC NetWorker and snapshot management	6
NetWorker Module for Microsoft Applications.....	6
PowerSnap Modules.....	7
Snapshot-assisted backup	7
Instant backup.....	7
Live backup.....	8
Proxy backup.....	8
Snapshot-assisted recovery	9
Rollback.....	9
Rollforward.....	9
File-level recovery.....	9
Restore from backup storage.....	10
Other snapshot-based options with NetWorker	10
SnapImage Module.....	10
NetWorker direct SCSI access.....	10
EMC Replication Manager and NetWorker.....	10
Considerations and recommendations	11
Components of the solution.....	11
Which solution is best?.....	12
Conclusion	12

Executive summary

In today's highly competitive business environment, data and applications must be available 24x7 and deliver consistent service levels, straining the resources of IT staff and stretching IT budgets. Better service levels are expected while the volume of data continues to grow. Companies are constantly evaluating new technologies to find solutions that meet or reduce the storage management requirements of their businesses.

Snapshot and replication technologies that create duplicate copies of data are essential to meeting stringent recovery time objectives (RTO) and recovery point objectives (RPO) in a frenetic 24x7 business environment. EMC® NetWorker® combines these solutions with traditional and deduplication backup to accelerate protection of data and applications and simplify management of a complete range of recovery services.

This white paper discusses various snapshot and replication solutions and how NetWorker integrates with each to provide tiered levels of backup and recovery protection.

Audience

This white paper is intended for anyone who is using, considering, or selling NetWorker.

Snapshots and replication

The decreasing cost of disk storage makes the use of snapshot and replication technologies attractive for multiple purposes including fast disk-based backup and recovery. By providing point-in-time virtual views or copies of file systems or volumes, these solutions provide an excellent means of faster recovery, easier management of large volumes of data, reduced exposure to data loss, and virtual elimination of backup windows.

Referred to by various names depending on the solution including snaps, snapshots, shadow copies, images, clones, and mirrors, there are typically two types of data replication products available, hardware-based (sometimes called controller-based) and software-based (also referred to as host-based). Hardware-based replication is integrated into disk arrays and enables the storage subsystem to create replicas. Typically these point-in-time snapshots are done at a block level and are generally independent of the operating system or file system. Software-based snapshots are implemented at the operating system and file system levels.

Snapshot data protection implementations vary by vendor including differing types and capacity for the number of snapshots. Each technique has its own benefit and requirements. The primary methods available are:

- [Split-mirrors](#)
- [Copy-on-write](#)

- [Continuous data protection](#)

Split-mirrors

Split-mirror, also referred to as disk mirroring, is a replication technique that duplicates every byte of the original data volume to another volume. These copies are known as business continuance volumes (BCVs), mirrors, and clones. A mirror can be temporarily suspended or split to create a point-in-time copy of data. During a split, the disk subsystem temporarily stops making updates to the mirror copy, which allows a frozen data point. The split-mirror can then be mounted read/write or used for backup and recovery. A recovery will impact the production volume; all other activities do not impact the production volume. After an offline backup is complete, the mirror is established and resynchronized with the product volume. A full data copy is available within the mirrored copy and therefore requires 100 percent of the capacity of the source. For example, 1 TB of data requires 1 TB of disk space for a mirror copy. With a mirror, if the original volume is lost, the alternate volume is an exact copy of the original. The issue of disk space must be considered if the intent is to store multiple copies.

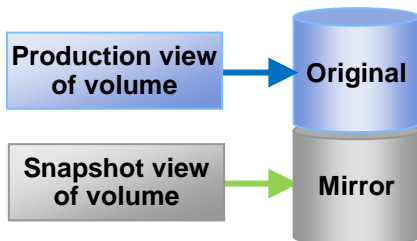


Figure 1. A mirror is a full data copy of the original volume.

Copy-on-write

Copy-on-write replication, also called snaps or snapshots, creates a virtual copy of an original volume. When the snapshot is initiated, the file system is frozen and a cache of disk space is created. Writes made after the initiation of the snapshot trigger a copy of the original block(s) to the cache. The production disk contains all current data, while the snap cache contains any original data that has subsequently been altered. A snapshot can be mounted for read/write access, or used for backup and recovery of the production volume. To recover the original disk to a point in time, the data from the snap cache is moved back to the original to re-create the volume as it existed at the time the snapshot was taken. Consequently, the copy-on-write snapshot is dependent on the original volume and a mount or recovery operation from the snapshot must hit the production disk. The required disk cache size will vary depending on the rate of change as well as the frequency and retention period of snapshots. Typically copy-on-write requires far less space than a mirror – on average 10 percent to 20 percent of the source volume size for each snapshot. Space requirements depend on how many writes and changes are made to the source volume and how long the snapshot is active.

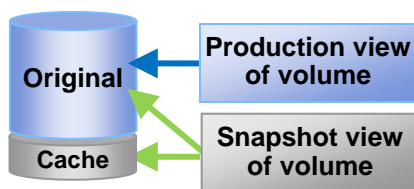


Figure 2. A copy-on-write snapshot relies on a cache area as well as the original volume to restore to a point in time.

Continuous data protection

A new class of replication called continuous data protection (CDP) automatically saves a copy of every change made to data while enabling the user or administrator to restore data to any point in time. While this sounds similar to mirroring, the differences lay in the capability of CDP-based solutions to capture and index every version of the data versus capturing only the most recent copy. For recovery, CDP is different from traditional backup in that you don't have to specify the point in time to which you would like to recover until you are ready to perform a restore. CDP volumes typically contain a full baseline copy of the production volume as well as byte- or block-level differences introduced as incremental writes and changes are made to the production system. This means that the CDP protection storage required is greater than that of the production volume and will depend on the size of the volume, change rate, and duration of the recoverable period.

EMC NetWorker and snapshot management

EMC NetWorker integrates management of multiple snapshot and replication technologies with traditional backup and recovery to tape and disk – as well as next-generation backup with deduplication. NetWorker provides the interface and intelligence to trigger point-in-time data copies from a number of array-based and host-based replication technologies. This capability coordinated with NetWorker application modules enables capture of transactionally consistent data states to ensure proper recovery and restart of applications such as Microsoft SQL Server, Microsoft Exchange Server, and Oracle. Tying replication and backup together with a single management tool is the ideal solution for simplifying recovery management.

The integration of snapshots and replication with backup and recovery is available via two NetWorker options: [NetWorker Module for Microsoft Applications](#) and [NetWorker PowerSnap™ Modules](#). Both integrate with EMC and other industry-leading replication technologies to manage the creation of point-in-time copies of data for backup and recovery.

NetWorker Module for Microsoft Applications

To simplify protection and recovery of Microsoft environments, the NetWorker Module for Microsoft Applications (NMM) provides a single, unified solution that leverages

Microsoft Volume Shadow Copy Service (VSS) for snapshot-based protection and recovery of Exchange, SQL Server, SharePoint, Hyper-V, and Active Directory.

Microsoft VSS provides a backup infrastructure for Microsoft Windows operating systems and applications, as well as a mechanism for creating consistent point-in-time copies of data. NetWorker and NMM make use of the VSS framework to enable consistent and efficient protection and recovery for major Microsoft Server business applications. This includes the following functionality:

- Enables creation of point-in-time copies of application data
- Leverages VSS software and hardware providers to enable flexible and efficient copy creation
- Facilitates backup to secondary media while applications are online and in use

PowerSnap Modules

PowerSnap Modules coordinate with NetWorker application support for Oracle, SAP, SQL Server, and IBM DB2 to create consistent replicas of messaging and database applications that reside on the supported storage technologies. Similar to NMM, once a point-in-time replica is created, PowerSnap verifies that the copy is clean and mountable to ensure recoverability and to enable off-host backup and restore operations to be performed as necessary. Snapshot policies are established and assigned in the NetWorker's management interface. In addition, PowerSnap provides command line utilities to enable the browsing and management of replicas. From these utilities administrators can accomplish the following:

- Recover whole snapshots
- Recover individual files and directories from snapshots
- Generate diagnostic reports

When recovery of an application from a snapshot is required, the restore process is managed from the command line utilities or the PowerSnap-enabled Application Module interface.

Snapshot-assisted backup

There are several backup types available when using NetWorker in conjunction with snapshots.

Instant backup

A point-in-time copy of data that is initiated and stored as a snapshot session or instance is called an instant backup. An instant backup is a block-level snapshot created from the application server and made available on disk but not written to tape. The snapshot is registered within the NetWorker media database to facilitate tracking for recovery.

Live backup

Snapshots tend to be transient and therefore backup to another media is typically done for disaster recovery. Once a replica is created, it can be mounted for backup to another backup device, such as disk or tape – and now also can be directed to EMC deduplication solutions including EMC Avamar® and EMC Data Domain®. In this process, called live backup or rollover backup, data is sent to a secondary storage medium and the snapshot can either be retained or deleted. This type of solution protects the data from both physical failures (such as the destruction of storage) and logical failures (such as an accidental deletion). Control of this process is orchestrated using NetWorker snapshot policies. Figure 3 shows the Create Snapshot Policy screen in the NetWorker Management Console.

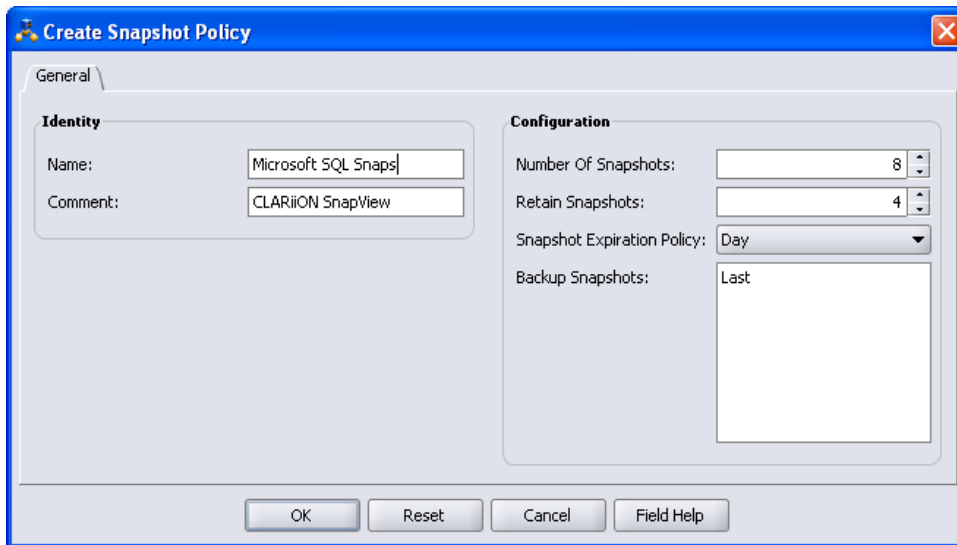


Figure 3. Snapshot management policies in the NetWorker Management Console

Proxy backup

To lessen the impact on file and application servers a proxy client or proxy host can mount and back up a volume or file system belonging to another client. The proxy client moves the actual data created by the application server within the snapshot or mirror to the backup storage such as tape or disk. While the proxy host is a server, this type of backup is referred to as serverless because it does not require the original application host to facilitate data movement. Figure 4 illustrates proxy host backup to disk or tape.

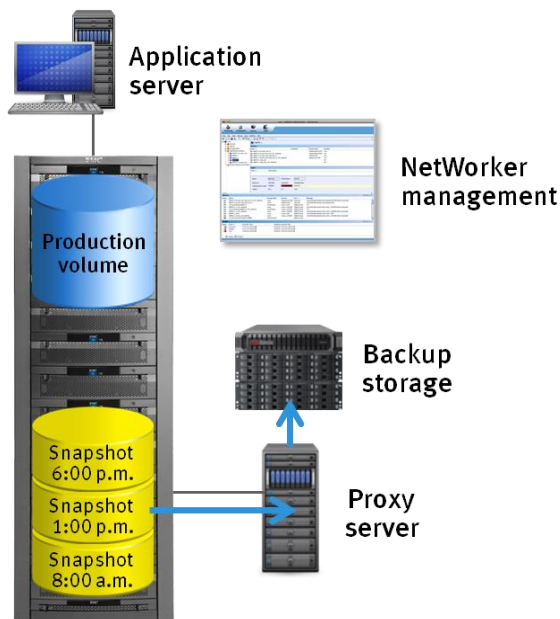


Figure 4. Backup from a snapshot via a proxy host

Snapshot-assisted recovery

There are several PowerSnap restore types.

Rollback

A rollback is the process of returning data to an earlier point-in-time copy in response to a recovery operation, and is a complete restore from a point-in-time copy to a standard volume without host involvement. Rollback restores are destructive by nature and can occur without having to retrieve data from the secondary storage (tape/disk).

Rollforward

A rollforward is the process of progressing data from a rollback using one or more instant backups. For example, if three snapshots were created at 10 A.M., 11 A.M., and 12 P.M., the user can perform a rollback to the 10 A.M. snapshot and then a rollforward to the 11 A.M. snapshot or even the 12 P.M. snapshot. You may perform a rollback from a more recent copy to approximate the same effect.

File-level recovery

When NetWorker initiates an instant backup, it captures an index to enable file-level recovery from a mounted replica. The recovery process is also referred to as instant recovery or file-by-file recovery.

Restore from backup storage

Data that has been saved through the live backup process to tape or disk – including Avamar and Data Domain – is recoverable in the same manner as any basic NetWorker restore. Save sets and individual folders or files can be restored via the recovery interface or from the command line.

Other snapshot-based options with NetWorker

SnapImage Module

A separate option for NetWorker, the SnapImage™ Module is a host-based software snapshot capability that enables high-speed backup to disk and tape. SnapImage provides live block-level backup targeted not at application servers, but at high-density file systems that have millions of files and directories. These systems are typically difficult if not impossible to back up in a reasonable time window. One of the key differences between this solution and PowerSnap is that the snapshot taken is not retained – it serves only the purpose of enabling efficient backup of a file system.

The SnapImage software runs on the host to be protected and takes a snapshot of its file system, creating a list of blocks to back up. In the process, SnapImage also generates a file catalog that is passed to NetWorker to enable individual file recovery. Similar to the [copy-on-write method](#), SnapImage creates a cache to track original blocks as changes are made on the production file system during backup. The SnapImage client communicates with tape devices using NDMP. During such recoveries, the solution identifies the list of blocks required for NetWorker to bring back the given file(s).

NetWorker direct SCSI access

Direct SCSI backup and recovery enable direct backup and recovery of Small Computer System Interface (SCSI) devices without the requirement of mounting them on the backup host if an access path is available to these devices over a storage area network (SAN). You can also use this feature to migrate to the NetWorker software to perform backup and recovery of business continuance volume (BCV) devices on an EMC Symmetrix® server (as well as backup and recovery of raw devices) over a SCSI bus.

The direct SCSI backup and recovery feature enables raw backups for the NetWorker software directly by using a SCSI target, which is usually accessible from a SAN proxy host. Typically, in a Symmetrix storage environment, these devices can be viewed from a primary application host and from a proxy backup host. The direct SCSI backup and recovery feature allows you to protect BCV devices from a proxy backup host as a raw backup.

EMC Replication Manager and NetWorker

EMC's Replication Manager is a software solution that manages the creation and disposition of point-in-time replicas of databases and file systems. Like NMM and

PowerSnap, Replication Manager eliminates the need for complex scripting typically involved with integrating replication technology and application environments and automates the scheduling and expiration of replicas. Replication Manager's value proposition extends beyond backup and recovery, providing access to replicas for the purposes of information sharing and repurposing for tasks such as reporting or testing.

Replication Manager is a distinct standalone EMC product line; however, it can integrate with third-party backup software such as EMC NetWorker to create tape backups of replicas. When backing up to tape/disk with a traditional backup, Replication Manager mounts the replica and starts a backup script automatically to execute the backup job. In this way, it is possible to integrate NetWorker with Replication Manager to perform live backup from replicas created by Replication Manager. This solution is ideal in scenarios where PowerSnap is not a fit for a customer requirement such as particular replication backup and recovery techniques or operating system support. Additionally, PowerSnap and Replication Manager can coexist on the same host, enabling a mix of technologies that best serves the customer's requirements.

Considerations and recommendations

Implementing snapshot and replication management in conjunction with backup requires an investment not just in hardware and software but also in identifying needs and requirements.

- What are the RTO and RPO requirements for applications and associated data?
- Are you unable to back up and recover servers or data in an acceptable backup window?
- Who will perform data recovery?
- How easy and transparent should the interface be?
- What technologies are currently in place (hardware and software)?
- What is the available IT budget and implementation timeframe?

Components of the solution

When considering the use of snapshot management with backup, it is important to understand all of the software and hardware components that make up the solution. The "stack" or list of technology components that must be in place and configured correctly include the following:

- A supported storage array, such as EMC Symmetrix or CLARiiON®, or others, with the appropriate level of supported firmware and API
- A supported snapshot software or CDP solution such as VSS, EMC TimeFinder® or SnapView™, EMC RecoverPoint, or others

- A supported operating system for the NetWorker server, client, and any proxy hosts utilized as well as the necessary software and licensing
- As appropriate, a supported volume manager, multipathing, and/or cluster solution
- The appropriate infrastructure such as SAN connectivity components or others

Note: EMC Consulting provides assistance to install and configure a complete backup and recovery solution including snapshot management.

Which solution is best?

The answer to the question of which replication technology is best must be based on the specific details about customer requirements gathered in an inquiry phase. The types of applications and availability of supported environments may trigger the choice of one implementation over another.

- Mirrors and clones are typically deployed for more critical data and in cases in which data changes a great deal in a short time, but they come at a higher cost than copy-on-write snapshots that can be used effectively for short-term protection and for data with minimal changes over time.
- Customers with Microsoft Windows environments may favor a VSS implementation over other options.
- To help protect environments that have large, highly dense file systems, some customers may choose a solution such as SnapImage.
- CDP delivers the most aggressive RTOs/RPOs by enabling any-point-in-time recovery and may be the right choice for high-value business applications and databases.

As customer environments and requirements vary greatly, it is recommended that NetWorker customers who want to invest in the capabilities of replication invest in experienced help for assessment, validation, and implementation. EMC is uniquely qualified to help customers assess needs centered on storage, backup, and replication. [EMC services](#) help customers successfully plan, build, and manage environments using a solutions framework that helps with the most demanding challenges.

Conclusion

EMC has a complete recovery management vision supported by the most robust portfolio of recovery management products on the market, which meet every requirement of a deployment framework. NetWorker and its options for snapshot management deliver on the promise of combining protection technologies to provide higher degrees of protection and recoverability for business data and information assets directly mapped to an organization's RPOs and RTOs.