

# Backup-to-Disk: The Key to Recovery Management

*Building the Foundation for Effective Operational  
Recovery Management*

---

## **Abstract**

This paper examines the benefits of backup-to-disk (B2D) and the central role it plays in operational recovery management. It identifies five immediate benefits of B2D, from faster and more reliable backup and restore to increased efficiency when backing up both SAN and NAS. Similarly, it describes five ways IT can take advantage of B2D to improve operational recovery. Finally, it reviews the key criteria an enterprise B2D product should support: concurrent read/write, Network Data Management Protocol (NDMP), direct file access, and single-step recovery.

---

Copyright © 2007 EMC Corporation. All rights reserved.

EMC believes the information in this publication is accurate as 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](http://EMC.com)

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

H3436

---

## Table of Contents

<b>Introduction .....</b>	<b>4</b>
<b>The benefits of backup-to-disk .....</b>	<b>4</b>
<b>Putting backup-to-disk to work .....</b>	<b>5</b>
The role of tape in a backup-to-disk strategy .....	5
<b>Key features to look for in backup-to-disk solutions .....</b>	<b>6</b>
Concurrent read/write .....	6
Network Data Management Protocol .....	6
Direct File Access .....	7
Single-step recovery .....	7
<b>Building a recovery management solution with a backup-to-disk foundation.</b>	<b>8</b>

---

## Introduction

Data backup has been a challenge for decades: it's traditionally slow, costly, labor-intensive, and unreliable. Additionally, tape, the traditional medium of choice, has its limitations. And now, heightened interest in corporate and IT governance, and the corresponding pressure to improve operational efficiency have organizations reviewing every business and IT process, seeking options to speed up processes and reduce costs. These factors are driving the need for a better way not only to protect data, but to implement a recovery management strategy.

Recovery management is a new, strategic approach to data protection that focuses on fast, reliable recovery as the aggregate goal of all protection activities. It includes technologies to recover the right data at the right time, easily and reliably. Combined technologies for backup, replication, continuous data protection (CDP), analytics and reporting, and management services create a solution that delivers higher levels of recovery than any single technology, no matter how robust.

To move to this new data protection strategy, IT managers are turning to redundancy, RAID, built-in high availability features, and enterprise backup-to-disk that encompass both SAN and NAS. In short, the key to a successful strategy is backup-to-disk (B2D).

## The benefits of backup-to-disk

B2D goes beyond mirroring, replication, or a simple copy of existing data on disk. Rather, it plays a central role in the organization's operational recovery management effort.

Backup-to-disk performs a job similar to backup-to-tape, but it allows the organization to take full advantage of the inherent capabilities of disk. Two important features are random access, through which data can be retrieved in non-sequential order, and concurrent read/write, in which the disk can read and write data on different parts of the disk at the same time.

Through B2D, organizations will experience a number of benefits:

- Predictable operational backup and recovery due to the greater reliability and consistency of disk storage
- Faster data recovery as a result of the random access capabilities of disk and direct file access
- Faster data backup because of the greater performance of disk
- Increased flexibility due to concurrent read/write capabilities available in advanced B2D products
- Increased administrative efficiency through the use of a single backup tool for both SAN and NAS backup and recovery

These advantages allow the IT organization to consistently meet more demanding recovery-time objectives (RTOs) and recovery-point objectives (RPOs). Due to the reliability and consistency of disk backup, IT can deliver higher-quality protection and recovery. B2D eliminates the errors and problems that, despite advances in tape technology, continue to hinder tape backup.

Backup-to-disk generally requires some changes to the backup process. The backup software must be reconfigured to see the disk as a backup target and treat the disk as disk so the organization can take full advantage of concurrent read/write and other disk capabilities. The changes, however, are modest, and the latest backup software supports a wide range of options, including disk as disk and disk as emulated tape.

In addition, B2D allows the IT organization to increase the ROI of tape storage itself. Tape can now be used in the roles for which it is ideally suited, such as the primary offsite disaster recovery medium. This takes advantage of the portability of the tape medium. IT can also augment tape operations with B2D and

---

thereby reduce the impact of tape backup on business processes, which often have to wait while tape backups are completed.

## Putting backup-to-disk to work

B2D has emerged as a powerful addition to any operational recovery management strategy. Highly flexible, B2D gives IT numerous options:

- **Staged backup to reduce impact on production systems**—by backing up to disk before copying the backup to tape, organizations speed backup while continuing to use tape for offsite storage
- **Comprehensive enterprise backup**—backing up both SAN and NAS data through a single, streamlined B2D backup tool
- **Backup consolidation**—backing up multiple diverse systems with a single B2D backup and recovery product
- **Selective recovery**—recovery of selective data very quickly rather than having to wait through a complete restore process
- **Fast recovery**—using concurrent read/write capabilities to recover some files even while others are still being backed up

### *The role of tape in a backup-to-disk strategy*

B2D does not usually eliminate the use of tape backup. To the contrary, most organizations continue to use removable tape media to meet a variety of data retention and disaster recovery requirements because of its portability and low cost. B2D, however, can be deployed in a way that is nondisruptive to the existing systems environment and flexible enough to fit within a company's backup window and disaster recovery guidelines.

In fact, combining B2D with tape adds immediate value, particularly in the area of restore. Traditionally, customers had to choose between fast backups and fast restores. In order to maximize backup throughput and avoid both tape delays and the wear and tear of tape drives, backup administrators tune the environment to keep the tape drives streaming during backup. This entails configuring the backup application to allow multiple streams (multiplexing) of data from multiple sources to be interleaved into a single data stream sent to a device. With multiplexing, the drives are kept spinning, thus offering the best performance with the least amount of impact on the tape drives. A price is paid, however, when it is necessary to recover. To restore a specific piece of data the backup application must demultiplex or unwind the data stream that is intermixed with multiple sources. As a result, the more the environment is tuned to provide fast backups, the more recovery performance suffers.

With B2D the organization can avoid having to make this choice by backing up first to disk. B2D continues to enable multiple streams to be sent to a device concurrently, but on disk it no longer interleaves these streams. Each is written contiguously. This maintains backup performance while eliminating the need to demultiplex data for a restore. In effect, administrators can tune for both backup and recovery performance.

Once backed up to disk, the data can be subsequently moved from disk to tape through cloning or staging. Both cloning and staging take advantage of the concurrent read/write capabilities of advanced B2D backup software to speed the process. Restoring from tapes that have been created by staging or cloning after data has first been sent to disk is consistently faster than restoring from data sent directly to tape because demultiplexing is avoided, as noted above.

---

## Key features to look for in backup-to-disk solutions

Not all backup software can handle B2D. This is especially true when IT wants to treat the disk as disk, instead of tape. Furthermore, even those that do claim to treat disk as disk often may not offer the full set of features required to get the most out of a B2D strategy.

Currently, four key features set apart top-performing B2D backup products:

1. Concurrent read/write
2. Network Data Management Protocol (NDMP)
3. Direct File Access (DFA)
4. Single-step recovery

The following sections provide details of these features.

### ***Concurrent read/write***

Advanced backup products, including EMC® NetWorker®, can handle read and write operations simultaneously. Such concurrent read/write gives the organization several advantages.

First, it allows for fast restores even while the backup is still in process. In effect, data sets that have been backed up earlier in the process can be accessed even before the remaining data sets have been backed up. This is not possible with standard disk or tape backup.

Why would a company want to do this? During a backup, for example, a user may mistakenly delete or damage an important file. Rather than wait for the entire backup to be completed, the IT group can restore that file, if it has already been backed up, before the backup is complete. This can save the user hours of time otherwise spent waiting, a clear productivity win for the organization.

Second, concurrent read/write speeds cloning. Organizations that use cloning to copy backed up data from disk to tape can begin the cloning process on any completed save set, even while the backup-to-disk is still running. Traditionally cloning could begin only after all the save sets in a group had been backed up. As a result, a backup and cloning process can finish in nearly the same time it takes to complete only the backup portion.

### ***Network Data Management Protocol***

Network-attached storage has become well accepted for enterprise storage. NAS enables organizations to consolidate file system data, reduce file storage administrative overhead, and lower management costs. Because NAS devices typically have specialized operating systems optimized for a single function, protecting these systems via traditional backup methods, which were designed primarily for block-level storage, is problematic. Most NAS vendors, in fact, discourage the running of additional software, such as backup agents that may degrade performance, on these devices.

In response, the industry developed the Network Data Management Protocol to provide centralized control of backup and restore for NAS devices. By defining a protocol that can be easily implemented by NAS vendors and other system and tape device vendors, NDMP allows backup, restore, and other data transactions to be performed without the installation of specialized software. The NDMP protocol has

---

achieved wide acceptance among companies including backup vendors, NAS vendors, and tape library manufacturers.

NDMP is an open standard for centralized backup of storage devices in the enterprise, allowing a backup application to control native backup-and-restore facilities on NAS. With NDMP, vendors follow common interfaces for common architecture data transfers. Similarly, NDMP uses common commands to pass file metadata (index data) to and from the backup software, regardless of the software package being utilized. Because NDMP integrates with a standard backup program on the NDMP server, the data format on tape is consistent no matter which backup vendor is acting as the NDMP client.

In the simplest NDMP protection configuration, the tape device or library is directly attached to the NAS system to be backed up. Although this reduces traffic on the network, directly attaching the NDMP backup to the NAS system presents drawbacks in organizations where IT wants to simplify, streamline, and consolidate backup. In that case, NDMP can be configured to back up the NAS system through a central backup server.

EMC has adapted EMC NetWorker to take full advantage of its NDMP Client Connection to deliver backup and restore of mission-critical data residing on NDMP-enabled NAS file servers, not just with tape, but also with disk. NetWorker's native functionality enables enterprises to automate, centralize, and consolidate the protection of all elements in a heterogeneous environment including NDMP-compliant NAS. NetWorker's disk-based backup option solves the problems that have been plaguing conventional tape-based backup systems for decades.

## ***Direct File Access***

Direct file access (DFA), a unique part of the NetWorker B2D feature set, uses the EMC Advanced File Type Device (AFTD) configuration to enable direct disk access to recover save set data. In practice, the NetWorker recover command uses the direct path to the save set file, automatically bypassing NetWorker's standard media-management process, to read information directly from the AFTD.

The benefits of DFA are increased availability and connectivity and a reduction in the latency (wait time) involved in using the recover program to restore data from NetWorker storage media. In short, it delivers faster and easier recoveries. Since file systems support concurrent access, DFA also enables true concurrent access to AFTDs. This means that save sets can be recovered simultaneously through multi-read/multi-restore access to save sets. If for any reason a NetWorker client cannot directly read and access a save set file that is stored on disk, then the recover program automatically uses the standard access process through NetWorker rather than DFA to recover the data.

## ***Single-step recovery***

Single-step recovery, as its name suggests, gives organizations the ability to complete a recovery from disk or from data that has been cloned or staged from disk to tape media in a single step. NetWorker understands where the backup data resides and enables recovery directly from tape, if appropriate, without having to first bring it back to the original medium by undoing the clone/stage operation. Some solutions in the market cannot do this.

---

## **Building a recovery management solution with a backup-to-disk foundation**

To move to the next level of recovery management, IT must simplify and centralize backup and recovery operations across the wide variety of platforms and storage topologies, both SAN and NAS, within the enterprise. Whether the organization has one large data center or a far-reaching network of branch offices, management needs assurance that all the data is safe, and most importantly, readily accessible.

NetWorker makes it possible for IT to confidently protect data and make it readily accessible. NetWorker lets you centralize operations and provide automation for efficient operational recovery management. With its comprehensive B2D features, such as concurrent read/write, NDMP support, DFA, and single-step recovery capabilities, NetWorker helps IT meet the demand for higher service levels while making the most efficient use of its backup and recovery budget. By leveraging disk storage, IT can deliver fast and reliable backup and recovery operations. In capitalizing on the B2D capabilities of NetWorker, IT gains an effective alternative to the tape-only backup solution for managing data—an alternative that allows IT to raise recovery management to a higher level.

## **About EMC**

EMC Corporation (NYSE:EMC) is the world leader in products, services, and solutions for information storage and management. Through information lifecycle management (ILM) strategies, EMC helps enterprises of all sizes manage their growing volumes of information—from creation to disposal—according to its changing value. EMC information infrastructure solutions are at the heart of this mission, helping organizations manage, use, protect, and share their information assets more efficiently and cost-effectively. The result? Information with greater business value and at lower management cost.