

Backup Solutions for the Celerra File Server

Copyright © 2000 EMC Corporation. All rights reserved.
Printed September 7, 2000

No part of this publication may be reproduced or distributed in any form or by any means, or stored in a database or retrieval system, without the prior written consent of EMC Corporation. The information contained in this document is subject to change without notice. EMC Corporation assumes no responsibility for any errors that may appear. All computer software programs, including but not limited to microcode, described in this document are furnished under a license, and may be used or copied only in accordance with the terms of such license. EMC either owns or has the right to license the computer software programs described in this document. EMC Corporation retains all rights, title and interest in the computer software programs.

EMC Corporation makes no warranties, express or implied, by operation of law or otherwise, relating to this document, the products or the computer software programs described herein. EMC CORPORATION DISCLAIMS ALL IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE. In no event will EMC Corporation be liable for

- (a) incidental, indirect, special or consequential damages or
- (b) any damages whatsoever resulting from the loss of use, data or profits, arising out of this document, even if advised of the possibility of such damages.

EMC, EMC² (the EMC logo), and Symmetrix are registered trademarks and Celerra, TimeFinder, and where information lives are trademarks of EMC Corporation. All other trademarks used herein are the property of their respective owners.

CONTENTS

Overview	1
Generic Backup Choices	2
<i>Full Backup</i>	2
<i>Incremental Backup</i>	2
Backup Configuration Methods	2
<i>Local and Remote Backup</i>	2
<i>Manual Local Backup – The server_archive Utility</i>	2
Automated Backup Strategies	3
<i>Backups using the Network Data Management Protocol (NDMP)</i>	3
<i>NDMP Backup Architecture</i>	3
<i>NDMP Version 1 Backups</i>	4
<i>NDMP Version 2 Backups</i>	4
<i>Other NDMP Backup Scenarios</i>	4
<i>Three-Way Backup</i>	5
Multiprotocol (NFS and CIFS) Backup	6
<i>NFS Backup</i>	6
<i>CIFS Backup</i>	6
Comparing Features of NetBackup and BudTool	6

OVERVIEW OF BACKUP SOLUTIONS FOR THE CELERRA FILE SERVER

Backing up file systems provides a measure of security and file system integrity. The primary goal of enterprise-level backup is to back up and restore information in an intelligent, secure, timely, cost-effective manner over all enterprise-wide operating systems. The nature of network and enterprise-wide backup is highly complex, involving:

- Identifying the data to be backed up
- Managing complex interactions with the backup media device
- Cataloguing and controlling the backed up file systems

The Celerra™ File Server can be tailored to perform a number of backup procedures and to facilitate control and management tasks based on such customer requirements as availability, scalability and performance.

Celerra File Server backup operations use either an NDMP-compliant backup tool, the server_archive utility, or a network (via NFS or CIFS) to copy a file system or incremental changes to a file system to a tape device (or a file) attached to a server. To ensure effective disaster recovery, the user must perform a complete tape backup and send the tapes off-site along with a printed list depicting the contents of the backup.

GENERIC BACKUP CHOICES

Backups fall into one of the following general categories:

- **Full backup**—Backs up all the data in the file, directory, or specified file system
- **Incremental backup**—Backs up only the files that have been updated since a previous backup

Full Backup

Systems that use a full backup strategy copy the contents of all the specified files or directories to tape. Full backups are easy to manage and restore, but can require significant tape storage.

Incremental Backup

Systems that use incremental backup create a full backup, and then incrementally back up only the changed files. Incremental backups provide more tape storage options and also let you customize the backup operation to meet your business needs.

A full backup is a level 0 backup. Incremental backups are assigned levels from 1 (the first incremental backup) through 9. When the backup is performed, the date and level are saved. Incremental backups save data that changed since the last backup at the specified level.

For example, consider this incremental backup strategy.

- Backup 0 is full, and occurs on Monday.
- Increment 1, on Tuesday, collects data that has changed since Monday.
- Increment 2, on Wednesday, collects data that changed on Tuesday.
- Increments 3 and 4 collect data for the rest of the week.

By definition, each incremental backup usually takes less time and requires less storage than a full backup.

BACKUP CONFIGURATION METHODS

The Celerra File Server offers the following backup configuration methods:

- Manual local backup
- Automated local backup
- Automated network backup
- Multiprotocol (NFS and CIFS) backup

The following sections describe these types of backup.

Note: You can use the TimeFinder/FS tool to create a point-in-time disk copy of a file system, or to create a dynamically mirrored copy on disk of a file system. You can also restore a file system from a TimeFinder™/FS copy. TimeFinder/FS copies are created on Business Continuance Volumes on a Symmetrix®. The backup strategies described in this paper, however, primarily refer to tape backups.

Local and Remote Backup

Local backup requires having one or more tape drives locally attached to Celerra File Server so that data transferred during backup remains “local”; that is, the data does not traverse the network.

In remote backup, the Celerra transfers the backup data to another server on which the tape drives are connected. The data may or may not traverse the network. If there is a direct server-to-server connection, the data does not traverse the network. If there is a network connection between the two servers, then the remote backup is a form of network backup. The type of network backup selected determines what information flows across the network: either both data and control information flow across the network or only control information (with the data itself being backed up locally).

Manual Local Backup—The server_archive Utility

Celerra’s internal backup/archive utility (`server_archive`) is built into its DART (Data Access in Real Time) operating system and run from the attached Control Station. The `server_archive` utility relies on directly-attached tape; specifically, a tape drive directly attached to the Data Mover that exports the target file system(s) and provides a set of functions, features and formats for customers to perform backup, archive, or data migration tasks. This utility supports:

- tar, cpio, and emctar formats
- String substitution on restores
- Selective backup or restore using wildcard characters
- Multiprotocol backup.

The `server_archive` utility also allows customers to store CIFS data like M8.3, M256, dos attributes, creation date, and access control lists (ACLs).

The `server_archive` utility does not include any robotics control or tape manipulation features. Instead, the `server_mt` command is used to rewind, position, and otherwise manipulate the tape. The `server_archive` utility also allows the use of files in place of tape devices, and allows for spanning multiple tapes.

Note: The internal backup/archive utilities in Celerra are enhanced to handle file attributes of NFS or CIFS and thus can be used on file systems used by either or both of those protocols.

AUTOMATED BACKUP STRATEGIES

The backup strategies described in the rest of this paper allow various degrees of automation, depending on the devices and configuration. The scenarios that follow represent a few of the many possibilities.

Backups using the Network Data Management Protocol (NDMP)

Network Data Management Protocol (NDMP), is an open standard communication protocol specifically designed for enterprise-wide backup and restoration of heterogeneous network-attached storage. NDMP enables centralized data management while minimizing network traffic. Celerra File Server administrators can back up critical data using any combination of backup devices and NDMP-compliant software backup applications.

NDMP separates the management of the backup/restore process from the flow of data to and from the backup media. This allows the use of third-party backup tools — (such as Legato BudTool® 4.6.1 or Veritas NetBackup® 3.2+) — executing on another platform elsewhere on the network to control the backup, while the data flows to locally-attached media or, with NDMP v2, to remote servers. The NDMP-compliant backup tool provides centralized control and media management through the NDMP protocol, thereby allowing a clean separation of control and data paths.

EMC qualifies third-party backup tools for use with Celerra File Server. The customer provides the qualified NDMP-compliant backup software. Celerra supports NDMP versions 1 and 2. Both versions support UNIX and NT attributes. Version 2 supports remote tape backup.

NDMP Backup Architecture

The main components associated with backup are the:

- Symmetrix on which the file system being backed up resides
- File server
- Backup device(s)
- Client with the backup control software

The NDMP backup architecture uses a client/server model in which the backup control software is considered an NDMP client to the Celerra File Server. Figure 1. NDMP Backup Architecture, shows a simple NDMP backup scenario. In this figure, lighter lines indicate control flow and heavier lines indicate data flow.

Figure 1. NDMP Backup Architecture

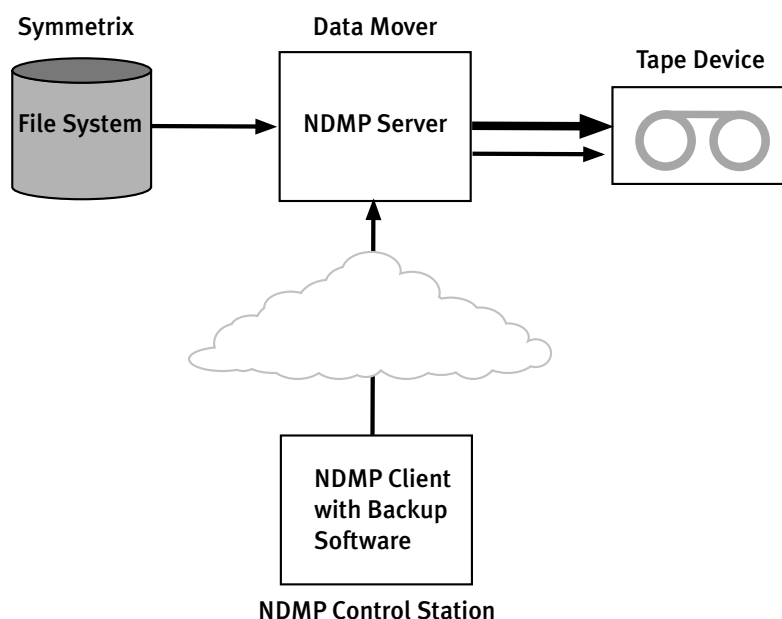
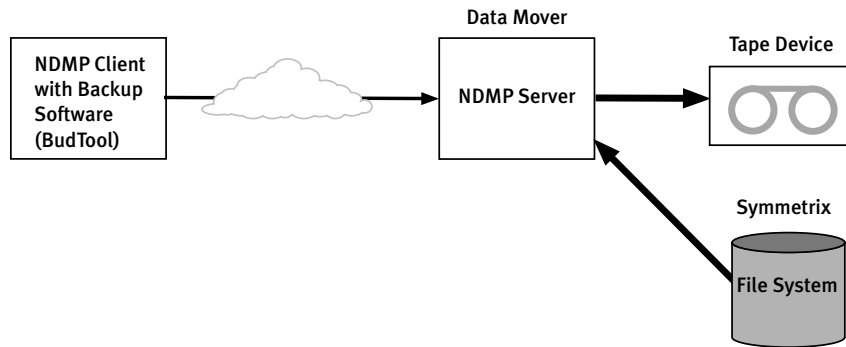


Figure 2. NDMP Version 1 Backup Architecture



NDMP Version 1 Backups

NDMP v1 allows backing up data only to locally attached tape devices. The data to be backed up is stored in file systems on the Symmetrix. During a backup operation, this data is copied to the backup device; for example, to a tape. File system data flows from the file system to the backup device(s) using a standard format, without traversing the network. In this scenario, only control data flows over the network to and from the backup software using NDMP, so network traffic is minimal. The backup software controls what is being backed up and also manages the database or catalogue of the data being backed up. Figure 2 shows how file system data (indicated by heavier lines) and control data (indicated by lighter lines) flow between these components.

NDMP Version 2 Backups

With NDMP v2, it is possible to direct the NDMP data stream from one Data Mover (acting as an NDMP data server) through a fast network connection to a remote Data Mover (acting as an NDMP tape server). The remote NDMP server can then access a tape device. This is known as three-way backup, since typically three hosts are involved:

- The host (NDMP control station) running the NDMP client
- The host (Data Mover) running the NDMP DATA service
- The host (Data Mover) running the NDMP MOVER (tape) service

In this scenario, both file system data and control information flow over the network, hence the need for a fast network connection. Figure 3 shows a typical NDMP Version 2 architecture. In this figure, lighter lines represent the flow of control information, and heavier lines represent data flow.

Other NDMP Backup Scenarios

An NDMP server (that is, NDMP running on one or more Data Movers in the Celerra File Server) can provide either of the following:

- An NDMP data service - Reads from disk and produces a data stream (in a specified format) or reads a data stream and writes to disk, depending upon whether a backup or restore is taking place.

Figure 3. NDMP Version 2 Backup Architecture

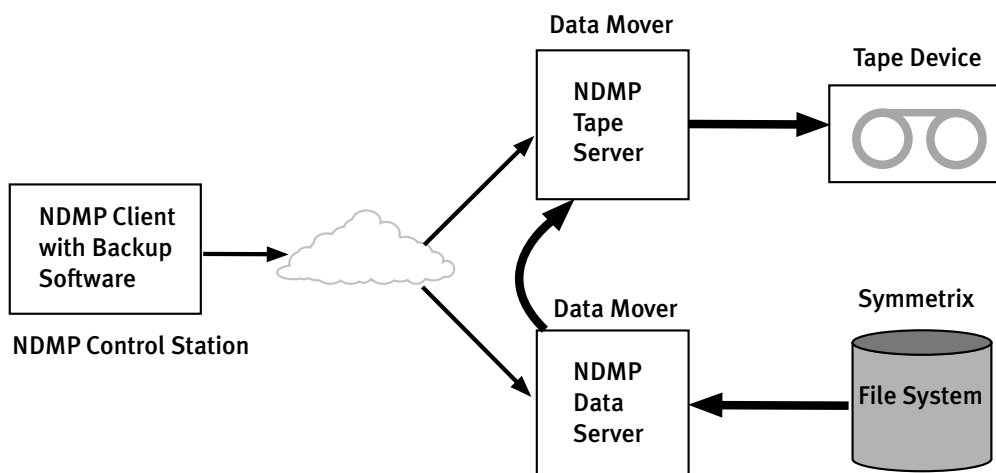
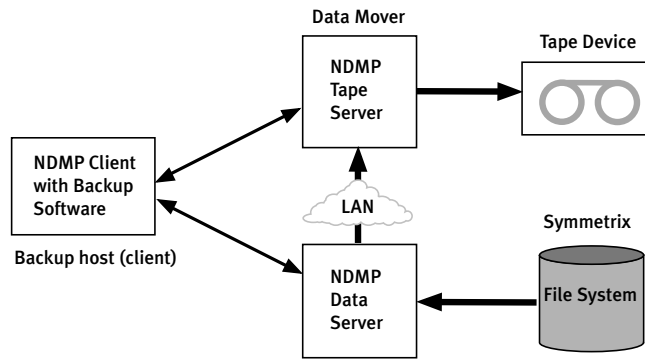


Figure 4. Backing Up Using a TCP/IP Connection Between Servers



- An NDMP mover (tape) service -Reads a data stream from the NDMP data server and writes it to tape, or reads from tape and writes a data stream, depending upon whether a backup or restore is taking place. This service performs all tape handling functions, such as resolving split-image issues. In addition, the mover service can enter the pause state while tapes are being changed by the NDMP client.

It is also possible to use NDMP to simultaneously back up to multiple backup devices physically attached to the NDMP host (that is, the Data mover). In this case, the Data Mover has two SCSI cards: one to provide connections to the Symmetrix and one to connect to each of the tape devices. Thus, a Data Mover can do two simultaneous backups across each SCSI using two different data streams or threads. The two simultaneous backups can share the same tape device robotics.

It is also possible to back up a host that supports NDMP but does not have a locally-attached backup device by sending the data through a raw TCP/IP connection to another NDMP server. Figure 4 illustrates this scenario.

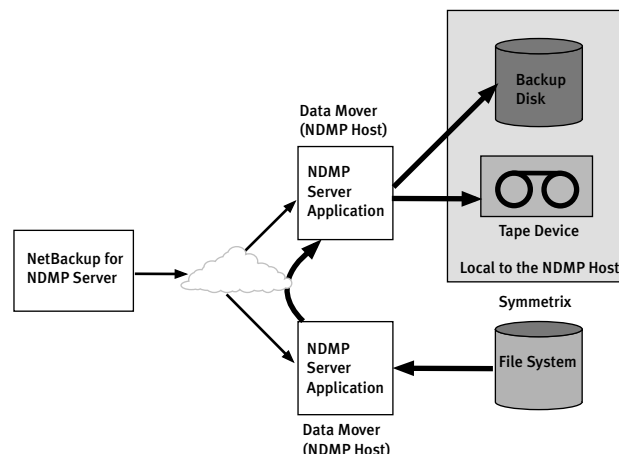
Three-Way Backup

The Veritas NetBackup tool also enables backups over the network. Figure 5 shows an example of a three-way backup scenario using the NetBackup tool. Notice that the data flow goes through the network.

After installation, using NetBackup with NDMP involves the following process:

- Setting up the network and configuring devices
- Setting up the Storage Unit that references the tape library attached to it
- Setting up the Backup Class. This references the file system on the Data Mover to which the Storage Unit is attached.
- At the NetBackup Backup/Restore screen, specifying the client to browse for Restores (where backup was done) and specifying the client to restore to (same as backup client or different). If restoring to a place different from where backup was done, providing an alternate path to a that location or Data Mover.
- At the NetBackup Backup/Restore screen, setting the backup type to “NDMP.”

Figure 5. NDMP Backup with Veritas NetBackup Tool



MULTIPROTOCOL (NFS AND CIFS) BACKUP

Network-based client backup can be used for backing up Celerra NFS or CIFS data using virtually any backup software that includes support for UNIX and/or NT with non-shared file systems.

NFS Backup

After being installed on a designated NFS machine, the backup client software—NetWorker, NetBackup, and others—can perform backup/restore operations over NFS through the NFS client machine. The backup software provides access to Celerra File System data using mounted NFS volumes and enables centralized control and media management.

Note: This technique is suitable only for NFS, since NT attributes are not preserved.

CIFS Backup

After being installed on a designated NT machine, the backup client software—NetWorker, NetBackup, and others—can perform backup/restore operations over CIFS through the CIFS client machine. The backup client software provides access to Celerra File System data using mounted Share volumes and enables centralized control and media management.

Note: This technique is suitable only for NT, since UNIX attributes are not preserved.

COMPARING FEATURES OF NETBACKUP AND BUDTOOL

The following table summarizes some of the differences between the features that the backup tools NetBackup and BudTool support when used with the Celerra File Server.

Tool	Multiprotocols	Robotic tape device	TCP/IP direct connection	Three-way backup
NetBackup	Yes	Yes	Yes	Yes
BudTool	Yes	Yes	No	No