Data Protector NDMP Integration Quickstart Guide

How to set up HP Data Protector NDMP integration and HP StoreOnce Backup Systems to protect data stored in HP X9000 and NetApp FAS3070 NAS filers

Technical white paper

Table of contents

Executive summary

This white paper provides quick start information for integrating Data Protector with an HP X9000 Network Storage System or NetApp FAS3070 Universal Storage System, for NDMP (Network Data Management Protocol) based backup tasks where the backup target is an HP StoreOnce Backup System.

Key take-away points:

- When using HP Data Protector for direct NDMP backups of HP X9000 and NetApp filers, the HP StoreOnce Backup System can be used as a deduplicating disk backup target.
- Additional Data Protector licenses are needed to enable NDMP backups.
- The StoreOnce Backup System must be configured to emulate tape devices (Virtual Tape) for the backup targets.
- Backing up to the HP StoreOnce Backup System delivers backup storage capacity savings of 90%+ with deduplication in the range 12:1 to 16:1 after around 30 backups. These deduplication ratios are similar to the best deduplication ratios achieved using regular file system backup, that is file system backups using the largest Data Protector configurable block size.
- StoreOnce Backup System replication can be used to create DR copies of virtual tapes at a remote location.

Audience

This document is intended for solution architects, project managers, engineers, and support personnel involved in planning, designing and configuring NDMP backup/recovery solutions using Data Protector with HP X9000 NAS technology, NetApp FAS filer technology, and HP StorageWorks Backup to Disk technology using HP StoreOnce deduplication for backup data reduction.

Familiarity with the following topics is recommended:

- HP X9000 architecture
- HP Data Protector
- HP StoreOnce Backup System (formerly: HP StorageWorks D2D) architecture
- NetApp FAS3070 architecture
- NDMP (Network Data Management Protocol)

Disclaimer

The configurations in this document are HP-recommended configurations. They are provided as a reference only, as configurations vary with specific customer needs. Where memory, processor count and speed, and I/O storage recommendations are given, these should be considered as minimum recommendations.

NAS concept

Network-attached storage (NAS) is file-level computer data storage connected to a computer network providing data access to heterogeneous network clients.

Although it may technically be possible to run other software on a NAS unit, it is not designed to be a general purpose server. For example, NAS units usually do not have a keyboard or display, and are controlled and configured over the network, often using a browser.

A fully-featured operating system is not needed on a NAS device, so often a stripped-down or customized standard operating system is used, such as Linux.

NAS systems contain one or more hard disks, often arranged into logical, redundant storage containers or RAID arrays (redundant arrays of inexpensive/independent disks) often referred to *primary storage*.

NAS removes the responsibility of file serving from other servers on the network.

NAS uses file-based protocols such as NFS (popular on UNIX systems), SMB/CIFS (Server Message Block/Common Internet File System) (used with MS Windows systems), AFP (used with Apple Macintosh computers), WebDAV, HTTP/HTTPS and FTP/FTPS to serve data to applications or users.

NAS units rarely limit clients to a single protocol.

You can now back up data residing on a filer to what is often called *secondary storage* in two different ways:

- On the Application Server
- Direct from the filer to tape using NDMP

The first method would be a "normal" backup, using the Data Protector Disk Agent, which is installed on each Application Server. During backup, data is transferred via the LAN to the system to which a tape device is connected.

The second way uses the NDMP integration of Data Protector and performs the backup locally on the NDMP server. This has two major advantages:

- The backup data does not need to be transferred over the LAN.
- There is no performance degradation on the Application Server.

NAS NDMP 2-way or local backup concept

The NDMP architecture uses a client-server model in which Data Protector is the NDMP client (DMA-Data Management Application) to the NDMP-Host Data Mover (NDMP server or DSP-Data Service Provider).

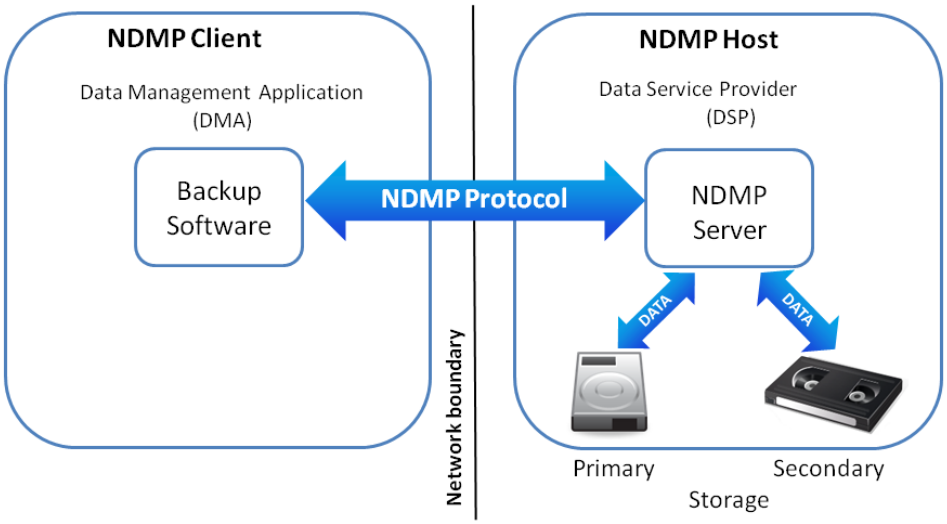
During an NDMP backup, backup data flows from the primary storage system to the Data Mover, and then on to the secondary storage system, such as an attached tape library backup device, without traversing the Ethernet network. Only the backup software's control data, catalog data, and tape library commands are send over the network, using NDMP protocol.

In a 2-way or local backup scenario (Figure 1) both primary and secondary storage are directly attached to the filer using either SCSI or Fiber Channel connections.

The Data Mover maintains a state machine for each NDMP client connection that executes and maintains backup and restore processes.

Note: Please refer to the filer vendor documentation to find the supported NDMP session concurrency scheme. Up to version 6.2, Data Protector only supports the 2-way or local backup concept.

Figure 1: NDMP 2-way backup architecture



NDMP backup

Preparation and prerequisites

StoreOnce Backup System (secondary storage)

1. Set up the StoreOnce system and create at least one virtual tape library with the appropriate number of tape drives. Select the desired tape emulation and number of tape slots. Make sure the library is mapped to an available Fibre Channel port.
2. Connect the configured library to your SAN segment.

SAN router

1. Create an Alias for each WWN (secondary storage: StoreOnce backup device, primary storage: NAS appliance).
2. Create a zone and include the Aliases.
3. Enable the zone.

NAS appliance (primary storage)

1. Set up the HP X9000/NetApp FAS3070 system and create at least one file system. Apply LTU were applicable.
2. Connect the NAS appliance to your SAN segment (via FC).
3. Detect the newly attached tape library/device.
4. Configure DSP (NDMP) on the NAS appliance.

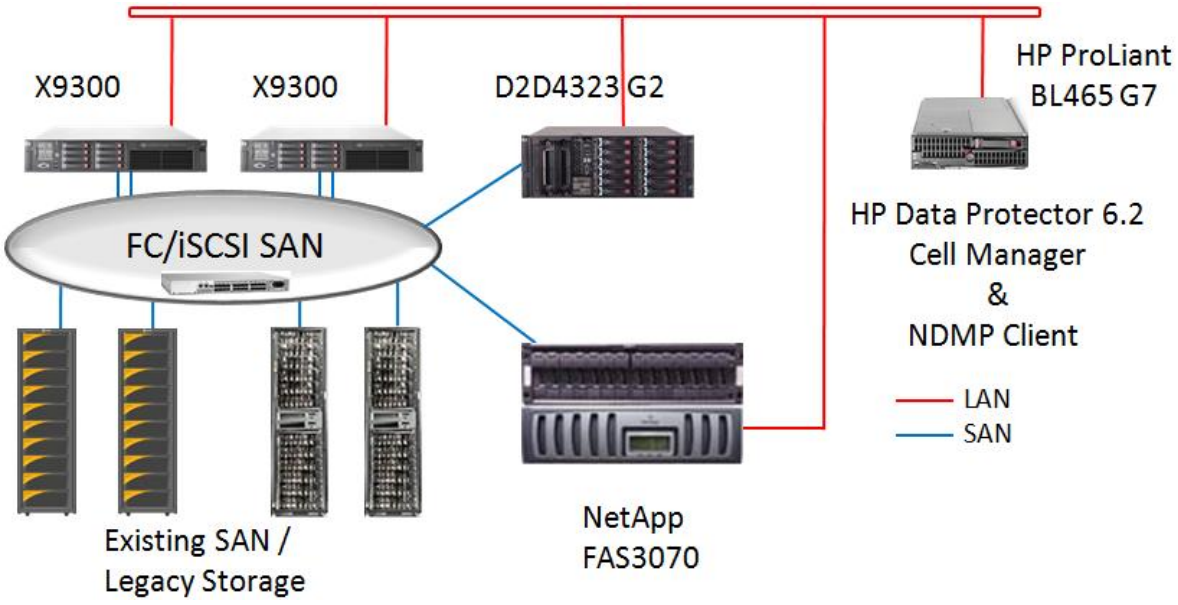
Data Management Application (DMA) and NDMP client

1. Set up a Data Protector 6.2 Cell Manager.
2. Install an NDMP Media Agent, which promotes the Cell Manager to an NDMP client.

   Note: The NDMP media agent can be installed on any client within the cell with a supported operating system. Always consult the latest support matrix for the components used to find any potential incompatibilities.

3. Ensure that you have correct Data Protector licensing for NDMP backups:
   - NDMP LTU for all platforms: 1xTB B7022BA, 10xTB B7022DA, 100xTB TD186AA
   - Appropriate number of library slots LTU: 1x 61-250/unlimited slots B6957BA/B6958BA, upgrade to unlimited slots B6958CA
   - Appropriate number of tape drive LTU for SAN, UNIX, NAS: B6953AA
   - Alternatively, licensing to tape drive LTU
   - Advanced Backup to Disk LTU for all platforms: 1xTB B7038AA, 10xTB B7038BA, 100xTB B7038CA
4. Configure the NDMP client.
5. Configure the backup library on the Data Protector Cell Manager.

Figure 2: NDMP 2-way backup topology



Note: Data Protector supports library sharing in a SAN environment.

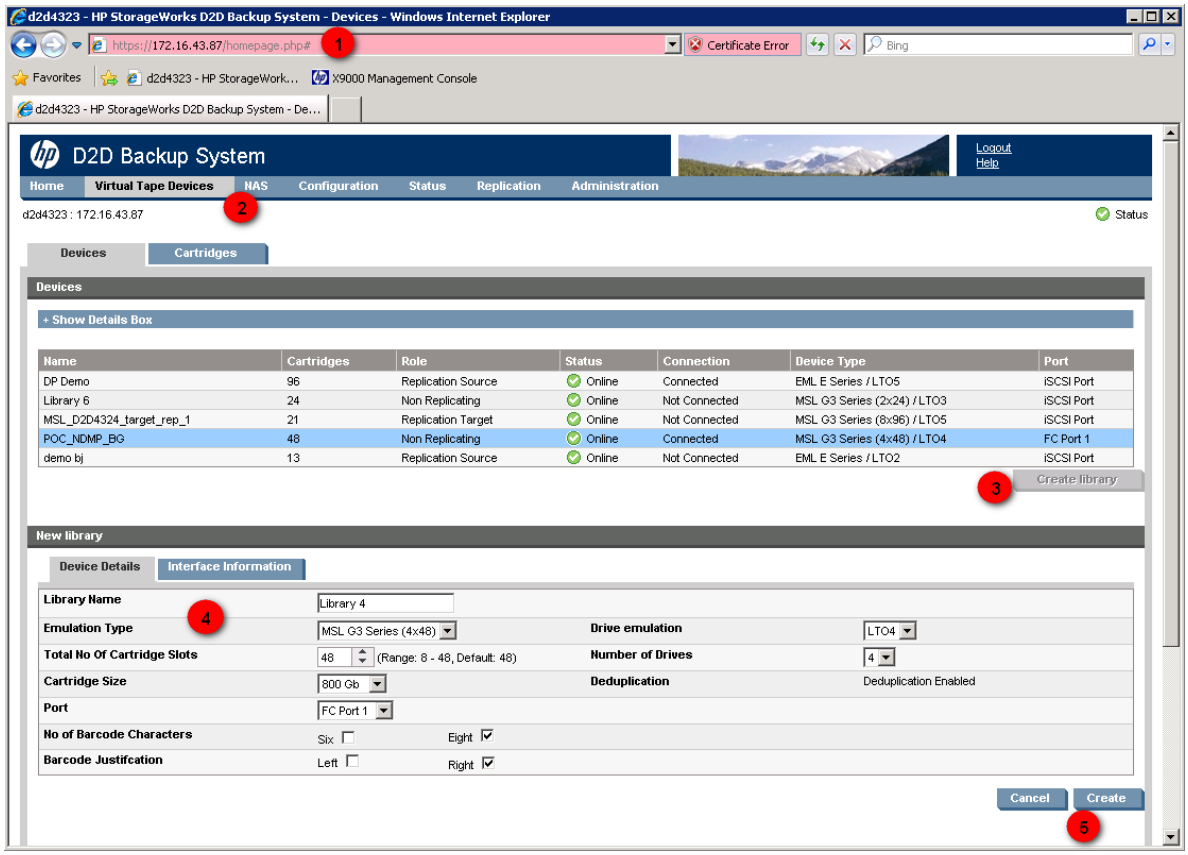
Configuration steps

The following sections describe the steps necessary for properly configuring the backup target (StoreOnce), the DSP (NAS appliance), and the DMA (NDMP client and Library configuration on the Cell Manager). They do not cover SAN and LAN configuration aspects. All DMA (NDMP client) settings are non-modified standard settings.

StoreOnce Backup System (secondary storage)

The StoreOnce backup system requires five steps to create and configure a new virtual tape library (VTL). These are shown in Figure 3, which depicts the StoreOnce management GUI:

Figure 3: StoreOnce Backup System Administration GUI

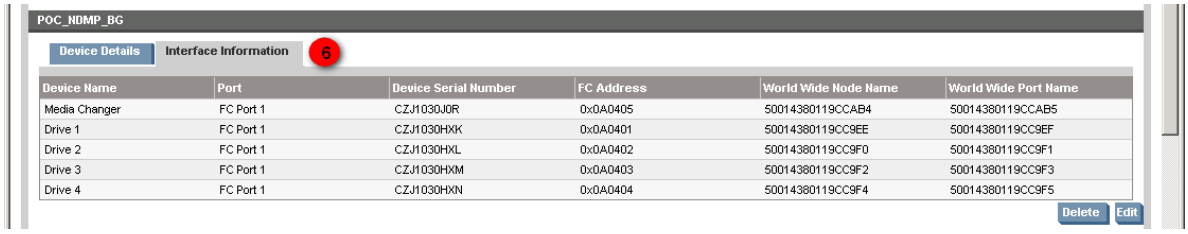


1. Log on to the StoreOnce web administration GUI. Enter the administrator user credentials and password.
2. Select **Virtual Tape Devices** from the menu list.
3. Select **Create Library**.
   This alters the lower part of the screen, allowing you to make changes to the library characteristics.
4. Change the library characteristics according to your needs. This includes:
   - Library name
   - Robotics emulation type
   - Tape drive emulation
   - Total number of cartridge slots
   - Number of tape drives
   - Cartridge size
   - FC port through which this library will be made accessible
   - Tape media labeling (barcode) details

   Note: Make sure the robotic emulation type and the drive emulation is supported by the DSP as well as the DMA. Always consult the latest support matrix for the components used to find any potential incompatibilities.

5. Click **Create** in order to create the virtual tape library according to your specification.The creation of the virtual library will make the new library appear on the StoreOnce management GUI.
6. To verify the interface information, select the **Interface Information** tab in the lower part of the screen:

Figure 4: Interface Information



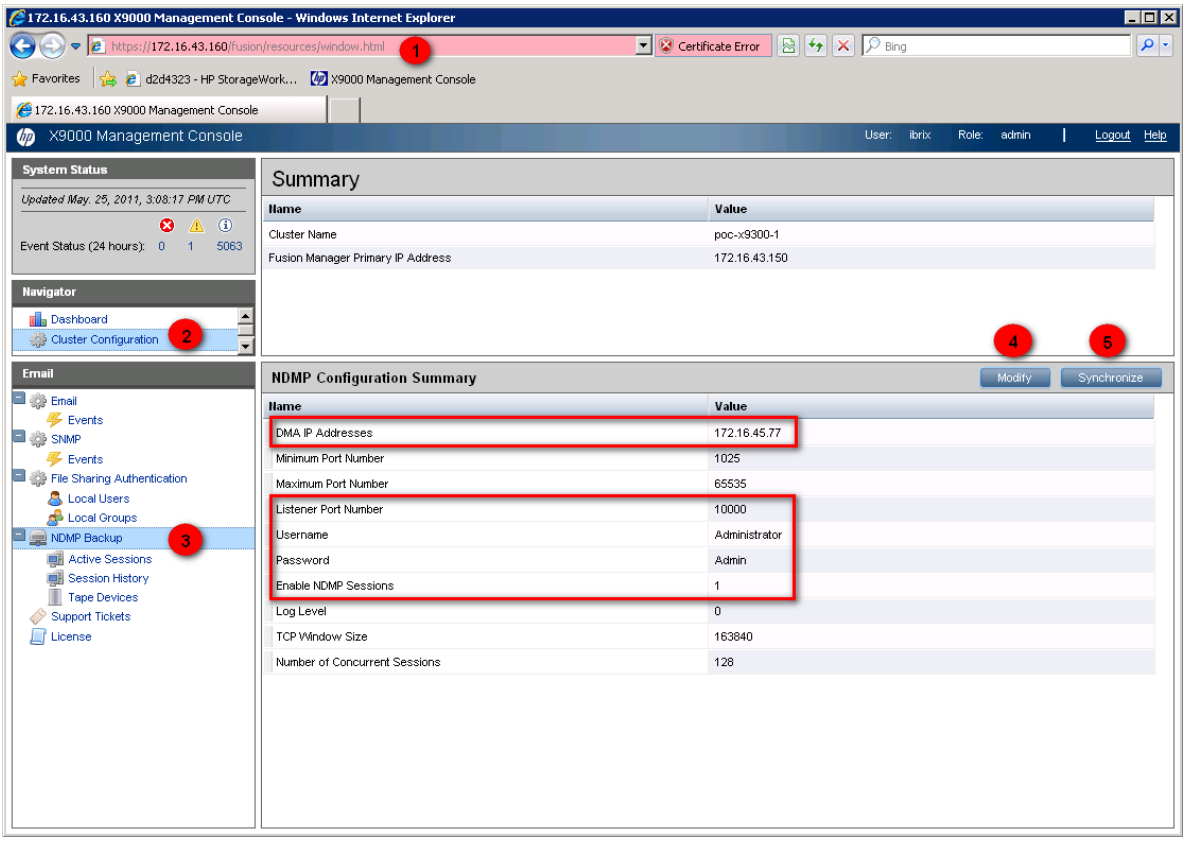
Note: "Device Serial Number" and "World Wide Node Name" are configuration details that are used in further configuration tasks such as setting up FC SAN zones on the SAN router or identifying tape devices attached to the DSP (NAS appliance).

DSP on NAS appliance (primary storage)

HP X9000 models

The HP X9000 Network Storage System requires the following steps to configure the cluster for NDMP backup. The steps are shown on Figure 5, which depicts a management GUI of an X9000 Network Storage Gateway consisting of two nodes:

Figure 5: HP X9000 Network Storage Gateway Administration GUI

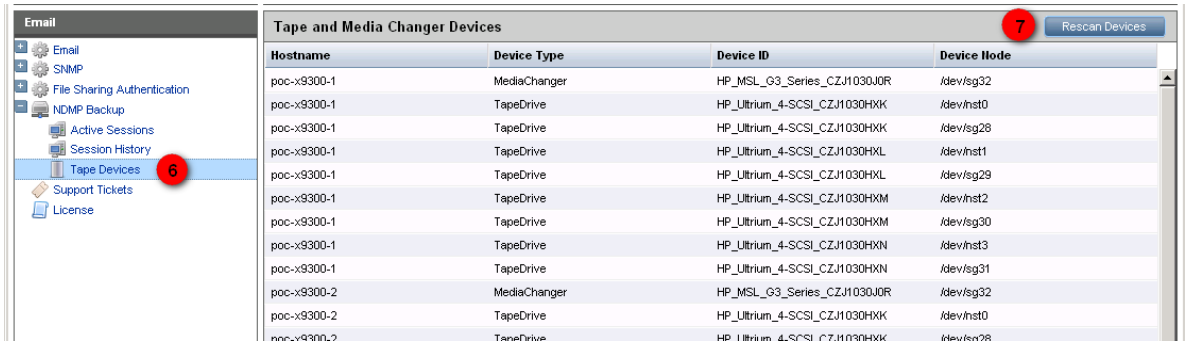


1. Log on to the HP X9000 web management GUI. Enter the administrator user credentials and password.
2. Select **Cluster Configuration** in the navigator list.
3. Select **NDMP Backup** in the Cluster Configuration.
4. Click **Modify** and make the necessary changes in the web form that appears. This includes the following details:
   - NDMP Sessions = Yes
   - Listener Port Number (typically the port is 10000)
   - Username and Password (must be a user with appropriate backup user rights)
   - DMA (NDMP client) IP Addresses
   - Enable NDMP Sessions = 1

   Note: The remaining configuration items can be left unmodified.
5. Click **Synchronize** to distribute the configuration changes to all nodes of the cluster.
6. When the configuration changes are distributed, make sure all backup libraries that have been configured in the previous step are discovered and visible on the HP X9000 Network Storage System.

   Select **Tape Devices** to verify which backup components are discovered. Refer to Figure 6 for details.

Figure 6: Discover Tape Devices



7. Select **Rescan Devices** if the newly attached devices do not appear.

Note: "Hostname" and "Device ID" are configuration details that are used in further configuration tasks within the DMA (NDMP client) configuration. Each segment server (Hostname) of the X9000 cluster may discover the attached backup devices (Device ID).

NetApp FAS3070

The NetApp FAS3070 platform has very limited access to NDMP configuration items through either the web GUI or the NetApp System Manager plug-in.

Note: Interaction to NDMP commands through the web GUI is limited to the following:

- Enable/Disable NDMP: the Enable/Disable NDMP Services page enables you to enable or disable the NDMP service on the storage system.
- Terminate NDMP sessions: the Terminate an NDMP Session page enables you to terminate a current active NDMP session.

There is no NDMP management functionality available through the NetApp System Manager plug-in.

It is recommended to use a FTP/SSH client such as PUTTY in order to manage the NDMP functionality using the CLI of the FAS3070 controller.

Verifying the newly-attached library components

To verify newly attached tape library components such as tape libraries (medium changers) or tape devices that are recognized and configured on the FAS3070 controller proceed as follows:

1. Log on to the FAS3070 controller using the telnet/ssh client. Provide administrator credentials and password to get the system prompt.
2. To verify the supported tape device types enter:

   ```
   tpc008> storage show tape supported [-v]
   ```

   If no options are given, the list of supported tape drives is displayed. The `-v` option is supported, which displays all the information about the supported tape drives, including their supported density and compression settings.

Note: Make sure you only select backup tape device types in your StoreOnce backup system configuration that appear on the list of supported drives (such as LTO4).

Displaying information about attached tape drives and libraries

To display information about attached tape drives and tape libraries (medium changers), and their associated serial numbers, enter the following command:

```
tpc008> storage show {tape | mc} [{alias | PPN | WWN}]
```

- `alias` is the logical name of the tape drive or medium changer.
- `PPN` is the physical path name.
- `WWN` is the worldwide name.

The following shows the result of a medium changer (`mc`) query:

```
Medium Changer:     fcs009:0.129
Description:        HP       MSL G3 Series
Serial Number:      CZJ04002H3
WWNN:               5:001:438007:5a2878
WWPN:               5:001:438007:5a2879
Alias Name(s):      mc4
Device State:       available
```

- `Serial Number` must match the device serial number of the configured library on the StoreOnce backup system.
- `Alias Name` is the device file to be used in the DMA (NDMP client) configuration.

The following shows the result of a tape device (`tape`) query:

```
Tape Drive:         fcs009:0.124
Description:        HP      Ultrium 4-SCSI
Serial Number:      CZJ04002D1
WWNN:               5:001:438007:5a27ba
WWPN:               5:001:438007:5a27bb
Alias Name(s):      st41
Device State:       available
```

- `Serial Number` must match the device serial number of the configured tape device on the StoreOnce backup system.

The `Alias Name` is the device file to be used in the DMA (NDMP client) configuration and can be verified by issuing the following command:

```
tpc008> sysconfig -t
```

The `sysconfig -t` command displays device and configuration information for each tape drive.

If you have a tape device that Network Appliance has not qualified, the `sysconfig -t` command output for that device is different from that for qualified devices. If the filer has never accessed this device, the output indicates that this device is a non-qualified tape drive, even though there is an entry for this device in the `/etc/clone_tape` file. Otherwise, the output provides information about the qualified tape drive that is being emulated by this device.

The following shows the result of a `sysconfig -t` device query:

```
Tape drive (fcs009:0.124)  HP      Ultrium 4-SCSI
rst41l  -  rewind device,        format is: LTO-2(ro)/3 2/400GB
nrst41l -  no rewind device,     format is: LTO-2(ro)/3 2/400GB
urst41l -  unload/reload device, format is: LTO-2(ro)/3 2/400GB
rst41m  -  rewind device,        format is: LTO-2(ro)/3 4/800GB cmp
nrst41m -  no rewind device,     format is: LTO-2(ro)/3 4/800GB cmp
```

```
urst41m -  unload/reload device, format is: LTO-2(ro)/3 4/800GB cmp
rst41h  -  rewind device,        format is: LTO-4 800GB
nrst41h -  no rewind device,     format is: LTO-4 800GB
urst41h -  unload/reload device, format is: LTO-4 800GB
rst41a  -  rewind device,        format is: LTO-4 1600GB cmp
nrst41a -  no rewind device,     format is: LTO-4 1600GB cmp
urst41a -  unload/reload device, format is: LTO-4 1600GB cmp
```

You can now verify the device file for the DMA (NDMP client) configuration. It is advisable to choose a no rewind device such as `nrst41a` or `nrst41h`.

Enabling NDMP services on the NDMP host/server

To enable NDMP services on the NDMP host/server issue the following command:

```
tcp008> ndmpd on
```

Note: The NDMP listener is on port 10000

Creating an NDMP password

To create a NDMP user password to be used in the DMA (NDMP client) configuration, issue the following command:

```
tcp008> ndmpd password username
```

This will create a NDMP user password similar to the following:

```
tpc008> ndmpd password Administrator
password ZVC5D0JUOJUw0BhV
```

Cut and paste the generated password into your “Import Client” task used on the DMA.

NDMP Client and DMA

The following is an overview of the steps required to set up the NDMP backup environment from a Data Protector perspective:

1. Push the NDMP Media Agent component to the designated NDMP client.
2. Import the NDMP host/server.
3. Create a dedicated NDMP media pool.
4. Create a backup library configuration.
5. Scan the library slots.
6. Format the media in the NDMP media pool.

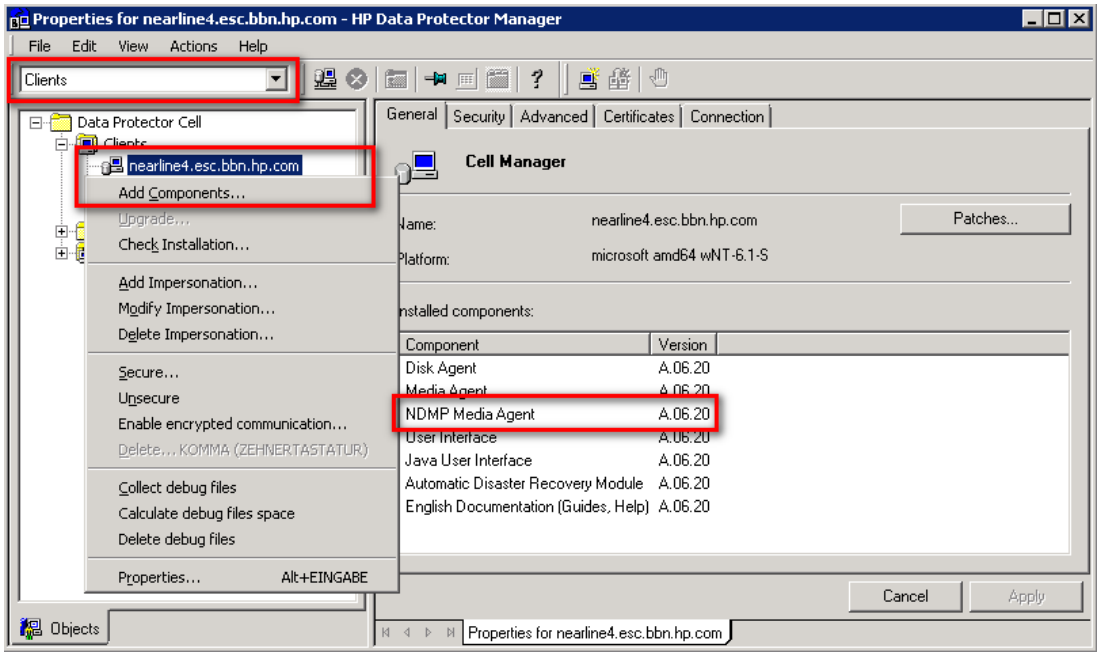
The following describes these steps in detail:

1. **Push the NDMP Media Agent component to the NDMP client.**

   You can install and upgrade Data Protector clients remotely by distributing the software using the Installation Server.

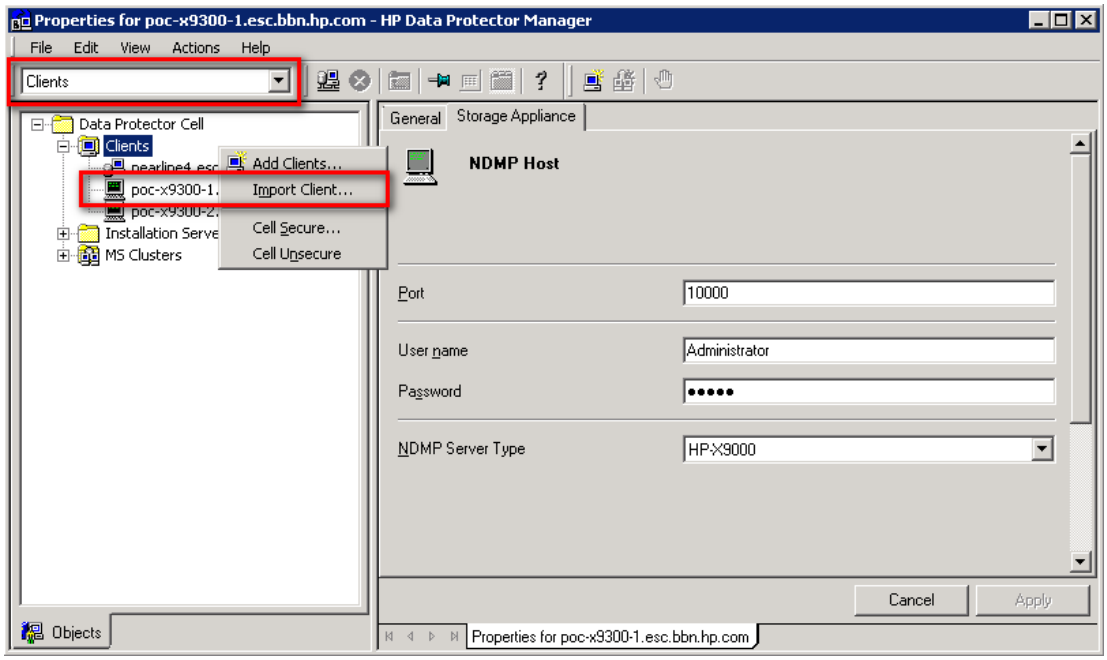
   Note: The NDMP client can be any client in the cell as long as the NDMP Media Agent is supported on the OP/Sys platform this client runs on. Even though the installation process reports that the standard media agent has been removed, it will still be listed on the client’s component list because the NDMP media agent provides the same functionality plus NDMP backup/restore functionality. Hence the NDMP media agent fulfills the legacy role of that client too.

Figure 7: Push the NDMP Media Agent component to the NDMP client



2. **Import the NDMP host/server.**

Figure 8: Import the NDMP host/server



Note: Importing a client means manually adding a computer to a Data Protector cell. On NDMP hosts/servers, no Data Protector component will be installed. When added to a Data Protector

cell, the system becomes a Data Protector client. Once the system is a client of the cell, information about the new client is written to the IDB, which is located on the Cell Manager.

3. **Specify information about the NDMP Server.**

   NDMP Server Type

   Select one of the available NDMP host types that matches your setup: NetApp, Celerra, BlueArc, Hitachi, or HPX9000.

   Port

   Specify the TCP/IP port number of the NDMP Server. The default is 10000.

   User Name

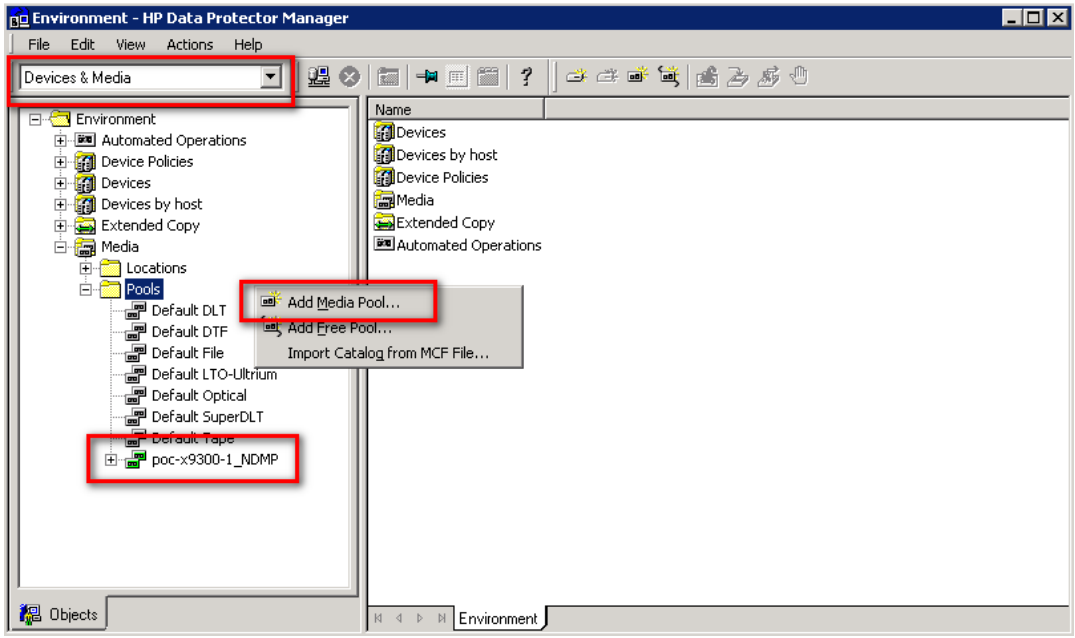
   Type the user name that Data Protector will use to establish the connection to the NDMP Server. Default might be “ndmp”. Check the vendor documentation for a valid “ndmp” user name and password.

   Password

   Type the password for the user. For NetApp FAS3070, use the password generated in one of the previous steps.

4. **Create the NDMP media pool.**

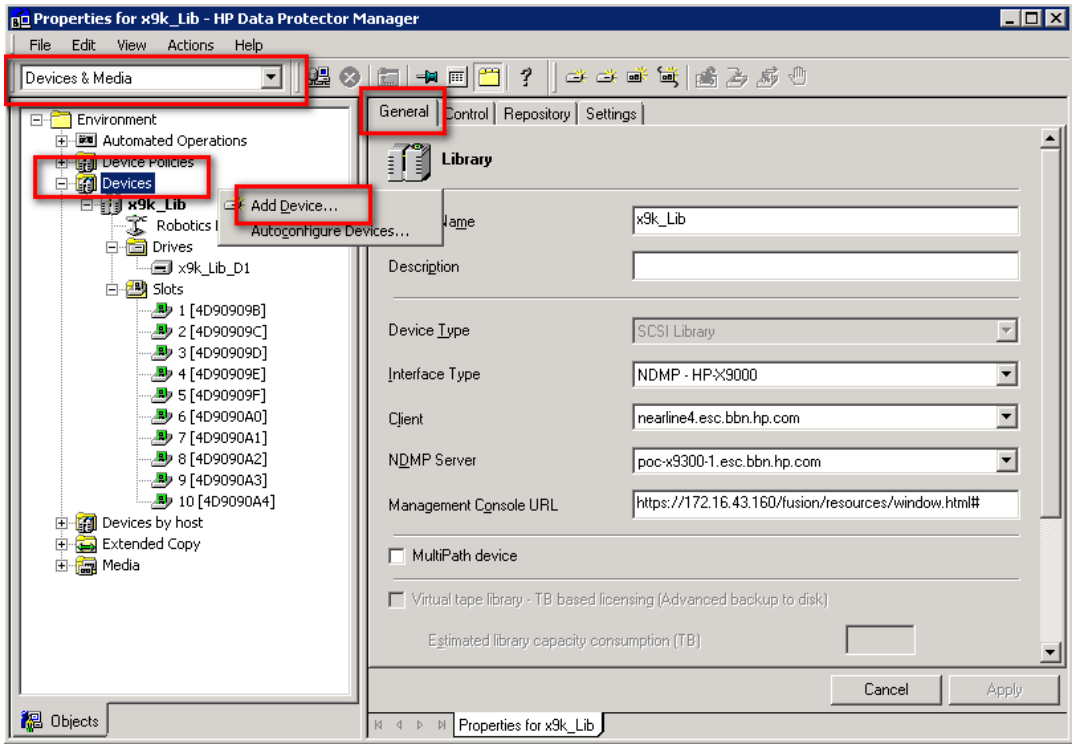
Figure 9: Create the NDMP media pool



Note: Although Data Protector provides default media pools, you should create your own media pool to suit your needs. Select the **Media Type** that matches the tape drive technology on your StoreOnce backup device. For example, select LTO-Ultrium if you have selected LTO4 on the backup device.

5. **Create a Backup Library configuration.**

Figure 10: Create a Backup Library configuration

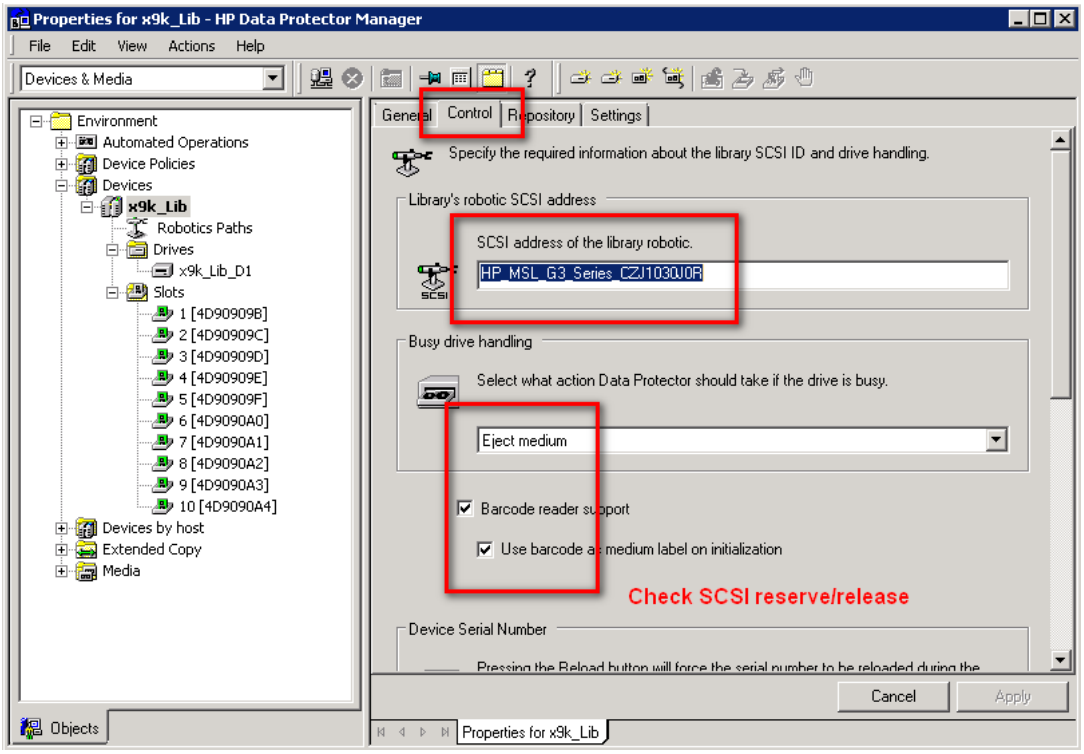


Once the device has been connected and configured on the NDMP host, the device can be configured with Data Protector. The library robotics and tape drives require the following common configuration items:

- In the **Device Name** text box, enter the unique name of the device.
- In the **Description** text box, enter a description (optional).
- In the **Device Type** list, select the SCSI Library device type.
- In the **Interface Type** list, if the device is to be shared with non-NDMP Data Protector media agent clients within the cell, select the SCSI interface type. If the device is to be used exclusively for NDMP backup, select the NDMP interface type specific to the NDMP host being configured (such as HP X9000, NetApp, Celerra, BlueArc, Hitachi).
- In the **Client** list, select the client that handles the NDMP communication (NDMP client).
- In the **NDMP Server** list, select the NDMP Server. If the SCSI Interface type has been selected this will be grayed out as the device robotic will be configured through SCSI on the selected Client.
- Enter a valid URL of the library management console in the **Management Console URL** text box. (Optional for the robotics configuration.)

6. **Enter the SCSI and media handling details.**

Figure 11: Enter SCSI and media handling details



SCSI library devices are large backup devices, also called autoloaders. They consist of a number of media cartridges in the device's repository and can have multiple drives handling multiple media at a time.

A typical library device has a SCSI ID (Windows) or a device file (UNIX) for each drive in the device and one for the library's robotics, which moves media from slots to drives and back again. For example, a library with four drives has five SCSI IDs, four for the drives and one for the robotics.

A medium is stored in a slot in the device's repository. Data Protector assigns a number to each slot, starting from one. When managing a library, you refer to the slots using their numbers.

Continue with the following steps:

7. Enter the SCSI address of the library robotics.
   Note: This is dependent on which Interface Type has been selected. If NDMP Interface has been selected, use the SCSI address as it appears on the NDMP host. For example, on the X9000 GUI (Device ID, such as `P_MSL_G3_Series_CZJ1030J0R`), or as device file as it appears on the NetApp CLI (such as `mc4`). If the SCSI Interface Type has been selected, enter the SCSI address for the device robotic that is attached to the Client system.
8. In the Busy Drive Handling list, select the action Data Protector should take if the drive is busy.
9. Select **SCSI Reserve/Release** (robotics control).
10. Specify the slots for the device. Use a dash to enter slot ranges and then click **Add**. For example, enter `1-3` and click **Add** to add slot 1, 2, and 3 all at once. Do not use letters or leading zeros.
11. In the **Media Type** drop-down list, select a media type for the device that you are configuring.

12. Click **Finish** to exit the wizard. You are prompted to configure a library drive. Click **Yes** and the drive configuration wizard appears.

Note: Creating a tape drive configuration requires the same steps as for the library robotics except that you have to perform the following actions:

- Enter the correct drive index.
- Assign the previously created NDMP media pool to every drive configuration you create.
- Verify the correct tape block size.
- Use "Lock Names" when appropriate.

The SCSI address follows the same convention as described for the library robotics.

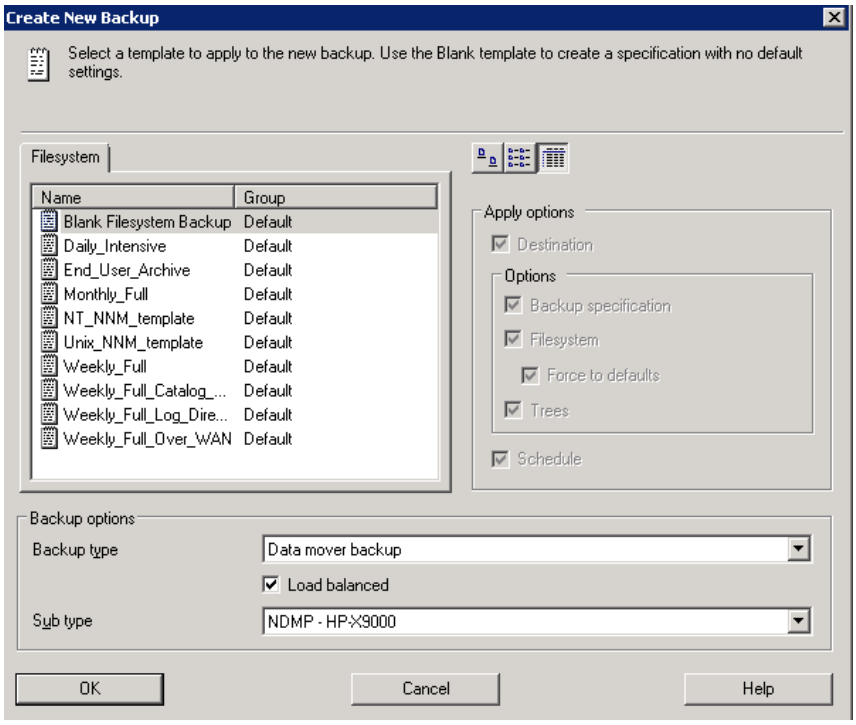
When the device configuration is finished you must scan the library slots and format the tape media to prepare the media pool for NDMP backups.

Creating an NDMP backup specification

Creating an NDMP backup specification in the Data Protector GUI follows a similar approach as for regular file system backups using the "normal" Disk agent (DA), except for a few details.

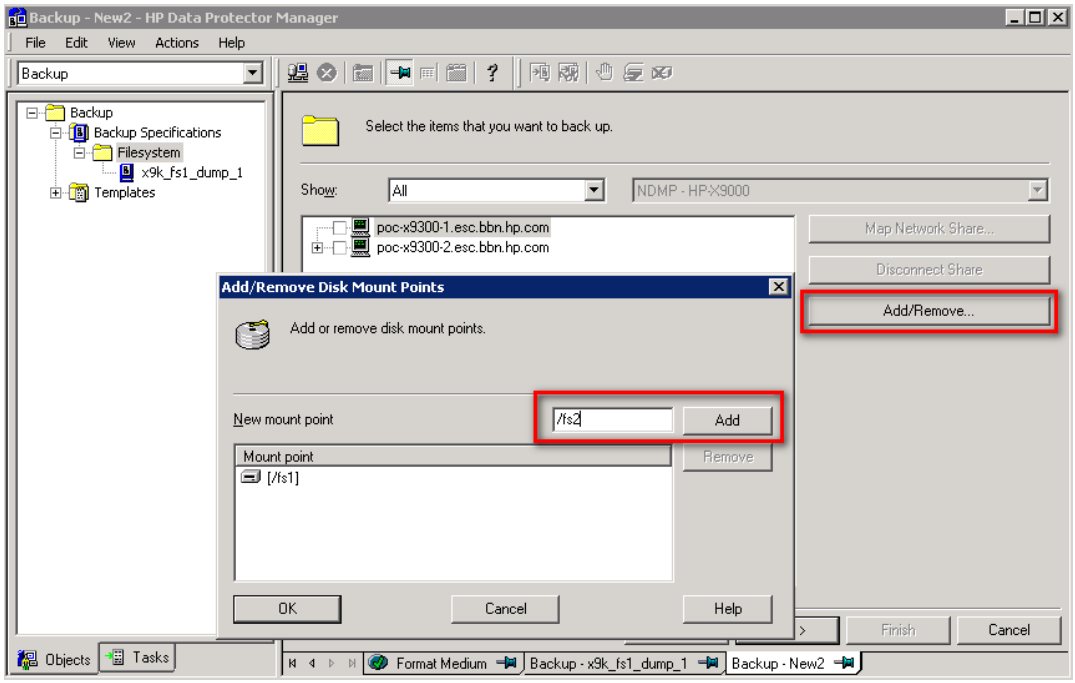
1. In the backup context, expand "Backup Specifications" and right-click on **Filesystem** to start the backup specification wizard by selecting **Add Backup** (see Figure 12).

Figure 12: Adding a backup specification using the GUI wizard



2. Select **Data mover backup** as the Backup type and your NDMP interface type in the **Sub type** field.
3. Continue in the wizard and start adding the backup source to the new backup specification (see Figure 13).

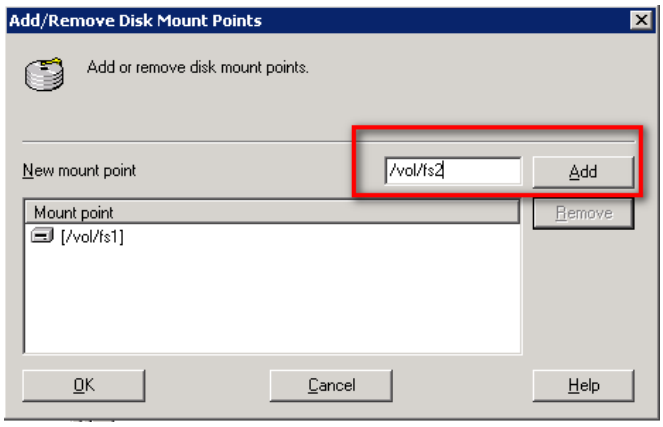
Figure 13: Adding the backup source for an HP X9000



4. As there is no Data Protector Disk Agent installed on any of the NDMP server platforms, you must manually specify the mount point to be included in the backup specification. You cannot browse existing file systems and have them added to the backup specification.

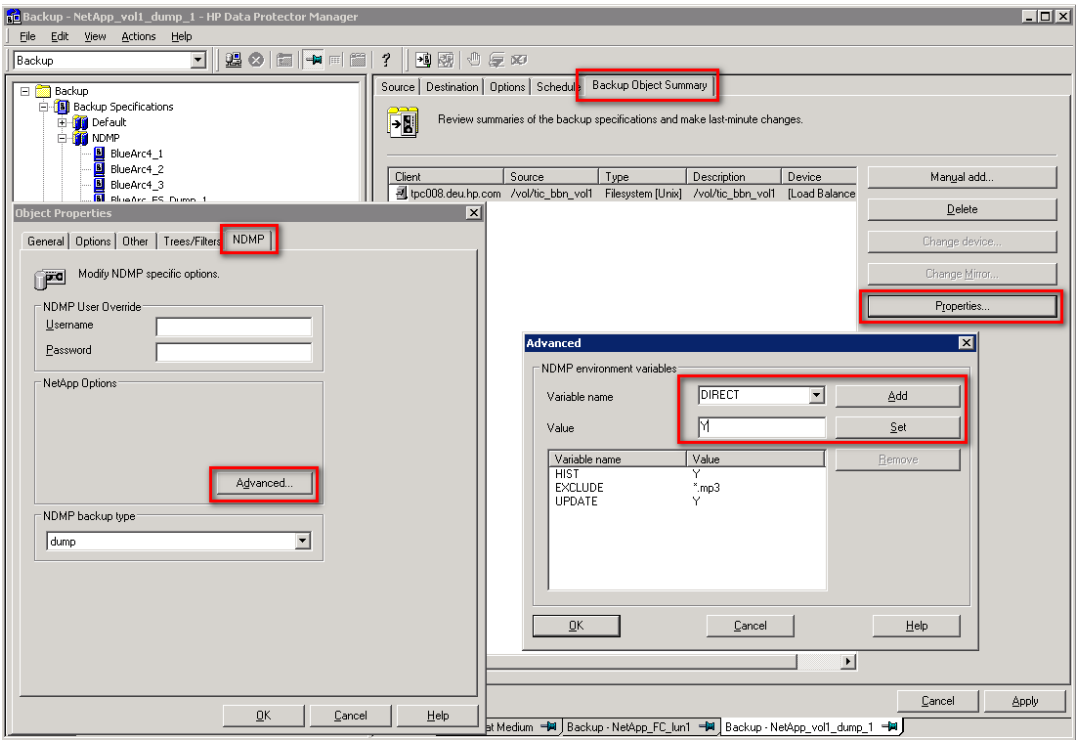
   Each filer vendor has a specific syntax on how to add mount points to the backup specification:
   - For an HP X9000, you must enter the file system name immediately after the root slash. For example: `/fs2`
   - For NetApp FAS3070, the syntax is: `/vol/fs2` (see Figure 14).

Figure 14: Add disk mount point syntax for NetApp FAS3070



5. Continue in the wizard and make your selection for the backup destination, global file system options and backup schedule. On the Backup Object Summary tab of the wizard you may configure each backup object individually (see Figure 15).

Figure 15: Backup summary



6. On Advanced NDMP Object Properties you can now specify your backup specific environment variable, such as EXCLUDE to exclude specific file types or directories.

   Note: Consult the NAS filer administrator handbook for supported environment variables and their usage.

Performance and implementation considerations

As described earlier in this paper, there are two methods of backing up the data on NAS filers. The first method uses the regular file system backup of the application server. The NAS share is backed up using the same process as the backup of a file system on a local disk. This method is well understood and simple, but not optimized for NAS filer protection.

The second method uses NDMP to enable a direct backup from the filer to the HP StoreOnce Backup System. This requires additional setup and licensing, as covered in detail in this paper, but is optimized for NAS filer protection. The performance and implementation considerations for both methods are summarized in Table 1.

Table 1: Summary of performance and implementation differences between the methods of protecting NAS filer data

	Regular file system backup	**NDMP-controlled backup**
Impact on production servers	Higher; the Data Protector Disk Agent runs on each protected machine. Requires a backup window.	None, all processing is done on the filer. No impact on the data network.
Impact on data network	Higher, all backup data is transmitted over the data network contending for bandwidth with the production data. Additional impact if the	None, the backup data is written over the storage network to the StoreOnce Backup

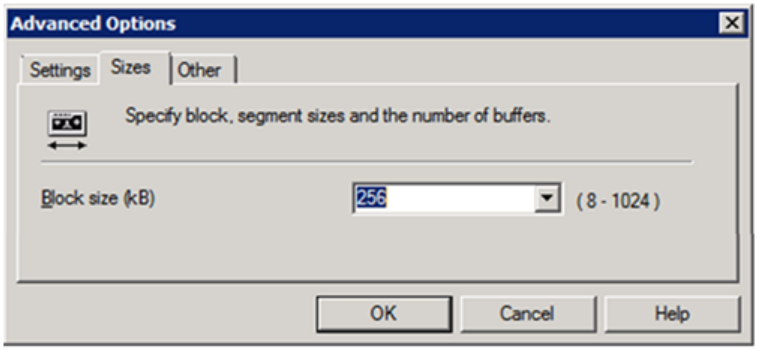
	StoreOnce Backup System is connected using the data network.	System.
Backup flexibility	It is a simple process to include and exclude files and directories for backup through the Data Protector GUI.	Standard dump format enables control over what is backed up via file and directory exclusion policies implemented using vendor-specific environment variables.
Backup throughput	Depends upon the 'front end' performance of the filer to the clients: the protocol used and filer connections to the data network (10 GbE), and the design of the backup infrastructure. Protection can be configured using multiple backup objects to enable multiple parallel backup streams to be written to the StoreOnce Backup System for best performance.	Depends on the 'back end' performance of the filer to the StoreOnce Backup System (8 Gb Fibre Channel), and the performance of the backup format (dump). Multiple virtual tape drives (one per backup object) can be used to enable multiple parallel backup streams during backup to be written to the StoreOnce Backup System for best performance. Data Protector supports load balancing for NDMP backups for optimal virtual drive usage.
Backup data deduplication (for a given data set and change rate)	The key variable is the block size set in Data Protector – larger block sizes have higher deduplication ratios.	Test results show that the block size set in Data Protector and the NDMP BLOCKING_FACTOR have no significant effect on the deduplication ratio.
StoreOnce Backup System backup targets supported	All; virtual tape drives via FC or iSCSI, and the Data Protector File Library on the NAS Share.	Requires a virtual tape target. Connectivity is over FC because there is no iSCSI support on filers.
Restore flexibility	All or selected files and directories are restored to their original or an alternate location.	All or selected files and directories are restored to the same or alternate locations on the same filer or the same model of the filer if the HISTORY environmental variable is set to "YES".
Data Protector licensing (assuming this is a new installation)	For backup and recovery management: • Cell Manager • Disk Agents (no cost) • Media Agent (no cost) For the StoreOnce Backup System: • Advanced Backup To Disk license for the usable capacity of the appliance ---- or ---- • Drive and Slot licenses for virtual tape libraries used	For backup and recovery management: • Cell Manager • NDMP Media Agent (no cost) • Direct Backup Using NDMP license for the capacity of protected data on the filer For the StoreOnce Backup System: • Advanced Backup To Disk license for the usable capacity of the appliance ---- or ---- • Drive and Slot licenses for virtual tape libraries used

Configuring Data Protector block size and the NDMP blocking factor

Within the limits of the backup device, the Data Protector block size determines the size of the blacks (data units) processed. The backup device processes data in the specified block sizes. The block size is specified as backup device parameter as shown in Figure 16.

The default block size value for all devices is 64 kB. Increasing the block size can improve backup performance and the deduplication ratio for regular file system backup. For NDMP backup, changing the block size has no significant effect on deduplication ratios.

Figure 16: Specifying the Data Protector block size used to write the backup to a virtual tape drive.



For an NDMP-controlled backup, the data stream is not in a Data Protector format. The default format is *dump* (others are possible), which can use different blocking factors. The blocking factor determines the size of the units of backup data moved and processed by the NDMP Server.

In Data Protector, the blocking factor is specified as an NDMP environment variable in the backup object properties, as shown in Figure 17. The syntax and default values for the blocking factor parameter are different for an HP X9000 and a NetApp FAS3070 as shown in Table 2.

Figure 17: Specifying the blocking factor environment variable in the backup object properties.

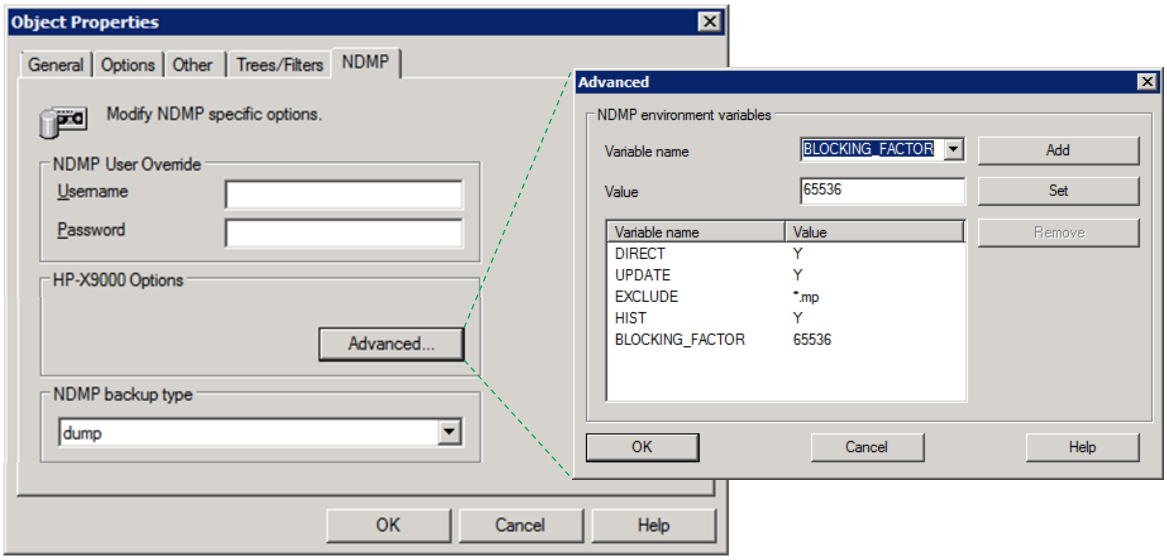


Table 2: HP X9000 and NetApp FAS3070 vendor-specific variables and values for the NDMP blocking factor.

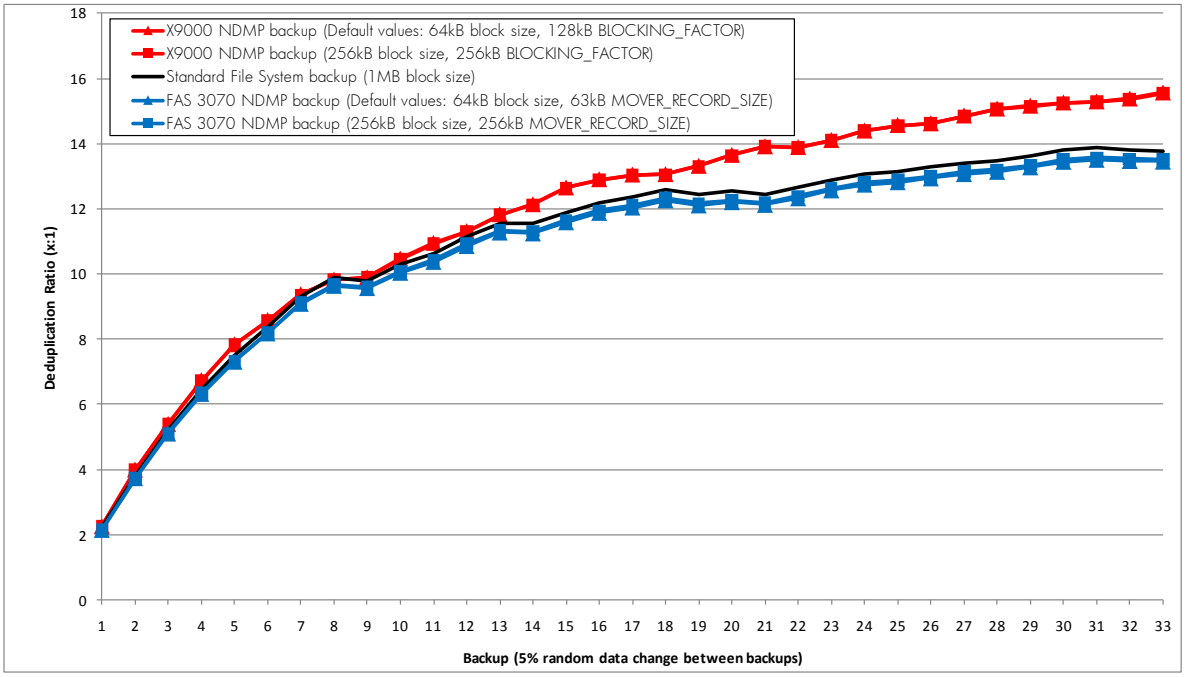
NAS filer	Blocking factor parameter	Default value	Other available settings
HP X9000	BLOCKING_FACTOR	128 kB	64–256 kB
NetApp FAS3070	MOVER_RECORD_SIZE	63 kB	4–256 kB

Deduplication performance test results

As part of the validation of Data Protector managed protection of HP and NetApp NAS filers using a HP StoreOnce Backup System as the backup storage, the backup data deduplication rates were measured. The test data set was made up of unstructured file data from a production environment. A backup scheme that simulated daily full backups was used. Between each backup, a random selection of 5% of the data was changed.

The tests were run multiple times from an HP X9000 and a NetApp FAS3070 to investigate the effects of varying the Data Protector block size and the NDMP blocking factor. The results are shown in Figure 18.

Figure 18: Comparison of deduplication ratios using different tape block sizes and different NDMP blocking factors versus a regular file system backup of the same data to an HP StoreOnce Backup System. The backup configuration was for full backups with a random 5% of the data changed between each backup.



Licensing considerations

As shown in Table 1, the file system and NDMP-controlled methods of protecting the NAS filer require different Data Protector licenses. Both require a Data Protector Cell Manager and either Advanced Backup To Disk, or Drive and Slot licenses for the StoreOnce Backup System.

The difference is that an NDMP-controlled backup requires a Direct Backup Using NDMP license. The licensing for this is calculated based on the configured capacity of the NAS filer. The result is that the NDMP-based method will have higher Data Protector licensing costs.

Conclusions

- The main benefits of NDMP direct backups are delivered when:
  - There is a need to minimize the backup load on the data network, application servers and backup media servers.
  - The filer is required to be backed up as a complete entity, although you can exclude files using exclude commands in the NDMP environmental variables set in the NDMP backup job.
- The benefits of NDMP direct backups should be traded off against the fact that NDMP backup s can be more expensive to implement (require additional licensing, such as an HP Data Protector Direct NDMP backup license), and more complex to manage and configure compared to regular file system backups.
- Backing up to an HP StoreOnce Backup System enables a deduplication ratio of between 12:1 and 16:1 over a period of 20+ backups with a simulated typical daily change rate of 5%. This provides an approximately 90% + increase in backup storage use compared with backing up the same data to a physical tape device or non-deduplicating disk.
- NDMP direct backups achieve similar deduplication ratios to file system backups using the largest Data Protector block size. (NDMP backups achieve higher deduplication ratios than file system backups using smaller Data Protector block sizes).
- The NDMP direct backup testing was conducted with various Data Protector tape block sizes along with HP X9000 NDMP BLOCKING_FACTORs and NetApp FAS3070 NDMP MOVER_RECORD _SIZEs. The results show that changing the tape block size and NDMP blocking factor had no appreciable effect on the backup data deduplication ratio. The recommendation is to use the default values in both Data Protector and the NAS filer.
- Testing showed that the HP X9000 implementation of NDMP backup delivers an approximately 10% + better backup data deduplication ratio compared with the NetApp FAS3070.

For more information

To read more about:

- HP Data Protector software, go to www.hp.com/go/dataprotector
- HP X9000 Network Storage System, go to www.hp.com/go/x9000
- HP StoreOnce Backup System, go to www.hp.com/go/d2d
- HP Data Deduplication, go to www.hp.com/go/storeonce

4AA3-5894ENW, Created August 2011