

# BACKUP AND RECOVERY OF DATA MANAGED BY EMC DISKXTENDER FOR WINDOWS

Recommended best practices

## Abstract

This white paper explains strategies for taking a backup of data managed by EMC® DiskXtender® for Windows. Four backup strategies are covered, and the recovery steps for each strategy are also explained.

February 2011

Copyright © 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](http://EMC.com).

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

Part Number h8176

## Table of Contents

<b>Executive summary</b> .....	4
Audience.....	4
<b>Introduction</b> .....	4
<b>Need for backup</b> .....	5
<b>Important considerations for a backup strategy</b> .....	5
<b>Controls available within DiskXtender</b> .....	6
Third-party backup mode for managed files.....	6
Special application filtering.....	7
Prevent delete from media.....	7
Metadata export.....	7
<b>Supported backup applications</b> .....	8
<b>Recommended backup strategies</b> .....	8
Strategy 1: Full backup from the host .....	8
Strategy 2: Stub file backup from the host, data backup from secondary storage (NAS).....	9
Strategy 3: Stub file backup from the host, retention or WORM media on secondary .....	10
Strategy 4: Block-based backup (VSS or SAN-based).....	10
<b>Conclusion</b> .....	11

## Executive summary

Data archived by EMC® DiskXtender® needs to be backed up regularly in order to guard against data loss due to accidental deletion.

Users need to use a well-defined backup strategy for this data. The strategy should take into account factors such as whether to back up only the stub files or the full data; whether to back up before archiving or after archiving; and how to prevent unnecessary fetches of data.

Controls available within the DiskXtender administrator allow you to tune your backup strategy.

Four possible backup strategies are explained in this white paper. Users are encouraged to adopt a strategy that is suitable for their environment.

## Audience

This white paper is intended for DiskXtender administrators. A basic familiarity with installation and usage of the product is assumed.

## Introduction

DiskXtender for Windows is a file archival solution from EMC. It enables users to archive less frequently used files to low-cost secondary storage. The archived files are replaced by a “stub file” on primary storage. In this case, the archived data is always available for access. When the user (or any application) accesses the primary data, DiskXtender detects that access, and retrieves (or “fetches”) the data seamlessly back to primary storage.

DiskXtender is mainly intended for managing reference or static data, but many customers use it to manage user home directories, or other data that may change with time. Thus it becomes necessary to have a backup of this data.

It is important to have a well-defined backup strategy in place when backing up data managed by DiskXtender. In the absence of such a strategy, backup software may negate some of the benefits achieved by DiskXtender. In this white paper, we describe the steps users should take to devise an effective backup strategy for the data that DiskXtender manages.

This white paper is organized as follows.

First, we describe the [important considerations](#) users should keep in mind when designing their backup strategy. Next, we describe the [controls](#) available within DiskXtender to manage the type of backup that you can perform. Finally, we list [four backup strategies](#), and explain how you can implement them. Readers can pick the strategy that best suits their needs, or they may design their own strategy using some variation of these four strategies.

## Need for backup

There are several reasons why you may want to back up your archival data.

- **To recover from accidental deletion:** If a user deletes the data on primary storage by mistake, then the data on secondary storage will also get deleted. If the secondary storage is “WORM” or some other type that prevents deletion, then the data on secondary storage will remain intact. However, the “stub file” that points to the secondary storage will be deleted. The stub file contains important information such as the object ID, clip ID, or some other pointer to secondary storage. Therefore, one needs to back up at least the stub files, even if the back-end storage is of the type WORM.
- **To recover from primary server failure:** If the primary server hardware fails, you will lose the stub files. For this reason, you need to have a backup of the stub files in order to recover data.
- **To recover a point-in-time version of your file:** If your data is changing often, you may sometimes need to recover a point-in-time version of your file. This can happen, for instance, if you want to compare the latest version with an older version of your file. A backup strategy enables you to keep multiple versions of your file.

## Important considerations for a backup strategy

Before you design your backup strategy, you need to consider the following points:

1. **Backup of complete data or backup of only the stub files:** If you back up your complete data, then you can restore it anywhere. However, this backup may take a long time to complete. If your secondary storage is WORM, or if it is being backed up separately, then you only need to back up the stub files. This helps to reduce the size of your backup data, and also reduces your backup window considerably. However, if two separate backups are being performed, the situation becomes somewhat more complicated because recovery of the stub file and recovery of the back-end data needs to be synchronized.
2. **Backup before archive or backup after archive:** Should you schedule your backup before or after you archive the data to secondary storage? This depends on many factors – including whether you want to take a full backup, or you only want to back up your stub files. For a full backup, it is preferable to take one complete backup before moving the data to secondary storage. However, if you only want to take a stub file backup, it is important to ensure that the first backup happens only after files are moved and purged.
3. **How to prevent unnecessary fetches of data:** When a backup application fetches a file, DiskXtender needs to distinguish between this access and any other application access. This is to prevent unnecessary load on the secondary storage system. In any archival solution, one important assumption is that only a small fraction of the data will be needed at any given time. In contrast, if the entire

archival data is fetched by the backup application, this assumption is no longer valid. Thus, DiskXtender needs to filter the read requests from the backup application by providing only the stub file, and not the actual file contents. There are controls available within DiskXtender to select the action to be taken when a backup application asks for file data. The following section provides more information on this function.

4. **Array-based or host-based snapshot techniques:** Today many storage arrays provide a snapshot facility. In these cases, you can make a point-in-time copy of storage data quickly without involving the host. Recent versions of Microsoft Windows provide Volume Shadow Copy Service (VSS), which can make a snapshot copy of a volume from the host. Can such a facility be used to take a backup of DiskXtender data? This is an important issue to be discussed. The following section provides more information on this function.

## Controls available within DiskXtender

There are several options available within DiskXtender for controlling the behavior of DiskXtender when it encounters a request for data from backup software. The following sections provide a brief explanation of each option. For detailed information on these topics, please refer to the appropriate *EMC DiskXtender Administration Guide*.

### Third-party backup mode for managed files

This option is available for each extended drive. There are three options:

- **Full backup:** DiskXtender provides the full file data to the backup application. Thus, a purged file data is read from secondary storage and given to the application<sup>1</sup>. This slows the backup considerably, but it has the advantage of making the backup fully contained. Such a backup can be restored to any other drive on the same server, or restored to any other server.
- **Fast backup:** DiskXtender provides only the extended attributes to the backup application (both for fetched and purged files.) This speeds up the backup considerably. However, the data can only be restored to the same server, and to the same extended drive.
- **Snapshot compatible:** DiskXtender provides the full file data for fetched files, but only the extended attributes for purged files. This provides the data on the extended drive “as is” without any change. Such a backup can be restored to the same server, and to the same extended drive.

The following table shows a summary of DiskXtender behavior for different file types in the three backup modes.

---

<sup>1</sup> The file is not fetched to the extended drive, if either “direct read” is enabled for the backup software, or if the backup application uses the “do not recall” flag.

**Table 1. DiskXtender behavior in the three backup modes**

Third-party backup mode	What is backed up		
	Unmanaged file	Purged file	Fetches file
Full backup	Full file data	Full file data	Full file data
Fast backup	Full file data	Stub information	Stub information
Snapshot compatible	Full file data	Stub information	Full file data

### Special application filtering

DiskXtender needs to filter accesses from special applications like backup applications and anti-virus software. The list of these special applications can be seen in the DiskXtender administrator (GUI). The list shows the names of executables, which are monitored by DiskXtender, and the action that DiskXtender takes for each of them. For backup applications, the action to be taken is called a “direct read.” If your backup application is not mentioned in this list, you can add it in the list. For more details, see the appropriate *EMC DiskXtender Administration Guide*.

### Prevent delete from media

This option, introduced in DiskXtender version 6.5, is available for select types of media services. It can be enabled per media service. If this option is enabled, delete transactions are not sent to the secondary storage. Thus, even when files are deleted on the extended drive, they are not deleted from the media. This function should be used with caution, because it can result in uncontrolled growth of storage space usage on the secondary storage. However, it is a powerful tool to retain data on secondary storage. You can use it to maintain “versions” of your files on secondary storage. This option can be utilized in a backup strategy where the application backs up the file stubs alone. It is useful with object storage systems like EMC Centera<sup>®</sup>. Since each file on EMC Centera is stored as a file clip, the “prevent delete from media” option enables users to keep multiple versions of the same file. Each version of the file stub can point to a different file clip on EMC Centera.

### Metadata export

If you do not use a full-fledged backup application<sup>2</sup>, you can use this option in DiskXtender to take a backup of your metadata. Metadata export has limited functionality; it is not comparable to a proper backup application.

<sup>2</sup> IMPORTANT: Metadata exports should not be enabled if you already use a backup software that is EA-aware. If the EA-aware backup software relies on the archive attribute to determine which files should be included in an incremental backup, then a metadata export can interfere with the process. This is because the metadata export process clears the Archive attribute. If the attribute is cleared by a metadata export, then the file is not included in the next incremental backup by the backup software because the backup software sees the cleared attribute and skips the file, assuming that the file was backed up in a previously scheduled process.

## Supported backup applications

The *EMC DiskXtender Software Compatibility Guide*, available on the Powerlink® website, provides a complete and updated list (including supported versions) of the backup software qualified for use with DiskXtender. The qualified applications include:

- EMC NetWorker®
- EMC Avamar®
- Microsoft Windows NTBackup
- Symantec Backup Exec
- Symantec NetBackup
- HP OpenView Storage Data Protector
- CA ARCserve Backup
- BakBone NetVault
- CommVault Galaxy Enterprise Edition

## Recommended backup strategies

In the following section, a few backup strategies are listed. You are encouraged to use one of these strategies, or a variation of these strategies.

### Strategy 1: Full backup from the host

This strategy will work well with any of the supported backup applications. You should use this strategy when:

- You have ample space available on your backup media.
- Recovery on any server is an important consideration in your system environment.
- Your backup window is manageable.
- Your data does not change very often.

The full backup strategy is simple – enable the “full backup” option in DiskXtender, and use any qualified backup application to take a full backup of data. Ensure that you take your initial complete backup before the data is purged by DiskXtender. Subsequently, the backup application will typically request file data only for changed files.

If a file is modified after its backup is taken, then the file will get deleted from secondary storage. You should ensure that the next backup occurs before the file is purged. You can set the purge rule in DiskXtender so that the file’s age is taken into consideration, and the age delay (for purging) is larger than the backup interval. Thus,

when the backup application asks for the changed file data, DiskXtender will supply the data directly from the primary disk.

Preferably, you should restore full backups to the same location (extended drive). Since the backup is fully contained, it can be restored to any server, or to any other location on the same server. While doing so, you should ensure that the new location is not a DiskXtender managed extended drive. This is because the backup files contain extended attributes (metadata). If these are restored to a DiskXtender managed extended drive, then the metadata may point to a stale location on the secondary storage.

One important consideration with this strategy is the frequency with which the backup software takes a complete backup. There are some applications, such as EMC Avamar, that take a complete backup only once (initially), and then take incremental backups from then on. However, other applications may take a complete backup of the data at regular intervals (such as once a month). For such cases, you will incur the cost of fetching the full data every time a complete backup is taken.

Another drawback of this strategy is that it consumes considerable space on backup media.

### Strategy 2: Stub file backup from the host, data backup from secondary storage (NAS)

This strategy works well when the data is archived on NAS secondary storage. You should use this strategy when:

- The secondary storage (NAS) is being backed up independently.
- Recovery on any server is not an important consideration in your system environment.

In this strategy, enable the “fast backup” option in DiskXtender, and use any qualified backup application to take a backup of your extended drive. The backup application should be EA-aware<sup>3</sup>. When the backup application requests the file data, DiskXtender supplies only the extended attributes to the application.

The backup frequency on your NAS secondary storage should be the same as that of primary storage.

Before you start your backup from the extended drive, ensure that you have moved the file data. You may enable realtime moves, or run a background scan before starting the backup. This will ensure that the extended attributes are populated by DiskXtender before the backup starts. Subsequent changes to the file data will not result in any changes to the extended attributes.

---

<sup>3</sup> An EA-aware backup application understands the NTFS Extended Attributes (called EAs). DiskXtender uses EAs to store its metadata. Most modern backup applications are EA-aware. Consult your backup application documentation to ascertain whether it is EA-aware or not.

Recovery of data in this strategy is a two-step process. The first step is to restore the data on the NAS secondary storage, and the second step is to restore the stub file on primary storage (extended drive).

This type of backup can only be restored to the same extended drive and to the same server.

### Strategy 3: Stub file backup from the host, retention or WORM media on secondary

You should use this strategy when:

- The data on your secondary storage is not being deleted (you are using WORM<sup>4</sup> media, or the data is under retention<sup>5</sup>).
- Recovery on any server is not an important consideration in your system environment.
- You want to reduce the space occupied by backup sets.

In this strategy, enable the “fast backup” option in DiskXtender, and use any qualified backup application to take a backup of your extended drive. You should start your backup only after files are moved. The backup application should be EA-aware. When the backup application requests the file data, DiskXtender supplies only the extended attributes to the application.

When files are under retention, they cannot be deleted or modified. Even the stub files cannot be deleted or modified. Thus the only reason to restore the data is if you lose the extended drive completely (due to a failure of the DiskXtender server).

Recovery of data in this backup strategy is easy since the back-end data is never deleted. To start the process, restore the stub file on the extended drive, and it will correctly point to the data (Clip ID in case of EMC Centera) on the back-end storage.

### Strategy 4: Block-based backup (VSS or SAN-based)

This strategy uses block-based backup software on the host (VSS) or on the primary storage array (if the extended drive is a LUN on SAN storage.) Block-based backup techniques bypass the file system. One example of block-based backup is VSS. VSS makes a snapshot copy of the extended drive, and takes a backup from this snapshot copy. Since the snapshot is a new drive, the DiskXtender filter driver does not monitor accesses to this drive. Many backup applications (for example, NetWorker) make use of VSS.

You should use this strategy if recovery on any server is not an important consideration in your system environment.

And, if you satisfy either one of these two conditions:

---

<sup>4</sup> If your media is not WORM, you can get the same effect by using the “Prevent delete from media” option.

<sup>5</sup> Retention should be specifically enabled in DiskXtender move rules. Even if you are using EMC Centera with retention capability, files saved by DiskXtender are not retained unless retention is specifically enabled in DiskXtender. Verify that retention is properly enabled before using this strategy.

- The secondary storage (NAS) is being backed up independently.
- The data on your secondary storage is not being deleted (you are using WORM media, or the data is under retention).

Since block-based backup bypasses the file system, it also bypasses DiskXtender. Therefore, the “3rd party backup mode for managed files” setting in DiskXtender is not relevant in this strategy. Irrespective of this setting, block-based backup backs up the file as it exists on the extended drive. If a file is purged, only the extended attributes are backed up. If a file is fetched, both the file data and the extended attributes are backed up. Finally, if a file is unmanaged, only the file data is backed up.

If you are using NAS storage, use the recovery procedure, as explained in [Strategy 2](#). If you are using WORM or retained storage, use the recovery procedure, as explained in [Strategy 3](#).

Block-based backup can only be restored to the same extended drive and to the same server.

The following table shows a comparison of these four strategies.

**Table 2. The four backup strategies**

	<b>Strategy 1</b>	<b>Strategy 2</b>	<b>Strategy 3</b>	<b>Strategy 4</b>
<b>3<sup>rd</sup> party backup mode for managed files' setting</b>	Full backup	Fast backup	Fast backup	--
<b>Secondary storage</b>	Any	NAS	WORM or retained	NAS, WORM, or retained
<b>Backup in relation to archive</b>	Before	After	After	After
<b>Space consumed on backup media</b>	Large	Small	Small	Small
<b>Restore to any location</b>	Yes	No	No	No
<b>File or block level</b>	File	File	File	Block

## Conclusion

We have presented four strategies to back up data managed by DiskXtender for Windows. Use any of these strategies to take a successful backup of data.