Best Practices Guide

Symantec® NetBackup® with
ExaGrid® Disk Backup with Deduplication
# Table of Contents

**EXECUTIVE SUMMARY** ................................................................................................................................. 2

**INTRODUCTION** ....................................................................................................................................................... 2

**AUDIENCE** .............................................................................................................................................................. 3

**EXAGRID BASIC CONCEPT** ........................................................................................................................................ 3

- **PRODUCT DESCRIPTION** ........................................................................................................................................ 4
- **TRADITIONAL NETBACKUP ENVIRONMENT** .................................................................................................................. 5
- **USING EXAGRID IN A NETBACKUP ENVIRONMENT** .................................................................................................... 6
- **USING EXAGRID IN NETBACKUP AND OST ENVIRONMENTS** ...................................................................................... 8

**EXAGRID TECHNOLOGY OVERVIEW** ....................................................................................................................... 10

- **SCALABLE GRID ARCHITECTURE** ............................................................................................................................... 11
- **POST-PROCESS DATA DEDUPLICATION** .......................................................................................................................... 13
- **NETWORK REQUIREMENTS** ........................................................................................................................................ 14
- **EXAGRID NIC USAGE** .................................................................................................................................................. 15
- **NETBACKUP LOAD BALANCING** .................................................................................................................................... 16

**BENEFITS OF USING EXAGRID AS A SYMANTEC NETBACKUP TARGET** .................................................................. 17

- **OST-SPECIFIC BENEFITS OF THE JOINT SOLUTION** .................................................................................................. 17
- **UNIFIED CONTROL AND TRACKING** .......................................................................................................................... 18
- **FLEXIBLE STORAGE LIFECYCLE POLICIES** .................................................................................................................. 19
- **OPTIMIZED SYNTHETIC BACKUPS** ............................................................................................................................... 21

**RECOMMENDED BEST PRACTICES** ............................................................................................................................ 23

- **BEST PRACTICES FOR EXAGRID SHARES** .................................................................................................................. 23
- **RETENTION POLICY LIMITATIONS** ............................................................................................................................... 23
- **EXAGRID HOSTNAMES BEST PRACTICES** ...................................................................................................................... 24
- **STORAGE SERVER BEST PRACTICES** ............................................................................................................................... 24
- **EXAGRID SHARE BEST PRACTICES** ............................................................................................................................... 24
- **DISK POOL BEST PRACTICES** ....................................................................................................................................... 24

**ADDITIONAL RESOURCES** ........................................................................................................................................ 24
Executive Summary

ExaGrid offers the only disk-based backup appliance with data deduplication purpose-built for backup that leverages a unique architecture optimized for performance, scalability and price. The combination of post-process deduplication, landing zone architecture and GRID scalability enables IT departments to achieve the shortest backup window and the fastest, most reliable restores and timely disaster recovery without backup window expansion or costly forklift upgrades as data grows.

This paper will focus on best practices and recommended settings, with the goal of effective integration of an ExaGrid appliance into a Symantec® NetBackup® (NBU) environment.

The use of ExaGrid appliances to replace onsite tape dramatically improves the performance of NBU environments. Some of the benefits include:

- Many more NBU backups are able to be retained in the existing storage footprint by eliminating redundant copies of backup data.
- Faster disaster recovery (DR) and data restoration can be assured by integrating ExaGrid appliances into NBU environments to replace or supplement onsite tapes for backup.
- Offloading heavy data deduplication resource loads that can slow backups from the NBU media server to a dedicated disk backup appliance with high-performance deduplication.

Using a Symantec OpenStorage (OST) enabled ExaGrid appliance and NetBackup together allows you to take advantage of several additional benefits of the combined ExaGrid-NetBackup solution, which are described in more detail in a later section of this document.

- Unified Control and Tracking is available since NetBackup tracks all copies of backup data including copies replicated to disaster recovery sites, the NBU catalog and tape copies via the NetBackup interface.
- Simplified Disaster Recovery with the ability to recover replicated backup data from an offsite ExaGrid appliance from within the NetBackup console.
- Flexible Storage Lifecycle Policies allow different data retention at local sites versus offsite disaster recovery locations by leveraging the unique ExaGrid architecture.
- Optimized Synthetics Support allows switching backup strategy to use fast, resource-friendly daily differential incremental backups plus a fast, resource-friendly daily optimized synthetic full backup, without requiring large amounts of backup data movement around the network.

Introduction

This guide provides a basic overview of how to use Symantec NetBackup to back up to an ExaGrid share. If you are familiar with Symantec NetBackup, you should be able to use the information provided in this guide to get a basic idea of how ExaGrid appliances can integrate with your existing environment. Information about the benefits of the combined solution is included. Additional information detailing specific benefits and capabilities, such as support for Optimized Synthetic backups, for those using OST is also included.

Best practices and additional recommendations are listed in the “Recommended Best Practices” section of this document for easy reference, as well as highlighted within some topics.
Audience

Symantec NetBackup customers, backup administrators, partners and others who are interested in configuration, backup best practices and recommended settings for using an ExaGrid appliance in conjunction with Symantec NetBackup, with the goal of optimal integration of an ExaGrid appliance into a Symantec NetBackup environment.

ExaGrid Basic Concept

ExaGrid appliances work seamlessly with Symantec NetBackup by presenting themselves as standard NAS shares. NetBackup jobs are simply redirected to point to the ExaGrid appliance instead of to a tape library. ExaGrid appliances are easily integrated into existing backup environments, as illustrated in Figure 1.

Figure 1 – ExaGrid Sits Behind Your Existing Backup Server and Replaces Tape Onsite or Offsite
Product Description

ExaGrid’s disk backup with deduplication system is purpose-built for backup. With ExaGrid’s GRID architecture and post-process deduplication, you get:

**Efficient Backup Storage Onsite:** ExaGrid’s post-process deduplication technology reduces backup disk space requirements from between 10:1 to 50:1.

**Fastest backups:** 30-90% faster than tape, via post-process data deduplication and full servers in a GRID

**Fastest restores:** fast onsite restore and instant VM recovery (with Veeam) from full recent backup copy maintained in high-speed cache on disk

**Integrated Replication and Disaster Recovery** – Replication of backups to offsite storage for disaster recovery purposes is fully integrated without sacrificing performance of critical backup and recovery features. WAN replication bandwidth is reduced by as much as 50:1 to a second site for data recovery if the local site is lost.

**Seamless scalability of backup system capacity:** high performance GRID architecture scales from 1 to 130TB full backup as data grows by adding full servers. This prevents backup windows from expanding, and eliminates forklift upgrades, and product obsolescence

**Fast Tape Copies** - Tapes can be created as needed for long-term offsite archival storage from backup data stored on the ExaGrid appliance.

**Support for Heterogeneous Environments** – ExaGrid appliances support CIFS for Windows and NFS for Linux and UNIX, as well as most major backup applications and specialized tools and utilities including Veeam Backup and Replication, Quest vRanger, Lightspeed, Redgate, SQLSafe, SQL Dumps, Oracle RMAN dumps, VMware backup, and direct UNIX TAR files.

Using an OST-enabled ExaGrid appliance offers several additional benefits of the combined ExaGrid-NetBackup solution, including unified control and tracking, simplified disaster recovery, flexible storage lifecycle policies and support for optimized synthetics.

For IT professionals, ExaGrid makes backup and data protection better by giving you faster, more reliable backups, cost-effective scalability as data grows without letting your backup windows expand and with no forklift upgrades, and budget protection both up-front and over time as data grows.
Traditional NetBackup Environment

Symantec NetBackup is a client–server software solution designed for enterprise customers. The core product functionality includes backup/recovery, archive/retrieval, and disaster recovery. A single console offers multi-site monitoring, analytics, and reporting, which allows customers to standardize operations and risk management.

NBU environments can be configured with a number of media servers, or one server acting as both a master and media server, managing shared or individual pools of storage units that allow for both onsite and offsite data copies to be made. Typically these copies are sent to onsite tape libraries, then to another tape sent offsite.

The traditional NetBackup environment and workflow, shown in Figure 2 above, is as follows:

1. NetBackup agent runs on each production server and connects to the NetBackup media server

2. NetBackup creates backup jobs that are managed and controlled with NetBackup’s Storage Lifecycle Policies

3. Data is cleared off the media server by migrating to a tape copy, which is kept onsite for data restore and to meet retention requirements.

4. Another tape copy is made and sent to offsite long-term storage for offsite disaster recovery.
Typical Challenges with the Traditional NetBackup Environment

The typical NBU environment scales by adding additional storage unit resources and associated media server resources to manage them. The most common challenges in NBU environments include:

- Efficient utilization of available storage resources for maximum utilization and throughput
- Extremely large client backups and increased storage costs
- Redundant data backups
- Performance bottlenecks (NBU servers, networking, tape drives)
- Large amounts of onsite and offsite tape media and the resultant management burden
- Catalog integrity, consistency and disaster recovery capabilities

Using ExaGrid in a NetBackup Environment

ExaGrid’s line of purpose-built disk-based backup with deduplication appliances is built specifically to provide an alternative to straight disk and tape volume pools in NBU, such as the offsite tape library in the backup process shown in Figure 3 below.

Compared to the traditional NetBackup environment and workflow described in the previous section, ExaGrid appliances added to the environment change the workflow as follows:

- NetBackup agent runs on the production servers with no changes, just as it did before adding ExaGrid appliances to the environment.
NetBackup running on the media server creates backup jobs that are managed and controlled with NetBackup’s Storage Lifecycle Policies. The only change here is that the jobs are mapped to NAS shares on the ExaGrid appliance.

The backup job lands onto the primary site ExaGrid appliance in full. The data deduplication process starts once the backup has landed.

If a second site is deployed only the data bytes that have changed since the last backup are replicated over the WAN to the secondary ExaGrid appliance.

When necessary to meet long-term retention requirements, a tape copy can be made and sent for offsite archiving.

**Benefits of the Integrated Solution**

- Seamless integration between NetBackup and the ExaGrid disk backup system, enabling quick installation
- Unified control and tracking of all copies of data via the NetBackup interface, including copies replicated to disaster recovery sites and tape copies
- Simplified disaster recovery, as NetBackup is aware of all copies of data – both onsite and offsite
- Flexibility to retain different amounts of backup data at a local site versus an off-site disaster recovery location, by better utilizing the unique ExaGrid architecture
Using ExaGrid in NetBackup and OST Environments

NOTE - The material in this section applies only to NetBackup environments that also have the Symantec OpenStorage capability installed and enabled. The description of the environment and workflow is very similar to, but not exactly the same as, the previous section of this document.

OpenStorage Description

Symantec Open Storage (OST) is a set of application programming interfaces (APIs) that allows ExaGrid hardware to be more tightly integrated with NetBackup. The capabilities delivered by OST enable ExaGrid to take advantage of NetBackup’s native disk features such as Media Server Load Balancing, Intelligent Disk Capacity Management and Storage Lifecycle Policies.

With the ExaGrid OST plug-in running on the media server, NBU can control the replication performed between ExaGrid appliances and can track the data movement in its catalogs. This capability enables a very powerful option for offsite replication or DR solutions. The description of the workflow is very similar, but not exactly the same as, non-OST NetBackup environments.

OpenStorage Workflow

When using Symantec Open Storage as shown in Figure 4 above:

1. NetBackup agent runs on the production server backups just as it did before adding ExaGrid appliances to the environment with no changes.
2. NetBackup running on the media server creates backup jobs that are managed and controlled with NetBackup’s Storage Lifecycle Policies. The only change here is that the jobs are mapped to NAS shares on the ExaGrid appliance.
3. The backup job lands onto the primary site ExaGrid appliance in full. The data deduplication process starts once the backup has landed. If a second site exists, only the data bytes that have changed since the last backup are replicated over the WAN to the secondary ExaGrid appliance.

Figure 4 – NetBackup Backup Data/Metadata Flow with OST Enabled ExaGrid Appliances
The NBU master server records and manages a backup image catalog. The catalog receives new metadata to reflect any changes.

When necessary to meet long-term retention requirements, a tape copy can be made and sent for offsite archiving.

**NetBackup Terminology**

Table 1 shows several of the NBU-specific terms and concepts used to describe the NBU and ExaGrid joint environment and best practices in later sections.

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Master Server</td>
<td>The primary server in an NBU environment controlling scheduling, tracking client backups, managing tape media and the NBU catalog (see below). This server can also be configured as a media server.</td>
</tr>
<tr>
<td>Media Server</td>
<td>This server manages the data movement between client and storage destination.</td>
</tr>
<tr>
<td>Catalog</td>
<td>A group of files containing vital information about performed backups. Included is information such as configuration, status, errors, what files and folders have been backed up, and tracking information on the location of the backed up data.</td>
</tr>
<tr>
<td>Full Backup</td>
<td>A backup that attempts to record all data in the set file list.</td>
</tr>
<tr>
<td>Differential</td>
<td>A backup that only records data that has changed since the last backup.</td>
</tr>
<tr>
<td>Incremental Backup</td>
<td>A backup that records data that has changed since the last full backup.</td>
</tr>
<tr>
<td>Storage Unit</td>
<td>A device where backup data is sent to be stored. This can be a library with a number of tape drives, direct attached tape drives, a configured disk pool, etc.</td>
</tr>
<tr>
<td>Storage Unit Group</td>
<td>A logical entity that pools the resources of a number of storage units for use by NBU policies.</td>
</tr>
<tr>
<td>Policy</td>
<td>The information provided to run a backup; client, file list, retention, schedule, etc.</td>
</tr>
<tr>
<td>Inline Copy</td>
<td>A backup duplicate made synchronously with the backup. This type of duplication relies on the throughput speed of the slowest Storage Unit resource.</td>
</tr>
<tr>
<td>Vault</td>
<td>An NBU agent option that allows the configuration of logical profiles designed to automate the duplication of backups, perform catalog backups, eject media, and generate reports.</td>
</tr>
<tr>
<td>Volume Pool</td>
<td>Groupings of available storage that is free for use, or is in use, by NBU backup data. Typically this refers to groups of tape media and configured disk.</td>
</tr>
<tr>
<td>High Water Mark</td>
<td>This value is the capacity level for a Disk Storage Unit at which, upon being reached, the NBU administrator is notified or staging to tape resources begins.</td>
</tr>
</tbody>
</table>

**Table 1 – Common NetBackup Terms Used in this Document**
ExaGrid Technology Overview

Figure 5- ExaGrid Appliances Connect to Form a Scalable GRID

ExaGrid appliances are comprised of Intel Quad Core XEON processors, enterprise-class SATA drives with RAID6 hot spares, and ExaGrid GRID software. Each appliance plugs into a switch and is virtualized into a shared GRID using a separate GRID connection, as shown above in Figure 5. The NBU server is connected to the same switch, and sees the appliances as one or more NAS shares. Since each appliance includes the appropriate amount of processor, memory, disk and bandwidth for the rated data size, performance increases exponentially as more appliances are added to the GRID.

**ExaGrid Appliance Models**

The product line’s multiple appliance models can be combined into a GRID configuration of up to 320TB raw capacity, allowing full backups of up to 130TB, as shown in Table 2.

<table>
<thead>
<tr>
<th>ExaGrid Model</th>
<th>Raw Capacity</th>
<th>Usable Capacity</th>
<th>Capacity for Weekly Fulls</th>
<th>Capacity for Daily Fulls</th>
<th>Backup</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Total Backup Data</td>
<td>Total Backup Data</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Copies</td>
<td>Total Backup Data</td>
<td>Copies</td>
</tr>
<tr>
<td>EX1000</td>
<td>3.5 TB</td>
<td>2 TB</td>
<td>16</td>
<td>16 TB</td>
<td>75</td>
</tr>
<tr>
<td>EX2000</td>
<td>5.5 TB</td>
<td>4 TB</td>
<td>16</td>
<td>32 TB</td>
<td>75</td>
</tr>
<tr>
<td>EX3000</td>
<td>9.0 TB</td>
<td>6 TB</td>
<td>16</td>
<td>48 TB</td>
<td>75</td>
</tr>
<tr>
<td>EX4000</td>
<td>11.0 TB</td>
<td>8 TB</td>
<td>16</td>
<td>64 TB</td>
<td>75</td>
</tr>
<tr>
<td>EX5000</td>
<td>13.0 TB</td>
<td>10 TB</td>
<td>16</td>
<td>80 TB</td>
<td>75</td>
</tr>
<tr>
<td>EX7000</td>
<td>16.0 TB</td>
<td>13 TB</td>
<td>16</td>
<td>104 TB</td>
<td>75</td>
</tr>
<tr>
<td>EX10000E</td>
<td>23.0 TB</td>
<td>20 TB</td>
<td>16</td>
<td>160 TB</td>
<td>75</td>
</tr>
<tr>
<td>EX13000E</td>
<td>32.0 TB</td>
<td>26 TB</td>
<td>16</td>
<td>208 TB</td>
<td>75</td>
</tr>
<tr>
<td>EX7000SEC</td>
<td>16.0 TB</td>
<td>13 TB</td>
<td>16</td>
<td>104 TB</td>
<td>75</td>
</tr>
<tr>
<td>EX10000SEC</td>
<td>23.0 TB</td>
<td>20 TB</td>
<td>16</td>
<td>160 TB</td>
<td>75</td>
</tr>
<tr>
<td>EX13000SEC</td>
<td>32.0 TB</td>
<td>26 TB</td>
<td>16</td>
<td>208 TB</td>
<td>75</td>
</tr>
</tbody>
</table>

Table 2 – Wide Range of Available Backup Data Capacities
Scalable GRID Architecture

The typical business or organization is seeing data growth of 30% a year, which presents challenges to IT to ensure their backup system can scale easily to support that growth.

With disk backup solutions that have a front-end server architecture and add just disk shelves as data grows, you start with a shorter backup window, but as data grows the backup windows expand because you only add disk and are not adding more deduplication processing resources. Eventually, the backup window expands to a point where you must replace the front-end server with a more powerful server via a costly “forklift upgrade.” In contrast, ExaGrid’s scalable GRID-based approach adds full servers—including memory, processor, disk, and bandwidth. Figure 6 shows the differences between how the two different architectures cope with data growth over time.

This unique approach provides the following benefits:

- **No Expansion of Backup Windows as Data Grows** – By adding full servers, ExaGrid maintains consistently fast backup performance and a fixed length backup window as data increases.

- **Seamless Scalability with No Forklift Upgrades** – ExaGrid’s scalable GRID-based approach avoids forklift upgrades by adding modular capacity with full servers to the existing system in increments of 1, 2, 3, 4, 5, 7, 10, or 13TB.

- **No Obsolescence of Previous Model Systems** – All ExaGrid appliances can be seamlessly added to any existing ExaGrid deployment and can be mixed and matched with older appliances of any size, thereby eliminating obsolescence. Adding a new ExaGrid appliance is as simple as plugging it in and pointing backup jobs to the new NAS share.

![Figure 6 - Scalable GRID Architecture Maintains Backup Windows as Data Grows](image.png)
ExaGrid disk backup appliances include GRID computing software which allows them to virtualize and share data storage capacity with one another when plugged into a switch. This GRID system shown above in Figure 7 can expand as your data grows by adding appliances to the GRID, providing up to 320TB raw capacity and allowing full backups up to 130TB.

**Figure 7 - ExaGrid GRID Architecture Expands by Adding Nodes to Increase Capacity as Data Grows**
Post-Process Data Deduplication

The ExaGrid disk backup appliance uses post-processing to perform its deduplication. As shown in Figure 8 above, the backup data is written directly from the backup server to ExaGrid's disk landing zone at the highest possible rate with no inline processing to interfere, resulting in the smallest possible backup window. Once the backup job is complete and off the network, the data is protected and immediately available for restore or tape copy. Then the appliance deduplicates and simultaneously replicates the data in the background.

Because ExaGrid's product allows each full backup to first land on the landing zone, it caches that most recent backup for rapid restore. Since over 90% of restores are done from the most recent backup, this approach avoids overhead incurred from “rehydrating” data during critical restores. As a result, restore times from an ExaGrid are two to three times faster than competitive solutions that do not preserve a complete copy of the most recent backup and store only deduplicated data.

ExaGrid's innovative approach minimizes the amount of data to be stored by using byte-level data deduplication across all received backups. ExaGrid's byte-level delta technology stores only the changed bytes from backup to backup instead of storing full copies. This unique approach reduces the disk space required by at least 10:1, and up to 50:1, delivering unparalleled performance for the fastest backups and restores for about the same cost as a new tape library.
Network Requirements

This section details the network requirements for ExaGrid sites and ExaGrid appliances. As shown above in Figure 9, you may connect all NICs of all ExaGrid appliances in an ExaGrid site to the same subnet. The internal ExaGrid network (NIC1) over which ExaGrid appliances communicate can share the same subnet as the other ExaGrid NICs over which backup data flows.

Each ExaGrid site must be on its own subnet. Depending upon the model, ExaGrid appliances come with two, four or six network interface connectors (NICs). Appliances with multiple NICs allow backup data from multiple sources to be written to one or more shares.
ExaGrid NIC Usage

ExaGrid appliances use their NIC1 port, shown in Figure 10 above, to communicate with other ExaGrid appliances within an ExaGrid site and with ExaGrid appliances in other ExaGrid sites, as shown in Table 3 below. Single ExaGrid appliance sites do not use their NIC1 port and can be left unconnected. ExaGrid appliances within an ExaGrid site must connect their NIC1 ports to the same subnet.

<table>
<thead>
<tr>
<th>NIC</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>NIC1</td>
<td>Used in communication between ExaGrid appliances. Not used in single server, single site systems.</td>
</tr>
<tr>
<td>NIC2</td>
<td>Used as the primary network connection for all ExaGrid appliance models. Over this connection flows: backups, notifications, and browser sessions.</td>
</tr>
<tr>
<td>NIC3, NIC4, NIC5 and NIC6</td>
<td>Additional NICs available on some ExaGrid appliance models. Multiple NICs allow you to increase your backup throughput.</td>
</tr>
</tbody>
</table>

Table 3 - ExaGrid NIC Usage

ExaGrid appliances within one ExaGrid site cannot connect their NIC1 ports to the same subnet used by NIC1 ports of ExaGrid appliances in another ExaGrid site. Failure to observe this requirement may cause loss of access to data. NIC1 cannot be used to accept backups from your backup application.
NetBackup Load Balancing

Symantec’s OpenStorage allows NetBackup to automatically balance the load across multiple media servers in a group based on existing server load conditions (for example, CPU and memory utilization). If an ExaGrid appliance has multiple NICs, NetBackup will direct traffic to its least busy NIC, as shown in Figure 11 above.

1. NetBackup selects which media server to use for a job based on the server with the lowest NetBackup application load.

2. The NetBackup media server communicates with the ExaGrid appliance and selects the NIC through which the backup will be written based on the lowest network load.

Figure 11 – NetBackup Jobs Sent to the ExaGrid Appliance are Load Balanced and Link Balanced
Benefits of Using ExaGrid as a Symantec NetBackup Target

Efficient disk-based backup requires close integration between the backup software and the disk device. That is the advantage delivered by the partnership between Symantec and ExaGrid. Together, Symantec and ExaGrid provide a cost-effective disk-based backup solution that scales to meet the needs of demanding enterprise environments. Symantec NBU users typically take only a few minutes to configure the system and are fully operational within as little as 60 minutes.

OST-Specific Benefits of the Joint Solution

The close integration of ExaGrid appliances, NetBackup and Symantec’s OpenStorage (OST) provides a seamless solution, as shown above in Figure 12. Unified control and tracking of all copies of data is done via the NetBackup server processes, with results shown in the NetBackup user interface, including copies replicated by your ExaGrid appliance to a second ExaGrid site. NetBackup environments with ExaGrid and OpenStorage benefit from:

1. **Unified Control and Tracking** of all copies of data via the NetBackup user interface, including copies replicated by your ExaGrid appliance to a second ExaGrid site.

2. **Flexible Storage Lifecycle Policies** with the ability to specify different retention schemes for data written to the primary ExaGrid site and the data replicated to a second ExaGrid site.

3. **Simplified and Efficient Disaster Recovery** since NetBackup can maintain two copies of a catalog; one for the data written to the primary ExaGrid site and another for the data replicated by the ExaGrid appliance to a second ExaGrid site. The time required for catalog and inventory operations is virtually eliminated.

4. **Optimized Synthetic Backups** that are fully supported by ExaGrid appliances, which have been certified by Symantec for use with NetBackup.

Figure 12 – ExaGrid Brings Specific Benefits to OST-enabled NetBackup Environments
Unified Control and Tracking

Unified control and tracking of all copies of data is done via the NetBackup server processes, using Storage Lifecycle Policies (SLP) and OST as shown in Figure 13 above. Results are shown in the NetBackup user interface, including copies of backup data that are replicated by your ExaGrid appliance to a second ExaGrid site.

The ability to control and track replicated data at a second ExaGrid site is especially useful in the following situations:

- **Disaster Recovery Scenarios** - Because NetBackup is “aware” of the data at the second site, time is saved during NetBackup’s inventory and catalog phases.

- **Performing Test Recovery Operations** - Proving to yourself, members of your organization or interested third parties that a second copy of your data resides at a remote location and is recoverable can be done easily and quickly via the NetBackup user interface without interrupting your normal backup operations.

- **Monitoring Replication from One ExaGrid Site to Another** - The ExaGrid user interface provides detailed reports on deduplication and replication. NetBackup’s Activity Monitor also reports the status of ExaGrid replication. OpenStorage-enabled jobs of type “Duplicate” are jobs replicated by your ExaGrid appliance to a second ExaGrid site.

Duplicate jobs are queued until the ExaGrid appliance completes deduplicating and replicating the source share’s backup data. If replication takes longer than 12 hours, the NetBackup duplicate job will fail, but if the job is part of a Storage Lifecycle Policy it will automatically restart.
Flexible Storage Lifecycle Policies

Storage Lifecycle Policies Overview

Storage Lifecycle Policies (SLPs) are a mechanism which controls the creation and retention of multiple copies of the same backup data providing a plan or map of where backup data will be stored and for how long. As shown in Figure 14 above, the SLP determines the locations where the backup is initially written and the destinations where it is subsequently duplicated to as well as the period of time that each copy of the backup will be retained. It also automates the replication process, making SLPs an ideal technology for implementing standardized protection policies. Storage Lifecycle Policies ensure backup data always exists at the appropriate locations at the appropriate phases of the lifecycle.

Using Storage Lifecycle Policies allows you to specify different retention periods for the initial backup and for the duplicate copies. For example, you might specify one retention period for the original local backup and another for a duplicate at a disaster recovery site.

In this configuration, the following activities take place:

1. Backup jobs are created by the NetBackup media server and backed up to the ExaGrid appliance.

2. The ExaGrid appliance deduplicates the backup data, greatly reducing the amount of disk space needed to store the backup data.

3. The backup image is then replicated from the ExaGrid spoke to the ExaGrid hub, managed via optimized duplication through NetBackup's Storage Lifecycle Policies and/or manual duplication jobs. Only the bytes that have changed since the last backup are copied to the hub, taking advantage of ExaGrid's 50:1 WAN efficiency. The catalog on the NetBackup Master server (not pictured) is also updated, so that the two backup instances can be managed and reported on via NetBackup, and can be restored from either location.
Simplified and Efficient Disaster Recovery

Disaster Recovery Overview

NetBackup sees ExaGrid OST-enabled appliances as disk, enabling features such as intelligent capacity management, media server load balancing, reporting, storage lifecycle policies and optimized replication. Without OST, NetBackup media servers have to manage all duplicate backup copies, which means that data must be transferred across the LAN, WAN, or SAN from the primary site secondary storage to a NetBackup media server and then to the disaster recovery site storage medium.

NetBackup Catalog Replication and Recovery

NetBackup can maintain two copies of the NBU catalog for disaster recovery, as shown in Figure 15 above. Since the NetBackup catalog is aware of all copies of data that are retained, recovery of data from a NetBackup OST-optimized duplicate copy is the same as recovery from another duplicate. The potential time savings over recovery from a non-OST-optimized duplicate could be significant, and the time required for catalog and inventory operations is virtually eliminated.

1. One copy of the catalog is created by the NetBackup Master NDMP Host and the NetBackup media server that contains the metadata for the data written to the primary ExaGrid site (Copy 1)

2. A second copy (Copy 2) of the catalog can be created by an offsite NDMP host for the data that is replicated to an ExaGrid appliance.

3. The catalog copies are kept up-to-date by NetBackup. Through NetBackup's Backup-Archive-Restore GUI, the OST-optimized duplicate copy can be designated as the primary copy and then a full or granular recovery can be initiated.
Optimized Synthetic Backups

Optimized Synthetics Overview

Optimized synthetics use metadata to create a roadmap for recovering a full backup without actually doing the data movement to rebuild the full. Optimized synthetic cumulative incremental backups allow a low-cost cumulative incremental backup to be synthesized from a series of previously-done differential incremental backups.

Optimized Synthetics Concept and Workflow

The optimization provided by the manipulation of metadata alone allows backup administrators to drastically reduce the frequency of those time- and resource-costly full backups. Instead, they can perform many days of fast, resource-friendly differential incremental backups plus a fast, resource-friendly daily optimized synthetic full backup. This concept is illustrated above in Figure 16 above.

1. The optimized synthetic full is the accumulation of all the metadata from the Day 1 full backup plus Days 2-5 differential incremental backups of just the changed data.

2. The various pieces of protected data are tracked via metadata, and these are used to optimize the construction of the synthetic full or the cumulative incremental.

3. No data is moved when constructing the synthetic full. Only the references to the actual data are manipulated.

4. Full recovery from that optimized synthetic full backup leverages the metadata references to bring back only the data required.
Benefits of Optimized Synthetics

- **Reduce Backup Windows from Elimination of Full Backups** - Weekly full backups are no longer required. The efficiency of metadata allows a single full backup to be followed by many, many days of fast, resource-friendly incremental backups.

- **Greater Reduction of Backup Windows** - The effort and time required to optimize a synthetic full is reduced to only the manipulation of metadata by the backup application and advanced disk-based appliance. No longer must all the organization’s data be shuffled around to create a synthetic full – only the metadata needs to be processed.

- **Reduced Backup Application Load** - Synthetic full backups carry a natural benefit of reducing the backup application load on all the organization’s clients. With optimized synthetic full backups, the backup application load is further reduced from having to move all the data around over and over to merely processing the metadata, which is only a fraction of the organization’s actual data.

- **Improve both RTOs and RPOs** - Because optimized synthetic full backups are so much cheaper to create than non-optimized (traditional) synthetic full backups, backup application policies can generate them more frequently, improving RPOs. When the need arises to restore an organization’s data, there’s a recent optimized synthetic full ready to go. Via a single backup application operation, the advanced appliance uses its metadata to return exactly the data required, reducing the overall RTO.

- **Conserve Replication Bandwidth** - Advanced disk-based backup appliances that provide integrated replication capabilities only need to replicate the daily changes in data and the metadata associated with the synthetic full backups, which is a fraction of the organization’s data protection/replication needs.
Recommended Best Practices

Best Practices for ExaGrid Shares

Before you can use NetBackup to back up to an ExaGrid appliance, you must first create an ExaGrid share. You will then use NetBackup to write to the ExaGrid share. Experienced NetBackup users will be able to use the information provided in this guide and build more complex backup schemes. ExaGrid recommends that you first identify your backup needs and then create a backup plan based on those needs.

How many shares to create, which ExaGrid appliances will host which shares and what data will be backed up to a share must be part of your overall backup strategy.

- For systems with multiple ExaGrid appliances, create shares on ExaGrid appliances in such a way that the backup load is evenly distributed across ExaGrid appliances.
- You can have a maximum of 20 shares per ExaGrid appliance. Typical customers will create far less than that maximum.
- You can have a maximum of 80 shares per ExaGrid site. Typical customers will create far less than that maximum.
- Use unique share names across your ExaGrid appliance. You must use unique share names on each ExaGrid appliance.
- For hub sites, consider replicating all shares to the same spoke. Doing so will make disaster recovery easier and faster.

In addition to the above best practices, ExaGrid recommends using unique share names across your ExaGrid appliance. Doing so will make ExaGrid reports easier to interpret, avoid share name conflicts in disaster recovery scenarios and avoid share name conflicts if you decide to migrate a share to another ExaGrid appliance.

- If this is a multi-site configuration, you may specify an ExaGrid site to which this share will replicate its data. Doing so puts an additional copy of the share’s contents on another site, which provides an additional layer of protection for your data.
- If the share is being replicated to another ExaGrid site, you may select a specific ExaGrid appliance on the other ExaGrid site on which the share will be created. In multi-server ExaGrid sites, you may balance the processing load across several ExaGrid appliances.
- If your backup application supports OpenStorage, ExaGrid shares allow you to take full advantage of OpenStorage functionality. Some devices may use more than one IP address. You must include all IPs/names of all devices (physical or virtual) that will write to the share.

Retention Policy Limitations

OST gives administrators the ability to set a different retention policy on the local copy of data than the replicated copy. ExaGrid supports OST’s ability to specify a longer time horizon for the off-site retention, but does not support having a longer local retention and a shorter off-site. For example, you can set a 4-week retention locally and 52 week remotely, but you cannot set 52 weeks locally and 4 remotely.
ExaGrid Hostnames Best Practices

- Use the assigned fully-qualified hostname.
- Use host files to convert hostnames to IP addresses.
- While generally reliable, naming services such as DNS may unintentionally introduce processing delays or error conditions into the data protection infrastructure.
- When possible, avoid creating secondary hostnames to associate with alternate IP interfaces.

Storage Server Best Practices

- Have only one storage server defined per ExaGrid site. Use hostnames when configuring storage servers.
- Do not use IP addresses in place of hostnames as this will limit the options to route optimized duplication traffic.
- Keep the name of the storage server unique across the enterprise.

ExaGrid Share Best Practices

- For systems with multiple ExaGrid appliances, create shares on ExaGrid appliances in such a way that the backup load is evenly distributed across ExaGrid appliances.
- Create no more than 20 primary shares per ExaGrid appliance.
- Create no more than 80 primary shares per ExaGrid site.
- Use unique share names across your ExaGrid appliance. You must use unique share names on each ExaGrid appliance.
- For hub sites, consider replicating all shares.

Disk Pool Best Practices

- Have only one disk pool per Storage Unit.
- Devise a naming convention to easily distinguish between Disk Pools used by the ExaGrid source shares. For example, append “-Source” to a Disk Pool’s name.
- For NetBackup 7.1 and later, ensure that the Limit I/O streams option is unchecked.
- Use the native capabilities of NetBackup whenever possible

Additional Resources

Detailed procedures and additional information on any of the topics covered in this document may be found in these publications available from ExaGrid.

- “ExaGrid Disaster Recovery Guide”
- “Using Symantec NetBackup with an ExaGrid System”