

# EMC BACKUP AND RECOVERY WITH DEDUPLICATION FOR SAP

EMC Data Domain, EMC NetWorker, EMC Symmetrix VMAX, and VMware vSphere

- Replicate data offsite and meet recovery SLAs
- Simplify complex offsite backup processes

## EMC SOLUTIONS GROUP

### Abstract

This white paper describes how to enhance data protection for SAP systems across data centers with EMC® Data Domain® replication technology. This paper also discusses best practices for administration simplicity and efficiency by using the EMC NetWorker® Module for SAP with Oracle.

December 2011



Copyright © 2011 EMC Corporation. All Rights Reserved.

EMC believes the information in this publication is accurate as of its publication date. The information is subject to change without notice.

The information in this publication is provided “as is.” EMC Corporation makes no representations or warranties of any kind with respect to the information in this publication, and specifically disclaims implied warranties of merchantability or fitness for a particular purpose.

Use, copying, and distribution of any EMC software described in this publication requires an applicable software license.

For the most up-to-date listing of EMC product names, see EMC Corporation Trademarks on EMC.com.

VMware, ESX, VMware vCenter, and VMware vSphere are registered trademarks or trademarks of VMware, Inc. in the United States and/or other jurisdictions.

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

Part Number H8876

# Table of contents

|   |           |
|---|-----------|
| <b>Executive summary .....</b>                              | <b>5</b>  |
| Business case.....  | 5         |
| Solution overview .....                                     | 5         |
| Key results .....   | 6         |
| <b>Introduction .....</b>                                   | <b>7</b>  |
| Purpose .....   | 7         |
| Scope .....   | 7         |
| Audience .....  | 7         |
| Terminology .....   | 7         |
| <b>Key components .....</b>                                 | <b>9</b>  |
| Overview.....   | 9         |
| EMC Data Domain deduplication storage systems.....          | 9         |
| EMC Data Domain Boost software .....                        | 9         |
| EMC Data Domain Replicator software .....                   | 9         |
| EMC Symmetrix TimeFinder/Clone.....                         | 9         |
| EMC NetWorker .....   | 10        |
| EMC NetWorker Module for SAP with Oracle .....              | 10        |
| VMware vSphere .....  | 10        |
| SAP ERP IDES .....  | 10        |
| <b>Solution architecture.....</b>                           | <b>11</b> |
| Solution overview .....                                     | 11        |
| Hardware resources .....                                    | 12        |
| Virtual resources.....                                      | 12        |
| Software resources .....                                    | 13        |
| <b>Design and configuration .....</b>                       | <b>14</b> |
| Introduction .....  | 14        |
| Solution design overview .....                              | 14        |
| Backup process.....   | 15        |
| Recovery process .....                                      | 16        |
| Key configurations .....                                    | 17        |
| SAP storage design .....                                    | 17        |
| Offloading the backup with Symmetrix TimeFinder/Clone ..... | 18        |
| Configuring DD Boost.....                                   | 18        |
| Configuring DD Replicator .....                             | 21        |
| Configuring NetWorker .....                                 | 22        |
| Configuring NMSAP.....                                      | 24        |

|   |           |
|---|-----------|
| Configuring DD Boost.....                     | 25        |
| <b>Testing and validation .....</b>           | <b>26</b> |
| Test overview .....                           | 26        |
| SAP nonproduction scenarios.....              | 26        |
| SAP production scenarios.....                 | 26        |
| Test results .....                            | 27        |
| Deduplication ratio analysis .....            | 27        |
| RTO breakdown analysis .....                  | 28        |
| <b>Conclusion .....</b>                       | <b>30</b> |
| Summary .....                                 | 30        |
| Findings.....                                 | 30        |
| <b>References.....</b>                        | <b>31</b> |
| White papers .....                            | 31        |
| Product documentation.....                    | 31        |
| Other documentation .....                     | 31        |
| <b>Additional information .....</b>           | <b>32</b> |
| Overview.....                                 | 32        |
| Backup utility profile.....                   | 32        |
| Unmount VMDKs from the backup server .....    | 32        |
| Mount cloned VMDKs to the backup server ..... | 32        |
| Technical recovery solutions .....            | 32        |

## Executive summary

**Business case** SAP customers are facing many challenges trying to meet their backup and recovery requirements:

- Enterprise data is growing at accelerated rates, causing management and maintenance issues.
- Customers prefer a daily full database backup instead of a daily incremental one after the initial full backup, because of the complexity to restore an incremental backup.
- Governance may require keeping backup copies at offsite locations.
- End-user productivity depends on the availability of data.
- Backup windows keep shrinking due to increasingly high availability requirements, 24x7 global operations, and the need to minimize impact on business operations.
- Service-level agreements (SLAs) are becoming increasingly aggressive, requiring a quick restore of the production environment, even if the backup of this environment is stored offsite.

These challenges demand a solution that offers effective, affordable, and efficient data protection for critical SAP business operations.

**Solution overview** This solution demonstrates how customers can:

- Consolidate SAP backups on a remote site with enhanced data protection
- Quickly and efficiently restore the backup data from a remote site if the local backup is not usable

Enterprises running SAP can address these challenges by deploying EMC® NetWorker® with EMC NetWorker Module for SAP with Oracle (NMSAP), together with EMC Data Domain® and Data Domain Boost (DD Boost).

This solution significantly:

- Improves backup performance
- Reduces backup duration
- Optimizes network bandwidth utilization
- Reduces the backup footprint with deduplication
- Automates and centralizes the backup and recovery operations

To demonstrate the benefits and validate the integration, the solution uses two data centers: the SAP production system located at the primary site and the development and test systems located at the remote site. One Data Domain storage system resides at each site. Production system backups are replicated to the remote site through NetWorker clone-controlled replication using DD Boost file replication. All SAP servers are built on fully virtualized VMware® virtual machines, to facilitate management, resource allocation, and dynamic workload balance.

## Key results

This solution:

- **Reduces backup storage cost:** with a 27:1 deduplication ratio after 10 backups provided by Data Domain, eliminating the use of tapes, their manual handling, and their offsite storage services required. The deduplication ratio can be much higher as the backup accumulates.
- **Accelerates backup and restore operations:** with integration of NetWorker automation and Data Domain Boost when compared with the traditional tape shipping solution.
- **Eliminates performance impact:** by removing the overhead on the SAP production system, backing up from a clone copy mounted to a virtual host.

# Introduction

## Purpose

The purpose of this white paper is to demonstrate how SAP customers can:

- Reduce backup windows when backing up SAP data at the primary site by using NetWorker and Data Domain deduplication
- Reduce network requirements when replicating the deduplicated backup data from the primary site to a remote site
- Ensure SAP data availability with recoverability from the remote site when the SAP backup data no longer exists on the primary site

## Scope

This white paper covers the following topics:

- An overview of the technologies involved in this solution
- The solution reference architecture
- The technical design and configuration
- The test procedures used and their results
- The solution benefits

## Audience

This white paper is intended for customers, partners, and EMC employees, including IT planners, SAP architects, Basis administrators, and any other personnel involved in evaluating, acquiring, managing, operating, or designing an SAP landscape infrastructure.

## Terminology

This paper includes the following terminology.

**Table 1. Terms used in the white paper**

| Term         | Definition  |
|--------------|---|
| BACKINT      | The external backup tool, called by BR*Tools, which performs the SAP system backup at the file system level. This tool is SAP certified.                      |
| Storage node | A host computer (including the NetWorker server) that has the physical connection and ownership of the attached storage devices.                              |
| DEV          | The SAP development system.   |
| QAS          | The SAP quality assurance system, which is also referred to as a test system, and is usually copied from the SAP production system.                           |
| PRD          | The SAP production system.  |
| RTO          | Recovery Time Objective (RTO) is the maximum acceptable amount of time for restoring an SAP system after a failure or disaster.                               |
| BR*Tools     | A toolset provided by SAP for Oracle database administration tasks.   |
| ESX™         | VMware hypervisor architecture, built on compatible physical servers, which provide the foundation for building and managing a virtualized IT infrastructure. |

| Term            | Definition  |
|-----------------|---|
| Virtual machine | A tightly isolated software container that can run its own operating systems and applications as if it were a physical computer.  |
| VMDK            | Virtual machine disk files, a datastore descriptor.   |
| VMFS            | VMware Virtual Machine File System, which is a cluster file system that enables multiple VMware ESX hosts to concurrently access a single VMDK file.  |
| RDM             | Raw device mapping, which is a mapping file in a VMFS volume that acts as a proxy for a raw physical device. It offers VMFS manageability with the raw device access required by certain applications.  |
| NMSAP           | EMC NetWorker Module for SAP with Oracle.   |
| DDR             | Data Domain Replicator, which is EMC Data Domain replication technology that transfers only the deduplicated and compressed unique changes across an IP network. DDR requires a tiny fraction of the bandwidth, time, and cost required by traditional replication methods. |

## Key components

### Overview

This section briefly describes the following key solution components.

- EMC Data Domain deduplication storage systems
- EMC Data Domain Boost software
- EMC Data Domain Replicator software
- EMC Symmetrix® TimeFinder®/Clone
- EMC NetWorker
- EMC NetWorker Module for SAP with Oracle
- VMware vSphere®
- SAP ERP IDES

### EMC Data Domain deduplication storage systems

EMC Data Domain, high-speed, inline data deduplication storage systems dramatically reduce the amount of disk storage needed to retain and protect enterprise data. By identifying redundant data segments before they are stored, Data Domain systems require a storage footprint that is 10 to 30 times smaller, on average, than the original dataset. Backup data can then be efficiently replicated and retrieved over existing networks for streamlined disaster recovery and consolidated tape operations.

### EMC Data Domain Boost software

EMC Data Domain Boost (DD Boost) extends the optimization capabilities of Data Domain solutions. DD Boost significantly increases performance by distributing parts of the deduplication process to the backup server or application clients. This simplifies disaster recovery procedures, and serves as a solid foundation for additional integration between applications and Data Domain systems.

### EMC Data Domain Replicator software

EMC Data Domain Replicator software provides simple, fast, robust WAN-based disaster recovery for the enterprise. This software offers comprehensive flexibility for a variety of topologies in the distributed enterprise, from remote offices at the edge to large core data centers. Unlike other deduplication methods, Data Domain deduplication is inline, so replication completes as fast as possible to minimize risk and maximize currency of the restore point.

### EMC Symmetrix TimeFinder/Clone

EMC Symmetrix TimeFinder/Clone provides point-in-time copies of devices that can be used for backups, decision support, data warehouse refreshers, or any other process that requires parallel access to production data. TimeFinder/Clone allows you to make copies of data simultaneously on multiple target devices from a single source device. The data is copied from the source device creating a physical backup copy called a clone. The data is available to a target device's host immediately on activation of the TimeFinder copy session, even if the copy process has not completed.

## EMC NetWorker

EMC NetWorker backup and recovery software centralizes, automates, and accelerates data backup and recovery across your IT environment. NetWorker delivers centralized backup and recovery operations for complete control of data protection across diverse computing and storage environments.

## EMC NetWorker Module for SAP with Oracle

The EMC NetWorker Module for SAP with Oracle (NMSAP) is an SAP-certified add-on module for NetWorker and provides a backup and restore interface between the SAP BR\*Tool programs and the NetWorker server. BACKINT, which is part of NMSAP, provides the interface between BR\*Tools and NetWorker.

## VMware vSphere

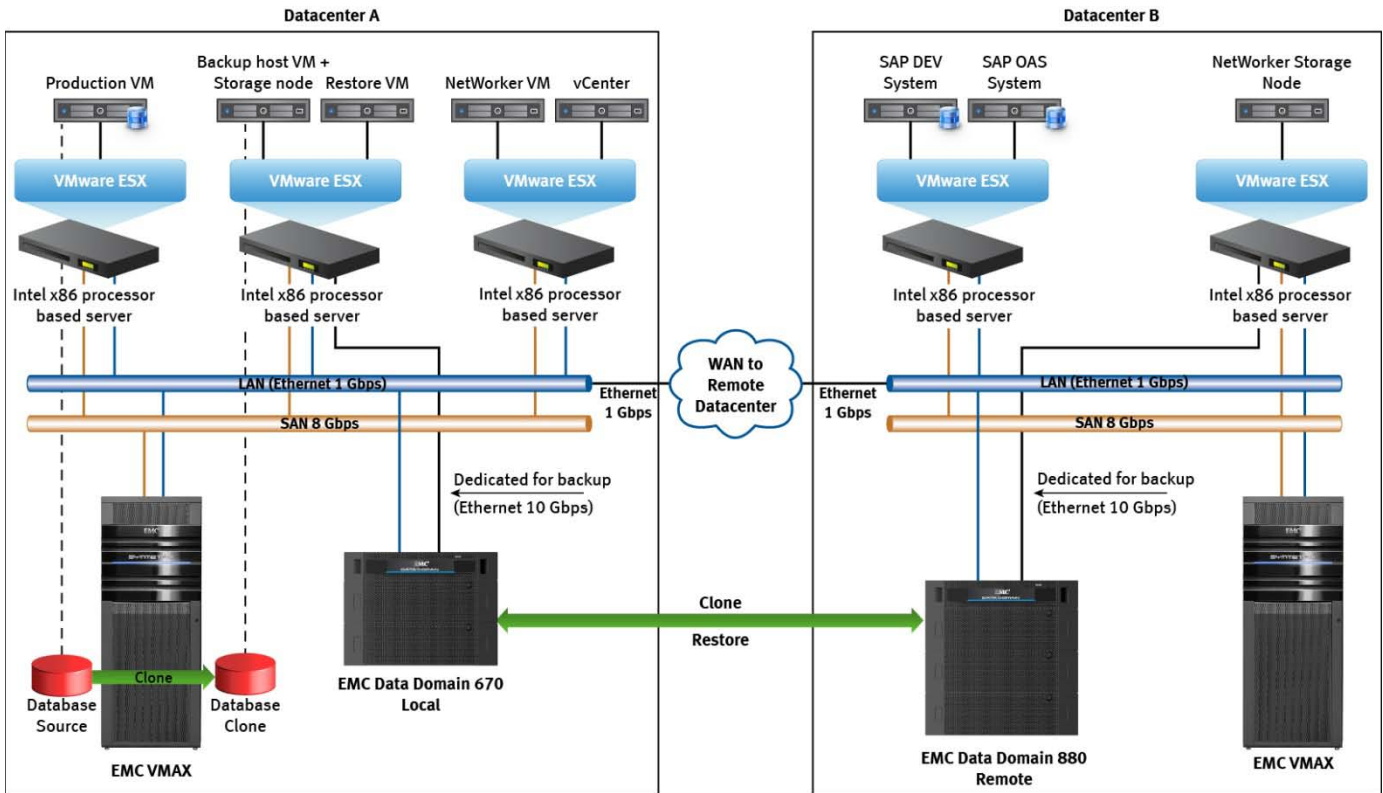
VMware vSphere is the foundation for a dynamic, self-optimizing IT infrastructure. It is a robust, production-proven virtualization layer that abstracts processor, memory, storage, and networking resources into multiple virtual machines. VMware vSphere allows enterprises to dramatically reduce hardware and operating costs by sharing resources across a virtual environment. The components of VMware vSphere include ESX, ESXi, and vCenter™.

## SAP ERP IDES

SAP ERP IDES is the *Internet Demonstration and Evaluation System* in the SAP ERP Central Component (ECC) system and represents a model company. It consists of an international group with subsidiaries in several countries. IDES contains application data for various business scenarios that can be run in SAP systems. The business processes in the IDES system are designed to reflect real-life business requirements and have access to many realistic characteristics. IDES uses easy-to-follow business scenarios to show the comprehensive functions of the ECC system. The focal point of IDES, however, is not the functionality itself, but the business processes and their integration.

# Solution architecture

**Solution overview** Figure 1 shows the high-level architecture of the solution environment.



GEN-001807

**Figure 1. Architecture overview**

This solution uses a two-data-center configuration to demonstrate and validate the efficiency of data replication across data centers as illustrated in Figure 1. The SAP production system is located at the primary site (Datacenter A), while the SAP nonproduction systems are located at the remote site (Datacenter B). In addition, due to the increasing deployment of cloud computing, all SAP application servers are built on fully virtualized VMware virtual machines to take advantage of the virtualized environment. This reduces costs, increases efficiency, and provides a high level of reliability.

To eliminate the impact on the production system during the backup, a clone copy of the production system is created and mounted onto a backup host with the NetWorker storage node installed.

A dedicated virtual machine (Restore VM) is created to verify the restore of the data backed up without destroying the production system.

The administration network is configured to run with 1 Gb/s bandwidth. Ten Gb/s network connectivity is configured between the storage node and Data Domain storage systems to maximize the backup performance, as recommended by EMC. Refer to the *EMC NetWorker Data Domain Deduplication Devices Release 7.6 Service Pack 2 Integration Guide* for more information.

## Hardware resources

Table 2 details the hardware resources for the solution.

**Table 2. Hardware resources**

| Equipment                                  | Quantity | Configuration  |
|--|----------|--|
| EMC Symmetrix VMAX™                        | 1        | Array Module: VMAX-1<br>Memory: 240 GB<br>Device pool: 240 x 300 GB FC 15K |
| EMC Data Domain duplication storage system | 2        | DD670 Data Domain<br>DD880 Data Domain                                     |
| VMware ESX server                          | 4        | CPU Type: 4 quad-core Intel Xeon X5570 2.26 GHZ<br>Memory: 132 GB          |

## Virtual resources

Table 3 details the virtual resources for the solution.

**Table 3. Virtual resources**

| Equipment            | Quantity | Configuration          |
|----------------------|----------|------------------------|
| EMC NetWorker server | 1        | 4 x vCPU, 8 GB memory  |
| SAP DEV host         | 1        | 4 x vCPU, 16 GB memory |
| SAP QAS host         | 1        | 4 x vCPU, 16 GB memory |
| SAP PRD host         | 1        | 8 x vCPU, 32 GB Memory |
| SAP Mount host       | 1        | 8 x vCPU, 32 GB memory |
| SAP PRD restore host | 1        | 8 x vCPU, 32 GB memory |
| Storage node server  | 1        | 4 x vCPU, 6 GB memory  |

**Software resources** Table 4 presents the software resources for the solution.

**Table 4. Software resources**

| Software                                     | Version                                 |
|--|---|
| Windows Server 2008                          | Enterprise SP2                          |
| Oracle Database                              | 11.2.0.2 64-bit X64 for Windows         |
| SAP ERP                                      | 6.04/NetWeaver 7.01 IDES 64-bit Unicode |
| EMC NetWorker                                | 7.6 SP2                                 |
| NetWorker Module for SAP with Oracle (NMSAP) | 4.2                                     |
| VMAX Enginuity™                              | 5875.198.148                            |
| EMC Data Domain operating system version     | 5.0.1.0-240411                          |
| VMware vSphere™                              | 4.1 U1                                  |
| EMC Solutions Enabler                        | V7.3.0.1                                |

# Design and configuration

## Introduction

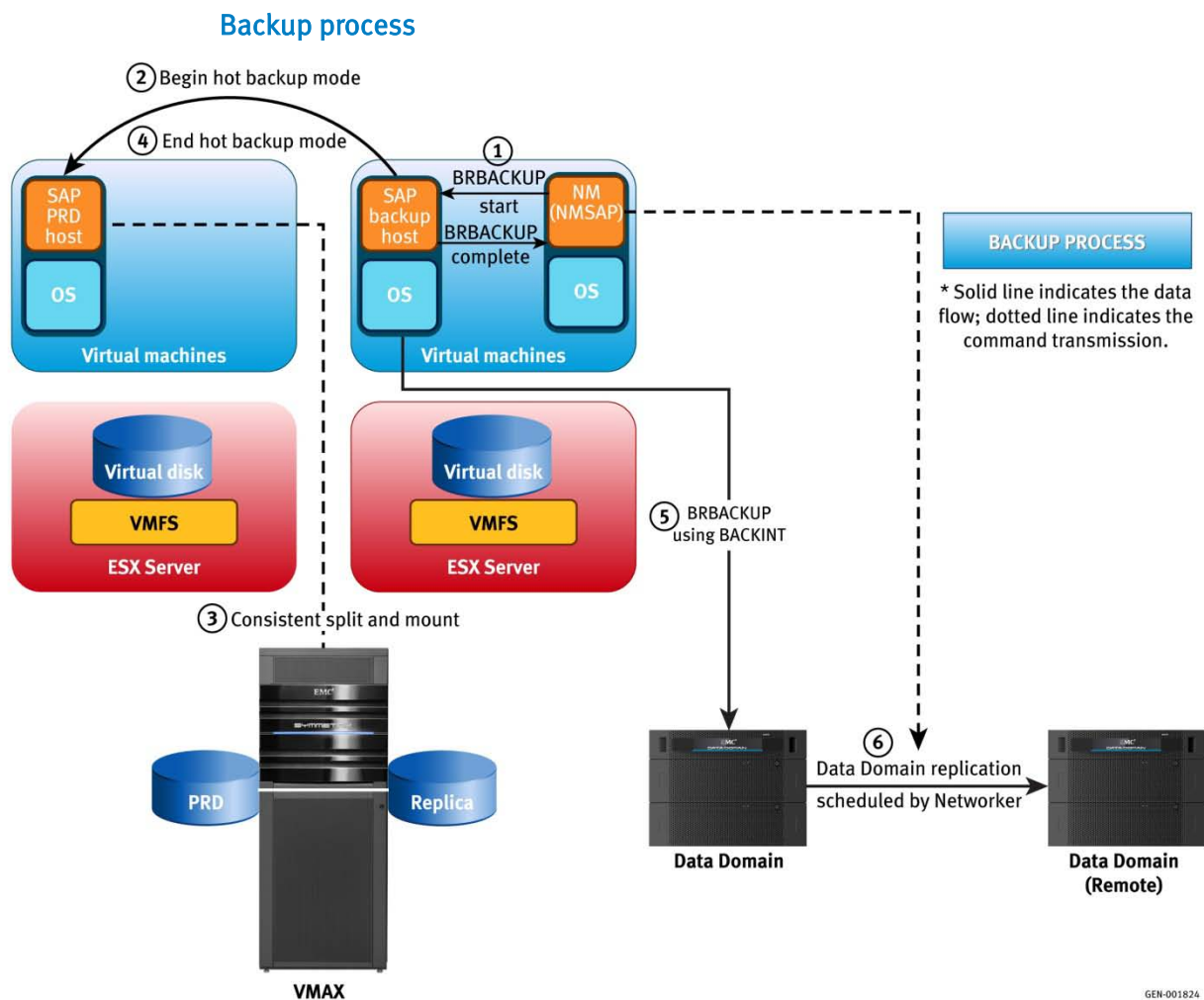
This section provides the configuration information for this solution including:

- Solution design overview
- Key configurations
- SAP storage design
- Offloading the backup with Symmetrix TimeFinder Clone
- Software components configuration, including:
  - DD Boost
  - DD Replicator
  - NetWorker
  - NMSAP

## Solution design overview

The goals of the backup process design are as follows:

- Integrate with customers' current backup strategy
- Provide nondisruptive backups
- Use storage subsystem clone and SAP-aware backup software capability
- Reduce backup storage footprint with deduplication storage systems
- Improve RTO with backup on remote storage system instead of offsite tapes



**Figure 2. Backup process**

The following steps briefly describe the backup process as illustrated in Figure 2.

1. NetWorker initiates the backup on the mount host by calling BRBACKUP.
2. BRBACKUP interacts with the SAP production host virtual machine and puts the database into hot backup mode.
3. BRBACKUP calls a script to create a clone copy of the SAP production system, including disk clone, split, and mount/unmount operations on the backup host.
4. BRBACKUP interacts with the SAP production server again, and takes the database out of hot backup mode.
5. BRBACKUP triggers a call to BACKINT and starts the backup operation from the TimeFinder clone mounted in the backup host virtual machine.
6. Data Domain Replicator replicates the backup data to the remote site for offsite protection. This is done under the control of NetWorker for future data retrieval for recovery.

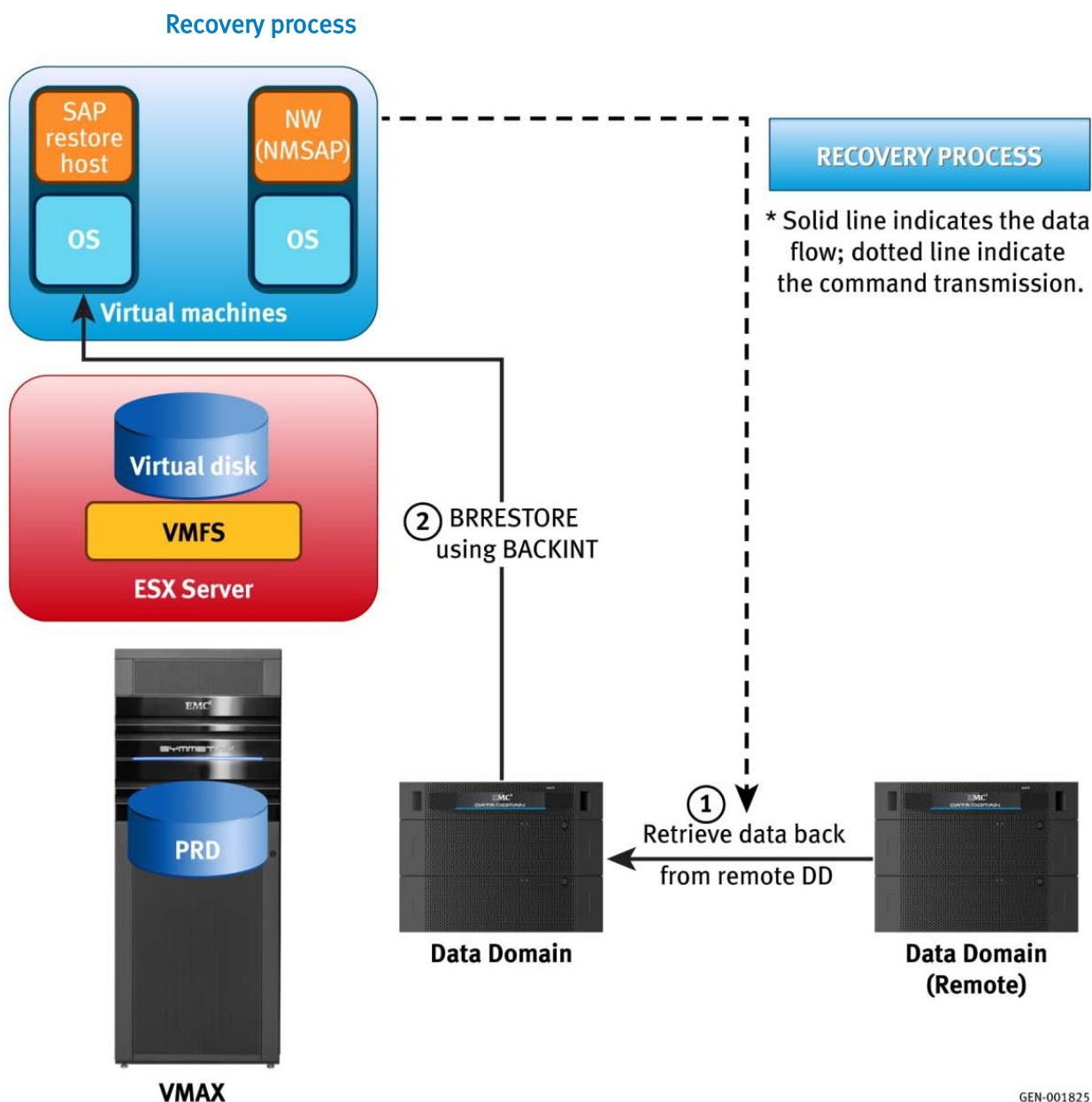


Figure 3. Recovery process

The following steps briefly describe the process of a restore from the remote site to a virtual machine host on the production site as illustrated in Figure 3.

1. The data is missing from the local Data Domain system.
2. From the EMC NetWorker Management Console a request is created to replicate back the data. Data Domain automatically identifies whether all the data required is available and transmits only the unique portions of data that are missing from the remote Data Domain.
3. BRRESTORE calls NetWorker to complete the recovery operation. The commands is as follows:

```
brrestore -u / -b <detail_backup_logfile> -m full
```

```
for example brrestore -u / -b behcqaaf.aff -m full
```

The SAP systems presented in this solution follow the SAP standard recommendation of three-system landscape shown as DEV, QAS, and PRD.

## Key configurations

To configure this solution, the testing team carried out the following procedures:

- Installed all SAP systems and application servers with VMFS on the VMware virtualized platform.
- Set a two-day backup retention period on the local Data Domain system and a three-month retention period on the remote Data Domain system—for lab validation purposes and to enable the backup sets to be replicated to the remote Data Domain system. In general, the remote Data Domain system was used to consolidate all backups from different sources across different locations.
- Created a clone copy of the SAP PRD database for backup and mounted it in the backup host virtual machine to eliminate the performance impact on the SAP production environment during the backup.
- Used NMSAP for integration with SAP BR\*Tools.
- Enabled DD Boost.
- Enabled NetWorker multiple data streams during the backup.
- Used the automated load software to generate additional SAP transaction data between each full backup.
- Backed up the production system archive logs every two hours, between the full backup to the Data Domain system, to minimize the data loss.

For more information about backup with multiplexing, refer to the *EMC NetWorker Release 7.6 Service Pack 2 Administration Guide*.

## SAP storage design

Table 5 describes the storage design for all the SAP systems built on virtual machines.

**Table 5. Storage design pattern for SAP systems**

| RAID type | Disk type | LUN type | LUN size | Purpose             | Filesystem   |
|-----------|-----------|----------|----------|---------------------|--|
| RAID 1    | VMDK      | Thin LUN | 90 GB    | Windows OS          | C:   |
| RAID 1    | VMDK      | Thin LUN | 32 GB    | SAP Binary          | E:\usr\sap   |
| RAID 1    | VMDK      | Thin LUN | 32 GB    | Oracle Binary       | F:\oracle\<sid>                                      |
| RAID 1    | VMDK      | Thin LUN | 32 GB    | Oracle Mirror log   | G:\oracle\<sid>\mirrlogA<br>G:\oracle\<sid>\mirrlogB |
| RAID 1    | VMDK      | Thin LUN | 32 GB    | Oracle Original log | H:\oracle\<sid>\origlogA<br>H:\oracle\<sid>\origlogB |
| RAID 1    | VMDK      | Thin LUN | 64 GB    | Oracle Archive log  | I:\oracle\<sid>\oraarch                              |

| RAID type   | Disk type | LUN type | LUN size | Purpose                 | Filesystem   |
|-------------|-----------|----------|----------|-------------------------|--|
| RAID 1      | VMDK      | Thin LUN | 16 GB    | Oracle trace and others | P:\oracle\<sid>\saparch<br>P:\oracle\<sid>\sapbackup<br>P:\oracle\<sid>\sapcheck<br>P:\oracle\<sid>\sapreorg<br>P:\oracle\<sid>\saptrace |
| RAID 5(3+1) | VMDK      | Thin LUN | 128 GB   | SAP data 1              | J:\oracle\<sid>\sapdata1   |
| RAID 5(3+1) | VMDK      | Thin LUN | 128 GB   | SAP data 2              | K:\oracle\<sid>\sapdata2   |
| RAID 5(3+1) | VMDK      | Thin LUN | 128 GB   | SAP data 3              | L:\oracle\<sid>\sapdata3   |
| RAID 5(3+1) | VMDK      | Thin LUN | 128 GB   | SAP data 4              | M:\oracle\<sid>\sapdata4   |
| RAID 5(3+1) | VMDK      | Thin LUN | 128 GB   | SAP data 5              | N:\oracle\<sid>\sapdata5   |
| RAID 5(3+1) | VMDK      | Thin LUN | 128 GB   | SAP data 6              | O:\oracle\<sid>\sapdata6   |

### Offloading the backup with Symmetrix TimeFinder/Clone

To reduce impact on the production during the backup process, a recoverable clone copy of the production database was created and mounted onto the backup host for backup purpose.

In this test scenario, we used Symmetrix Solutions Enabler to create a consistent Symmetrix TimeFinder/Clone and used vSphere PowerCLI to mount and unmount VMDK to the backup host virtual machine. This is needed before or after cloning in order to map VMDK files in a VMware environment. Symmetrix Solutions Enabler and vSphere PowerCLI are recommended to be installed on the backup host.

To create a consistent Symmetrix TimeFinder clone, refer to the *EMC Symmetrix TimeFinder Product Guide*.

For more information about how to use vSphere PowerCLI for mounting and unmounting VMDK in a VMware environment, refer to the sample scripts provided in the [Additional information](#) section on page 32.

### Configuring DD Boost

With DD Boost enabled, the Data Domain system can achieve better network utilization performance while deduplicating the data during backup. DD Boost saves network bandwidth by transmitting only the unique data segments from the backup client to the Data Domain systems.

To enable DD Boost, the Data Domain system and NMSAP need to be configured accordingly.

To configure DD Boost on the Data Domain system, perform the following steps:

1. Check the DD Boost license in EMC Data Domain Enterprise Manager as shown in Figure 4.

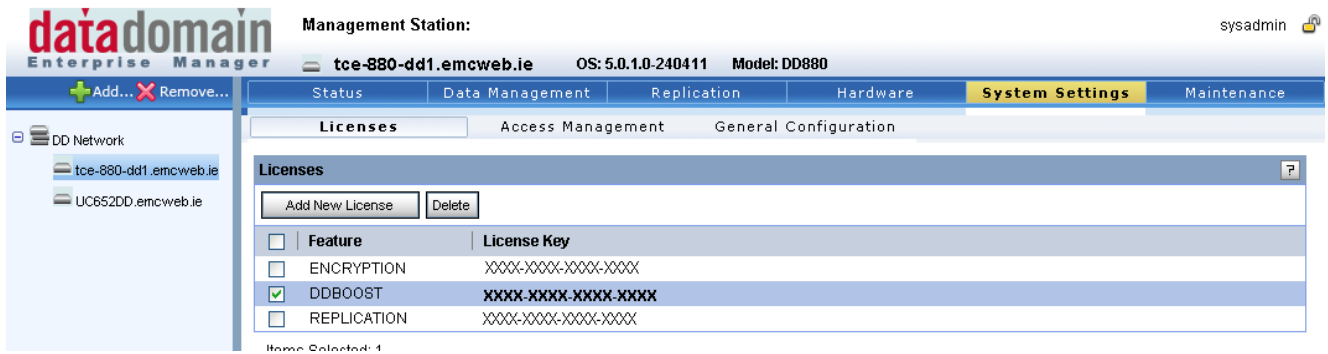


Figure 4. DD Boost license verification

2. Create the backup user and specify the DD Boost username as shown in Figure 5.

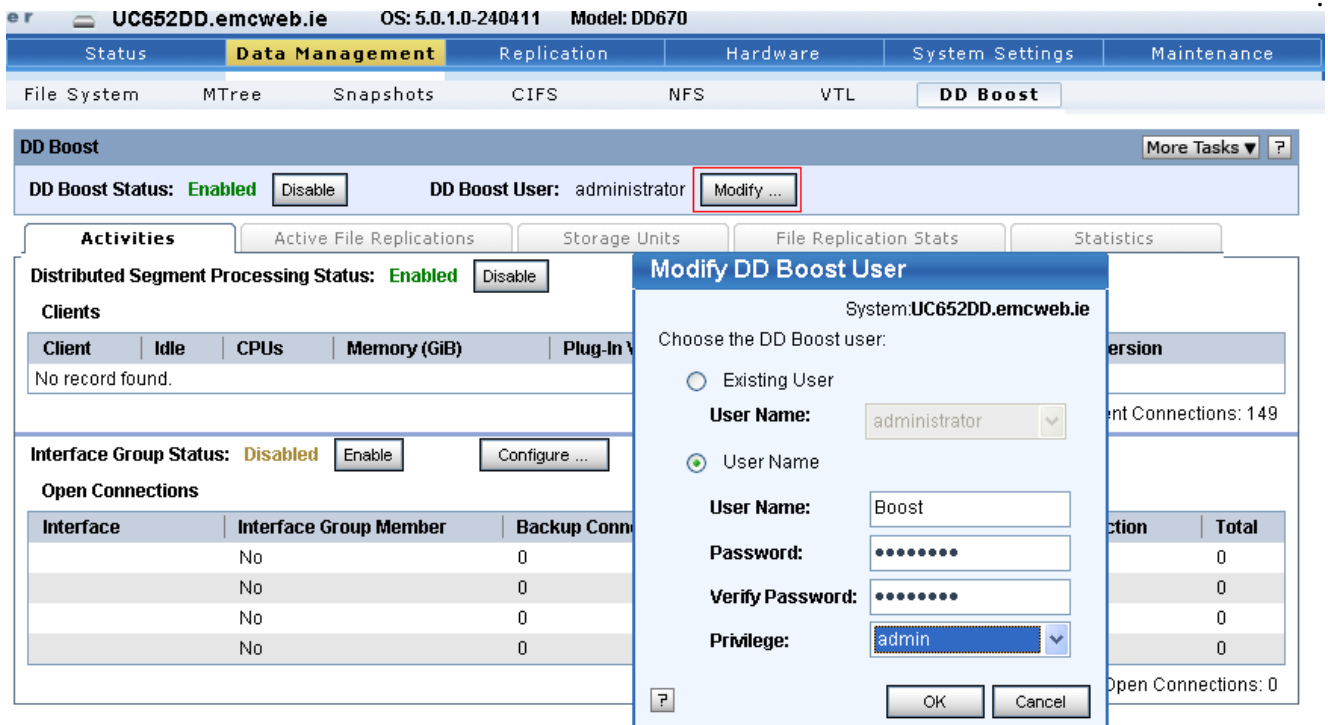


Figure 5. Enter the DD Boost user name and password

3. Ensure that the Data Domain file systems and NFS services are enabled as shown in Figure 6 and Figure 7.

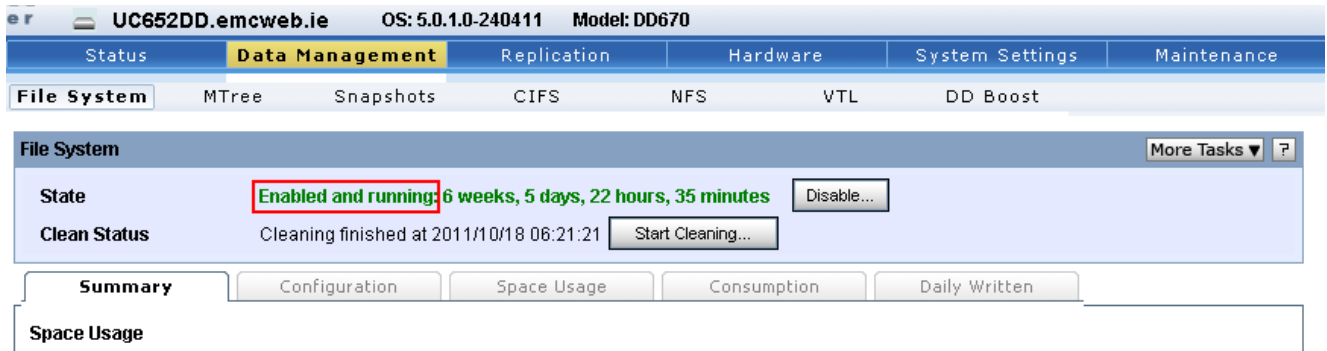


Figure 6. Data Domain File System service status

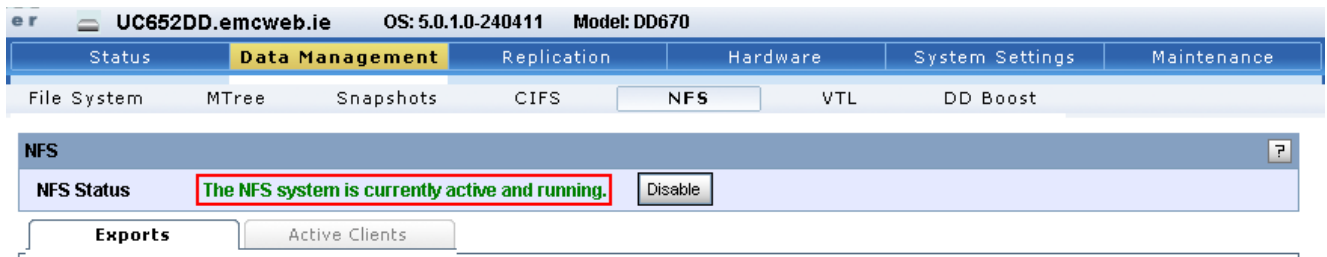


Figure 7. Data Domain NFS service status

4. Ensure that DD Boost is enabled as shown in Figure 8.

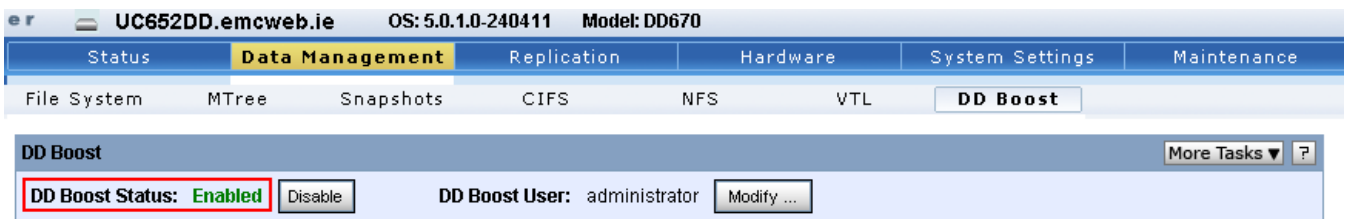


Figure 8. DD Boost service status

5. Create the storage unit as shown in Figure 9.

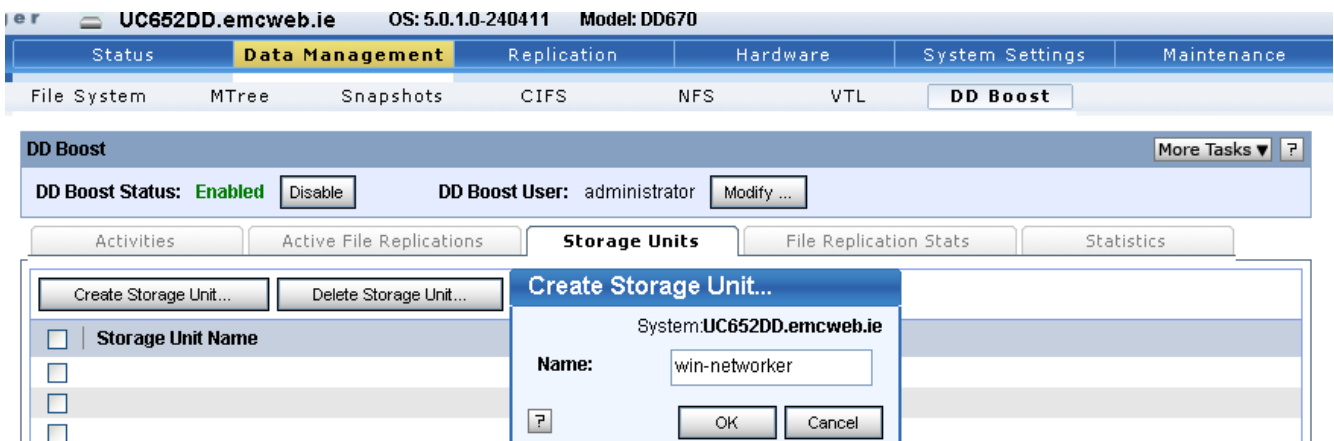


Figure 9. DD Boost storage unit creation

For detailed configuration information, see the *DD OS 5.0.1 Administration Guide* and the *EMC NetWorker Data Domain Deduplication Devices Release 7.6 Service Pack 2 Integration Guide*. Refer to the [Configuring NMSAP](#) section for the detailed setting on NMSAP.

## Configuring DD Replicator

Data Domain Replicator enables copy of SAP data from a local Data Domain device to a remote one to enhance protection. This SAP data replication is possible through the use of the NetWorker clone feature. The NetWorker clone is a complete and independent copy of the data, which can be used for data recovery or further clone creation.

Configure the clone for immediate use, or schedule NetWorker to create the clone later depending on the environment and storage needs. EMC recommends that you use the scheduling method for long-distance remote data centers. Creating an immediate clone can extend the backup time and may create a WAN bandwidth utilization conflict.

To configure a scheduled Data Domain clone, perform the following steps on the EMC NetWorker Management Console:

1. In the **Media** view, create a target media pool on which the clones are to be stored. Select **Backup Clone** for the **Pool type** field and **Data Domain** for the **Media type required** field as shown in Figure 10.

The screenshot shows the 'Create Media Pool' dialog box. The 'Basic' tab is selected. Under 'Identity', the 'Name' field contains 'prdclone'. The 'Target Media' section has 'Media type required' set to 'Data Domain' and 'Media type preferred' set to an empty dropdown. The 'Configuration' section has 'Enabled' checked, 'Pool type' set to 'Backup Clone', and 'Label template' set to an empty dropdown. Orange boxes highlight the 'Data Domain' dropdown and the 'Backup Clone' dropdown.

Figure 10. Create a Data Domain media pool

2. In the **Device** view, run **New Device Wizard** to create the Data Domain clone devices.
3. Specify the media pool created in Step 1 for the Data Domain clone devices.
4. Create a scheduled clone job in the **Configuration** view. Specify the clone pool created in step 1 as the **Write clone data to pool** attribute. Select **Clone save sets that match selections** to limit the save sets to be cloned as shown in Figure 11.

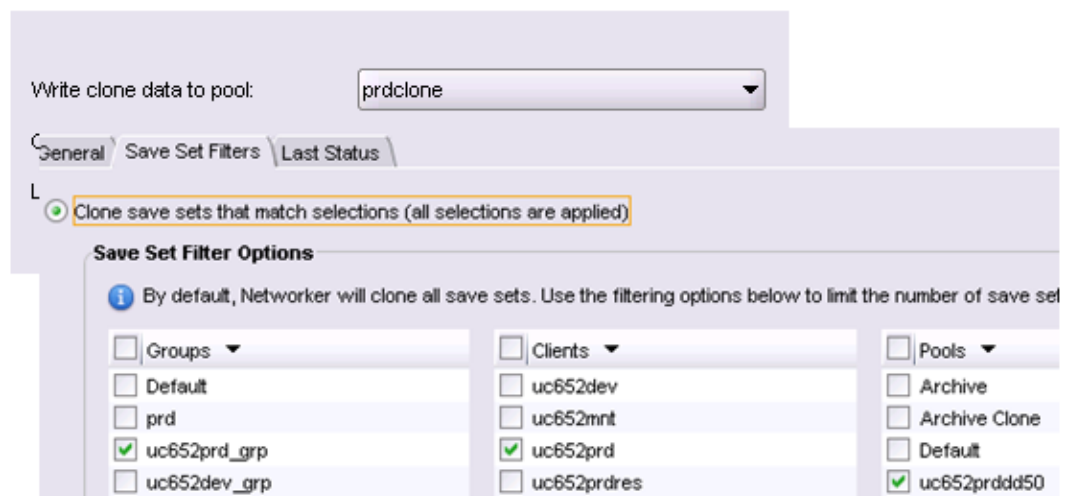


Figure 11. Create the Data Domain Replicator clone job

5. Review the Data Domain scheduled clone job in the **Monitoring** view as shown in Figure 12.

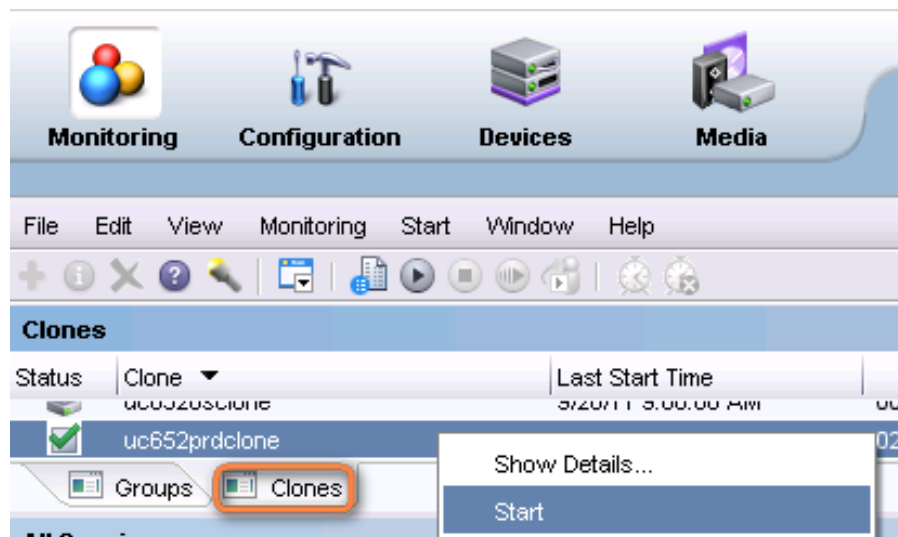


Figure 12. Start and monitor the Data Domain clone job

For more configuration information about how NetWorker integrates with Data Domain systems, refer to the *EMC NetWorker Data Domain Deduplication Devices Integration Guide*.

### Configuring NetWorker

Complete the NetWorker general configuration tasks by following the steps in the *EMC NetWorker Release 7.6 Service Pack 2 Administration Guide*. General tasks include creating:

- Label Template
- Media Pool
- Device
- Group and Client resources

The SAP client resource on the backup host requires special settings for the **Save set** name and the **backup command** as shown in Figure 13 and Figure 14, as part of NMSAP configuration.

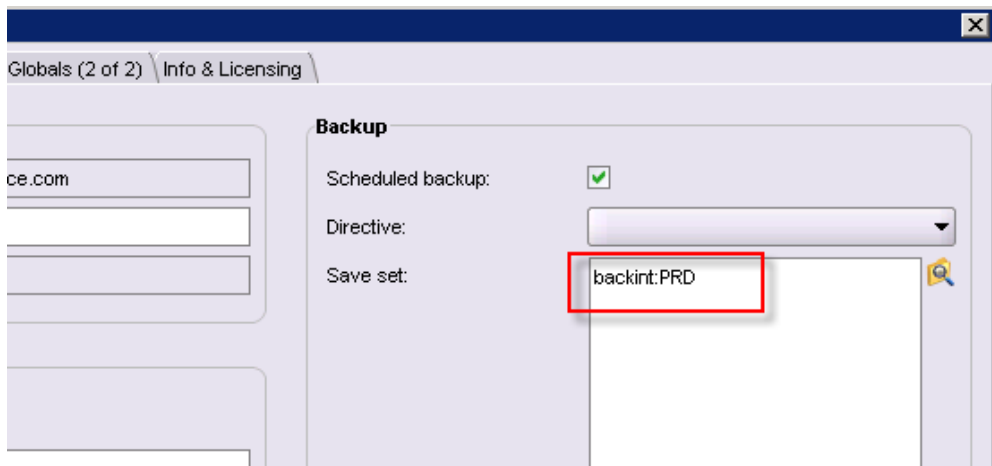


Figure 13. Set the Save set on the SAP client

The **Save set** name formatted as **backint:<sid>** indicates the backup type handled by NMSAP as shown in Figure 13.

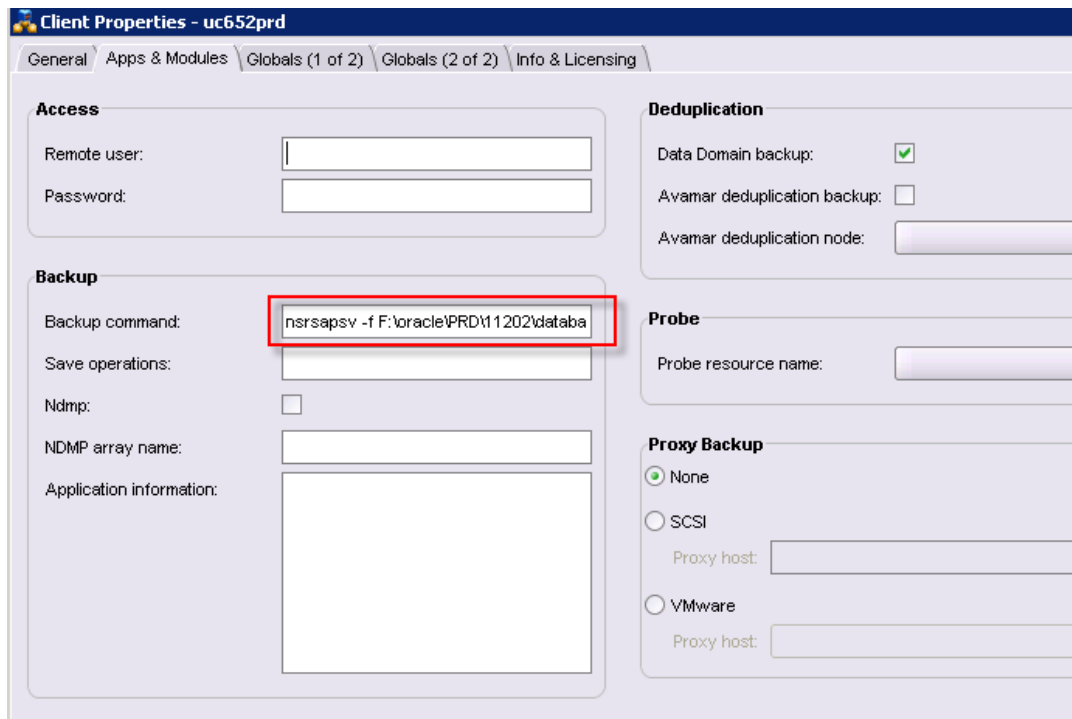


Figure 14. Set the Backup command on the SAP client

The full **Backup command** shown in Figure 14 is:

```
nsrsapsv -f F:\oracle\PRD\11202\database\nsrsapsv.cfg
```

This command specifies which configuration file NMSAP should call when the backup begins. When you specify the backup command, a manual configuration process is applied to achieve flexible control over the backup of SAP systems. In addition to the manual backup command, you can also use the NetWorker wizard to easily configure backup parameters for SAP systems.

For more information about the backup command profile, refer to the *EMC NetWorker Module for SAP with Oracle Release 4.2 Administration Guide*. For backup with the Data Domain system, you must create and bind the Data Domain device to the relevant pool associated with Net Worker group resources.

## Configuring NMSAP

NMSAP uses the SAP configuration file on the backup host to specify how and where to back up the SAP system as listed in Table 6.

**Table 6. NMSAP configuration files**

| Parameter                         | Description   |
|-----------------------------------|---|
| backup_type = online_split        | This option is only applied to a split mirror online backup.              |
| backup_dev_type=util_file_online  | Specify the NetWorker utility file for this parameter.                    |
| split_cmd=sample_clone_script.bat | This is a script to split the clone copy and mount it to the backup host. |
| Backup_mode=all_data              | Select the data to backup.  |
| Util_par_file=init<SID>.util      | This is the NetWorker configuration file.                                 |

Table 7 describes the BR\*Tools parameters.

**Table 7. BR\*Tools parameters**

| Parameter        | Description  |
|------------------|--|
| util_file_online | Use external backup programs for file-by-file backup. Use this parameter value for an online backup if it is supported by the external backup program. This value dynamically sets and ends the backup status of the tablespaces to be saved and greatly reduces the volume of the offline redo log files during the backup. |
| online_split     | Perform the splitting and saving of the mirror disks in the split command scenario when the database is open. This option is only applied to a split mirror online backup.   |
| all_data         | Back up the files of all tablespaces except pure index tablespaces or empty tablespaces.   |

The **util\_file\_online** parameter must work along with the backup utility profile, as demonstrated in the [Additional information](#) section on page 32.

For more information about using BR\*Tools to configure NMSAP, refer to the *EMC NetWorker Module for SAP with Oracle Release 4.2 Administration Guide*.

## Configuring DD Boost

The parameters in the backup utility profile are specific to DD Boost as shown in Table 8.

**Table 8. DD Boost configuration file**

| Profile characteristic | Value |
|------------------------|-------|
| NSR_DIRECT_ACCESS      | Yes   |

The parameter NSR\_DIRECT\_ACCESS determines the location where duplicate data is computed and the data transmission path.

In this solution, direct access mode is activated, which causes the data to be deduplicated on the backup client directly without storage node engagement. The deduplicated data is then transmitted from the SAP backup client to the Data Domain system directly, reducing the backup time significantly.

**Note** For more information about the explanation of backup utility profile, refer to the *EMC NetWorker Module for SAP with Oracle Release 4.2 Administration Guide*.

# Testing and validation

## Test overview

To validate this solution, the testing team:

- Backed up the production SAP database at the primary Data Domain site, and backed up the nonproduction SAP database at the remote Data Domain site with NetWorker and Data Domain.
- Replicated the deduplicated production backup data from the primary site to a remote site.
- Recovered the production SAP data from the remote site when the SAP backup data was no longer available at the primary site due to expiration.

The testing team captured the following performance matrix for comparison:

- Deduplication ratio (database size/unique data stored on DD system)
- Backup and restore job duration

## SAP nonproduction scenarios

To complete the SAP nonproduction system test at the remote site, the testing team:

1. Performed online full backup of the SAP systems to Data Domain using NetWorker.
2. Destroyed the SAP system database.
3. Restored the SAP systems from the backup created in step 1.

## SAP production scenarios

To complete the backup and recovery test with Data Domain Replicator, the testing team:

1. Performed the initial backup for PRD using NetWorker.
2. Replicated the deduplicated backup data from the primary site to the remote site enabled by clone-controlled replication with integration of Data Domain and NetWorker.
3. Launched the automated load software to generate one percent of the SAP system data.
4. Performed the PRD backup to the local Data Domain system and replicated it to the remote site.
5. Repeated steps 3 and 4.
6. Expired the initial backup on the local Data Domain system and restored the PRD system from the backup on the remote site.

## Note:

- The SAP production data was restored to a dedicated virtual machine. The test conducted was to demonstrate a complete recover of the PRD environment for re-purposing or compliance reasons.
- The automated load software simulated changes created by business activities that took place in the SAP PRD database.

- Benchmark results are highly dependent upon workload, specific application requirements, and system design and implementation. Relative system performance will vary as a result of these and other factors. Therefore, this workload should not be used as a substitute for a specific customer application benchmark when critical capacity planning and/or product evaluation decisions are contemplated.
- All performance data contained in this report was obtained in a rigorously controlled environment. Results obtained in other operating environments may vary significantly.
- EMC Corporation does not warrant or represent that a user can or will achieve similar performance expressed in transactions per minute.

## Test results

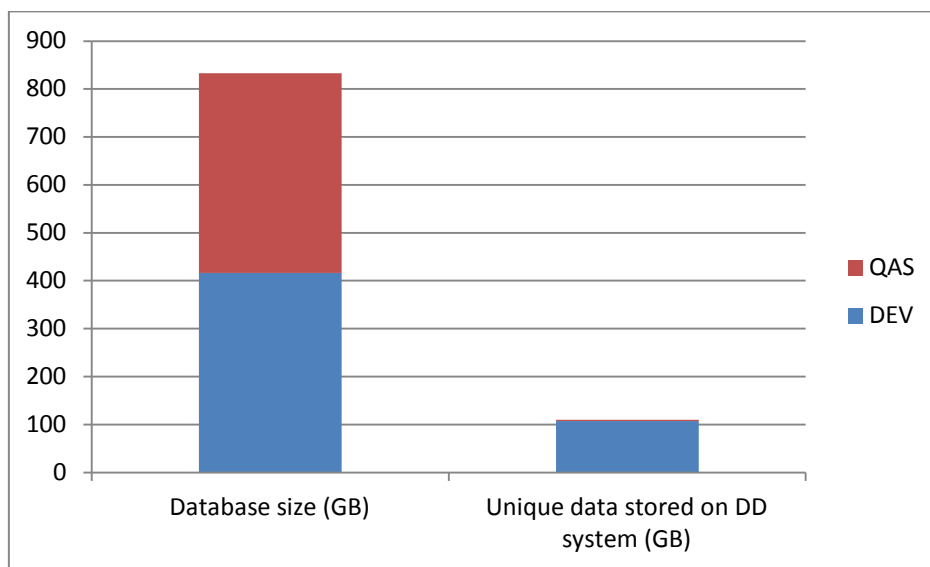
### Deduplication ratio analysis

**Table 9. Nonproduction deduplication ratio statistics in the Data Domain test**

| Backup sequence | Deduplication ratio analysis | Database size (GB) | Unique data stored on DD system (GB) | DD deduplication ratio |
|-----------------|------------------------------|--------------------|--------------------------------------|------------------------|
| 1               | DEV                          | 416                | 107                                  | 3.88                   |
| 2               | QAS                          | 417                | 3                                    | 7.58                   |

As shown in Table 9, the storage requirement of the two systems in a traditional backup scenario is approximately 833 GB. However, inline deduplication and compression significantly reduces the real space used in the Data Domain system to accommodate the 833 GB data set from 833 GB to 110 GB. This reflects a 7.58:1 deduplication ratio. The ratio can be much higher if you run sequent backups. This test scenario achieved a 27:1 deduplication ratio after 10 backups.

The test results demonstrate the great benefit of the inline deduplication provided by the Data Domain systems regarding storage requirements.



**Figure 15. Storage footprint comparison**

In the Data Domain system test, backup data physically written to the Data Domain system (as shown in the **Unique data stored on DD system** column in Table 9) is much smaller than the logical size of the backup content (as shown in the **Database size** column). Only unique data was transferred across the network during the backup testing performed using Data Domain with DD Boost, which was a great improvement in terms of LAN bandwidth requirements. It also reduces the storage required to keep the backed up information, thus allowing the retention period to be much longer, and raising the probability of a faster recovery when compared to traditional methods.

**Table 10. Production deduplication ratio statistics in the Data Domain test**

| SID | BRS sequence        | Backup duration (mm:ss) | Database size (GB) | Unique data stored on DD system (GB) | Local/remote DD deduplication ratio |
|-----|---------------------|-------------------------|--------------------|--------------------------------------|-------------------------------------|
| PRD | Initial backup      | 30:44                   | 454.31             | 133.24                               | 3.7                                 |
| PRD | Subsequent backup 1 | 19:12                   | 460.48             | 3.70                                 | 7.2                                 |
| PRD | Subsequent backup 2 | 18:18                   | 467.85             | 1.28                                 | 10.7                                |

As shown in Table 10, the production system test reflects the consistent deduplication result as the nonproduction systems.

- The subsequent full backup time is reduced by 40 percent compared with the initial full backup.
- The subsequent replication between Data Domain systems is shortened by 28 percent compared with the initial replication.
- When the backup on the local Data Domain system expires, the BR\*Tools automatically retrieves the backup stored on the remote Data Domain system with no manual intervention.

### RTO breakdown analysis

**Table 11. RTO breakdown comparison in different scenarios**

| Scenario                  | Locate media name to be restored | Retrieve media and prepare for restore | Restore from the media | Total time |
|---------------------------|----------------------------------|--|------------------------|------------|
| Tape offsite              | 0.5 hours                        | 4 hours                                | 1.5 hours              | 6 hours    |
| Remote Data Domain system | N/A                              | N/A                                    | 1.2 hours              | 1.2 hours  |

### Notes

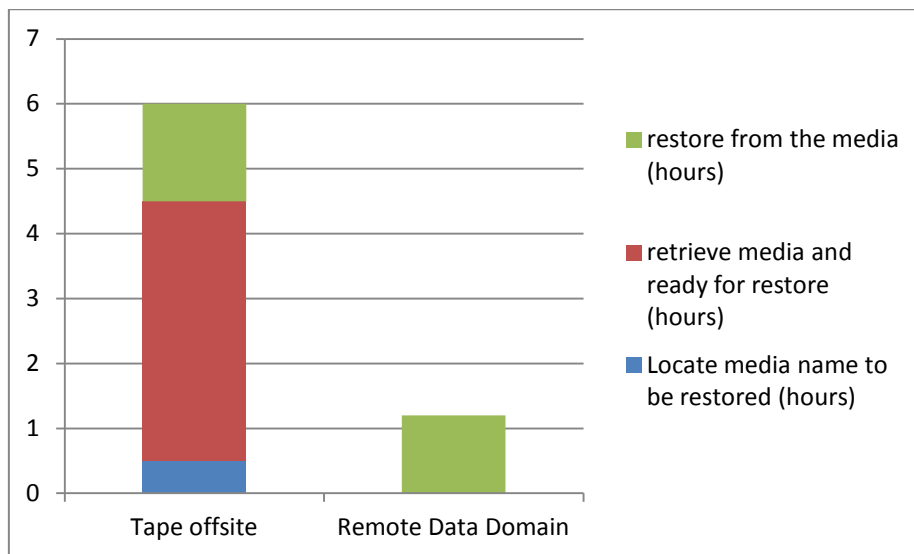
- Refer to the [Additional information](#) section on page 32 for some typical RTO and RPO (recovery point objectives) ranges for the main categories of recovery solutions.
- The tape restore scenario is based on the best practice for the most commonly used LTO-5 tape library with a 10 GB/minute restore speed.

In most cases, to restore data from the last backup on the remote site, you need to:

1. Check the backup/tape management log to identify the corresponding tape media that contains the required backup.
2. Retrieve the tapes from the offsite data center to the local data center. A professional data management service can currently provide the highest SLA within a 4-hour retrieval time.
3. Restore data from tapes to recover SAP applications to the point of time.

**Note** Compared with the tape storage scenario, restoring backups from the Data Domain system was much faster. The time taken to perform Steps 1 and 2 when restoring data from the Data Domain system was minimal.

Full backups are the fastest way for recovery because in most situations the latest backup of the past day is needed. While incremental backups reduce the backup time, the restore and recovery time will increase because in addition to restoring a full backup, the incremental backup needs to be applied to the database before log recovery can be started. This tradeoff needs to be considered with respect to the RTO that has to be covered by the service levels, since it usually takes longer.



**Figure 16. RTO breakdown comparison**

The time used in tape storage as recorded in Table 11 does not include the factor of manual error and resource shortage at the time of the outage, which, in reality, is most likely to happen, thus increasing the total RTO. However, by using Data Domain systems instead of the traditional tape storage, the data retrieval and restore processes can be seamlessly integrated and automatic.

## Conclusion

### Summary

EMC provides leading-edge technology to support the backup and recovery of SAP applications. This solution demonstrates how a Data Domain deduplication storage system with replication and NetWorker backup achieve the following benefits:

- Eliminates the impact on the production server during backup with Symmetrix TimeFinder/Clone technology.
- Reduces the backup footprint with a deduplication ratio of 27:1 after 10 backups through Data Domain deduplication storage systems. The deduplication ratio can be much higher as the backup accumulates.
- Automates and centralizes backup and recovery management with integration of NetWorker and Data Domain.

### Findings

Combining Data Domain systems with NetWorker brings SAP backup and recovery to a new level by:

- Automating restore operation by using BRRESTORE with seamless integration with NMSAP, NetWorker, and Data Domain
- Removing the performance impact of the backup on an SAP production environment using Symmetrix TimeFinder/Clone technology and a virtual machine-based mount host

The integration of Data Domain, DD Boost, and NetWorker increases SAP data protection levels and simplifies backup and recovery management tasks. Compared with the traditional offsite tape backup, this solution eliminates the risk of manual errors, transportation inconveniences, and media unreliability.

## References

### White papers

For additional information, see the white papers listed below:

- *EMC Data Domain Replicator - A Detailed Review*
- *Optimized Backup and Recovery for SAP Landscapes – Enabled By EMC Data Domain and EMC NetWorker – A Detailed Review*

### Product documentation

For additional information, see the product documents listed below:

- *EMC NetWorker Data Domain Deduplication Devices Release 7.6 Service Pack 2 Integration Guide*
- *EMC NetWorker Release 7.6 Service Pack 2 Installation Guide*
- *EMC NetWorker Release 7.6 Service Pack 2 Administration Guide*
- *EMC NetWorker Module for SAP with Oracle Release 4.2 Installation Guide*
- *EMC NetWorker Module for SAP with Oracle Release 4.2 Administration Guide*
- *EMC DD OS 5.0 Administration Guide*
- *EMC Data Protection Advisor 5.8 Administration Guide*
- *EMC Symmetrix TimeFinder Product Guide*
- *Master Guide for SAP ERP 6.0*

### Other documentation

For additional information, see the documents listed below:

- *Best Practice Backup and Restore for SAP System Landscapes*
- *SAP Business Continuity Management for SAP System Landscapes Best Practice Solution Management – May 2008*
- *vSphere PowerCLI Administration Guide*

## Additional information

### Overview

This section provides the sample configuration files used in this white paper.

### Backup utility profile

Table 12 lists the backup utility profiles.

**Table 12. PRD – initPRD.utl**

| Profile characteristic | Value           |
|------------------------|-----------------|
| parallelism            | 10              |
| savesets               | 50              |
| pool                   | uc652prddd50    |
| server                 | win-NetWorker   |
| client                 | xxx.xxx.xxx.xxx |
| ssNameFormat           | new             |
| NSR_SAVESET_RETENTION  | 2 days          |
| NSR_SAVESET_BROWSE     | 2 days          |

### Unmount VMDKs from the backup server

A sample script is provided below to unmount VMDKs from the backup server that stores the data files, redo logs, and trace logs prior to resynchronizing the disks. For more information about a full set of PowerCLI commands, refer to the *vSphere PowerCLI Administration Guide*.

```
>Connect-VIServer -server <vCenter IP address> -user <username> -password <password>
>$vm="PRD_Mount"
>$DS1=get-harddisk -vm $vm | where-object {$_.Filename -match "-prd_sapdata1"}
>remove-HardDisk $ds1 -confirm: $false
```

**Note** This script can be run on any host with vSphere PowerCLI.

### Mount cloned VMDKs to the backup server

Refer to the sample script below to mount cloned VMDKs to the backup server that stores the data files, redo logs, and trace files.

```
$esxhost="esx server name"
Get-VMHostStorage -VMHost $esxhost -rescanallhba
$vm="PRD_Mount"
$DS1=get-datastore | where-object {$_.Name -match "-prd_sapdata1"}
$a1=$DS1.Name
New-HardDisk -VM $vm -DiskPath "[a1] PRD/PRD.vmdk"
```

### Technical recovery solutions

Table 13 lists the durations of typical RTOs and the RPOs for typical recovery solutions.

**Table 13. Recovery solution comparison**

| Recovery solution                                | RTO (technical recovery)   | RPO                                  |
|--|--|--------------------------------------|
| Database restore and recovery                    | Gradual/intermediate<br>12 to 72 hours                                   | 0 (complete recovery)                |
| Tape shipping; pickup truck access method (PTAM) | Gradual/intermediate<br>48 to 168 hours                                  | 24 to 168 hours                      |
| Standby database (asynchronous log shipping)     | Intermediate/Immediate<br>1 to 8 hours                                   | 10 minutes to 24 hours               |
| Remote point-in-time copies                      | Intermediate/immediate<br>4 to 24 hours                                  | 4 to 24 hours                        |
| Asynchronous replication                         | Immediate<br>30 minutes to 8 hours<br>(in combination with HA solutions) | 5 minutes or more                    |
| Synchronous replication                          | Immediate<br>5 minutes to 8 hours<br>(in combination with HA solutions)  | 0 (depending on continuation policy) |

Source: Section 3.6.3, Recovery Options, *SAP Business Continuity Management for SAP System Landscapes Best Practice Solution Management*