EMC NETWORKER AND
EMC DATA DOMAIN BOOST
Best Practices Planning

Abstract
Conceptually understanding EMC® NetWorker® with EMC Data Domain® Boost provides a clear view of the business value and technical merits of integrating these two solutions in an IT infrastructure. This white paper moves past the conceptual stage to solution planning and deployment. Best practice guidelines are presented with the goal of eliminating implementation challenges. Knowledge and experience gained from a large number of customer deployments are logically presented for the overall benefit of those using a DD Boost solution with EMC NetWorker.

November 2010
Table of Contents

Executive summary ........................................................................................................................................ 4
Introduction .................................................................................................................................................. 4
  Audience .................................................................................................................................................. 4

NetWorker with DD Boost background ....................................................................................................... 4
  Configuration wizard ............................................................................................................................... 5
  Clone-controlled replication ..................................................................................................................... 6
  NetWorker reporting ............................................................................................................................... 7

Terminology and naming guidelines .......................................................................................................... 7
  Naming guidelines ................................................................................................................................... 7
  Data Domain hostname ............................................................................................................................ 9
  NetWorker Data Domain device type ....................................................................................................... 9
  Storage unit ............................................................................................................................................. 9
  Media pool ............................................................................................................................................... 10

Network planning ....................................................................................................................................... 10
  Example network topologies .................................................................................................................. 11
  Recommended network configuration ...................................................................................................... 13
  Data Domain Boost distributed segment processing ............................................................................... 14
  Authentication based on reserved ports .................................................................................................... 15
    Throttling clone-controlled replication ................................................................................................. 17
  Enable SNMP to allow NetWorker monitoring of Data Domain devices ................................................... 18

Deployment considerations ......................................................................................................................... 18
  NetWorker server considerations ............................................................................................................ 18
  Storage node considerations .................................................................................................................... 18
    Memory considerations .......................................................................................................................... 19
    Processing power/CPU ............................................................................................................................ 19
  Media pools ............................................................................................................................................. 20
  Client considerations ............................................................................................................................... 20
  Network considerations ........................................................................................................................... 22
  DD Replicator low-bandwidth optimization .............................................................................................. 22
  Seeding remote Data Domain systems ..................................................................................................... 22

Multiple NetWorker datazone configurations ............................................................................................ 22

Conclusion ................................................................................................................................................... 23

References .................................................................................................................................................... 24
**Executive summary**

Data deduplication is the enabling technology for next-generation data protection solutions. By reducing the size of backup datasets by an average of 10x-30x, backups can be retained onsite longer for fast operational restores, and replicated offsite efficiently over existing network links for disaster recovery and multisite tape consolidation. EMC® NetWorker® backup and recovery software delivers centralized backup and recovery operations across diverse computing and storage environments. NetWorker integrates with EMC Data Domain® deduplication storage systems through Data Domain Boost software. With DD Boost, parts of the deduplication process are distributed to the NetWorker storage node to significantly increase performance and simplify management.

Implementations of NetWorker with DD Boost should be well planned and documented to speed deployment and eliminate challenges that would otherwise be encountered using ad-hoc techniques.

Deployment is often followed by a series of trials or a period of testing intended to prove the solution functions as planned. In this white paper, best practices are examined and discussed to assist in eliminating the bottlenecks associated with deployment and functional testing of the solution.

**Introduction**

This white paper focuses on backups and the creation of clone copies with NetWorker software. Network configurations, clone-controlled replication, and centralized tape operations are examined. Recommended best practices, as well as alternate strategies, are covered with the goal of enhancing NetWorker solution planning and deployment.

**Audience**

System administrators and vendor staff performing and associated with NetWorker deployments are encouraged to use this paper to take advantage of substantial real-world knowledge gained by EMC from assisting customers.

**NetWorker with DD Boost background**

Data Domain Boost software provides integration between Data Domain storage systems and NetWorker software. It provides NetWorker with visibility into the properties and capabilities of the Data Domain storage system, control of the backup images stored in the system, and efficient wide area network (WAN) replication to remote Data Domain storage systems.

DD Boost for EMC NetWorker distributes parts of the deduplication process to the NetWorker storage node, which improves backup throughput by up to 50 percent and
EMC NetWorker and EMC Data Domain Boost

reduces data transferred from the storage node to Data Domain systems by 80 percent to 99 percent.

DD Boost for NetWorker has two components, the DD Boost library that is integrated into the storage node and the DD Boost server that runs on Data Domain systems with DD OS 4.8 or later. The DD Boost library is provided as the NetWorker Data Domain Device Type feature starting with NetWorker 7.6 Service Pack 1 and provides the following key enhancements for disk-based data protection strategies:

- Simplifies device setup and configuration by using wizards
- Increases aggregate backup throughput
- Provides NetWorker clone-controlled replication – Backup cloning available using EMC Data Domain Replicator, which provides network-efficient replication that is controlled, monitored, and cataloged by the NetWorker software
- Integrates NetWorker advanced reporting of the Data Domain storage systems
- Provides recovery of replicated backup images in their entirety or at a granular level via the NetWorker user interface
- Tape consolidation – Using the NetWorker clone-controlled replication functionality, backup images can be moved to a centralized location where they can then be cloned to tape.

**Configuration wizard**

The Device Configuration Wizard is started by right-clicking on the new Data Domain resource under the Device pane of the NetWorker Administration Console. The wizard walks through discovering Data Domain storage systems, by either hostname or network address, and setting up the devices (defined as storage units in Data Domain) within the system. Once connected to the system, users have a complete view of all the storage units configured and have the ability to create new ones.

Figure 1. The Device Configuration Wizard enables fast, repeatable operations that provide easy and efficient implementation of Data Domain systems
Clone-controlled replication

One of the key features of NetWorker is its ability to be aware of and fully control every data copy made using its cloning functionality regardless of the device used. Whether with tape or disk devices, NetWorker can manage data at the saveset or volume level, fully tracking the location of any copy within its media database. Each copy is tracked independently by a unique identification number. NetWorker 7.6 SP1 integrates this functionality with Data Domain storage systems by leveraging robust NetWorker cloning workflows with Data Domain Managed File replication. This feature is called clone-controlled replication. There are two main options to execute NetWorker cloning – immediate and scheduled. Immediate cloning starts Data Domain Managed File replication directly following a backup group completion. This is set up as part of the Group properties. Immediate cloning is typically not used if there are many backup groups scheduled to run concurrently within a backup window because of potential resource contention.

Scheduled cloning is new functionality added to NetWorker 7.6 SP1 that enables an administrator to create a schedule policy to execute clone jobs – including Data Domain Managed File replication – at a later time or date. By providing an easy and automated way to run clones at specified times, NetWorker helps ensure that users can meet required backup service levels while ensuring that disaster recovery copies are created in a timely manner.

Scheduled cloning is configured in the Clones resource in the Configuration pane in the NetWorker Administrator Console. By directing the jobs to a clone pool, data replicated by Data Domain systems inherits the policies assigned to that target clone pool. For more information regarding pools in NetWorker see the “Media Pools” section in Chapter 7 of the EMC NetWorker 7.6 SP1 Administration Guide.

By integrating Data Domain Replicator features together with the familiar NetWorker clone processes and workflows, NetWorker administrators gain simplified management of new, efficient ways of delivering disaster recovery readiness.
NetWorker reporting

To aid in management of operations with Data Domain systems, NetWorker has added seven new reports that provide Data Domain-specific backup and recovery statistics. This includes common reports such as Client summaries, Save Set summaries, and Daily summaries. There are also two types of drill-down reports, as identified by the drop-down arrow in the lower right-hand corner of the report, that can be used to show greater detail pertaining to the deduplication ratios achieved. These reports can be exported to various types of output including HTML, PDF, and plain table, CSV or even in various graphs.

NetWorker also now has the ability to collect SNMP traps from Data Domain storage systems to improve awareness of the total data protection infrastructure. By being able to report hardware-related alerts directly within NetWorker, administrators can feel comfortable that they are being made aware of all that is occurring in the backup environment in one single interface. Alerts for hardware failures like disk overheating, disk failure, or even a failed fan module are shown directly in the NetWorker Management Console (NMC). Each of these traps is selectable so the administrator can choose which traps are reported and which are not. This can be configured with the Configuration Wizard when the Data Domain storage system is initially set up or modified through the Properties dialog box by right-clicking on the Data Domain storage system in NMC.

Terminology and naming guidelines

Nomenclature, the assigning of names to specific components, within the NetWorker and Data Domain environment is an important consideration. It is best to use a naming convention that will be easily understood by the user, system engineer, and potentially any support personnel involved with the solution. Key recommendations for this nomenclature follow.

Naming guidelines

The following guidelines are offered to help you create consistent, easily identifiable names that facilitate configuration, reporting, and troubleshooting in the NetWorker and Data Domain environment.
Use the local hosts file to help diagnose and resolve naming issues. Use the `net hosts add` command on the Data Domain system to add hosts to the `/etc/hosts` file:

- Create names that are unique across all NetWorker datazones. Names should identify the network role, such as administration, backup, cloning, or production, with possibly a location or server name.

- Include source and target abbreviations in names to quickly identify whether network connections are correct. For example, add an abbreviation of the NetWorker storage node hostname in the Data Domain name and an abbreviation of the Data Domain hostname in the storage node name. Ensure that these names are included in the Data Domain `/etc/hosts` file.

- Specify all aliases (long and short names, IP address, and so on) for the NetWorker server and storage nodes in their respective NetWorker Client resources (Globals 1 of 2 tab).

- Ensure that all hostnames in the network can be consistently resolved from multiple locations in both directions. Examples include Shortname to IP, Longname to IP, IP to Shortname, and IP to Longname.

- In general, use short, easy to identify, descriptive names instead of IP addresses or fully qualified name strings for devices and storage nodes. Long names might not fit into some views. Examples of a long and short name:

  - NWDD365-1.brloa.lab.remcy.com:/DatazNW_Dir1
  - NWDD365-1:/DatazNW_Dir1

- Use standard alphanumeric characters, including “-” and “_” with no spaces or special characters, except pool names may not use underscores “_”.

- Use fixed formats (text field length and case) and include leading 0s in numbers, with a maximum of 50 characters.

- Avoid the use of dates in names where the dates might change or become meaningless in the future.

**Examples**

The following are some example name formats.

**NetWorker Data Domain devices**

- DD-DDsystem name-Device00-99
- Example: DD-Brandza-Device01

**Folders on a Data Domain system**

NetWorker Data Domain device names should refer to the NetWorker server and indicate whether it is for backup or cloning operation.
NW server or storage node_BACK or CLON_SU00-99
Example: DZBurl_BACK_SU01

**Volume labels for Data Domain**

DD-MEDIA or 5-letter descriptor-000-999
Example: DD-MEDIA-001

**Data Domain hostname**

The Data Domain hostname identifies the system. Hostnames are converted to IP addresses via host files or a name service such as DNS. Hostnames are not specifically NetWorker objects, but when used as storage server names, they tell the NetWorker storage nodes at the TCP/IP level how to connect to the storage server.

- For ease of use, employ the host's short name.
- 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. Each new hostname/IP address will require its own **DD device type license**. The best practice is to use the `ifgroup` command and consistently use the master IP address for management and configuration.

**NetWorker Data Domain device type**

The NetWorker Data Domain device type is a logical object defined within NetWorker that maps directly to a Data Domain system storage unit:

- These devices cannot be shared between multiple storage nodes.
- Multiple NetWorker Data Domain devices can be shared on a Data Domain system.
- Devices can be set for target and maximum sessions; see the **Deployment considerations** section.

**Storage unit**

A storage unit (SU) is a disk target within a Data Domain system that resides under the OST pane's LSU tab.

It is best to create storage units directly under the Data Domain storage systems as follows:

```plaintext
<Data Domain system>/<NetWorker Server shortname>_<NetWorker storage node shortname>_<chosen device name>
```

Example: DD660a:/Jupiter_saturn_dev001
**Media pool**

NetWorker media pools act as filters that tell the server which backup devices and associated volumes should receive specific data. Figure 2 shows pool configurations:

- Restrict the Data Domain device types in a pool to a single Data Domain system to their own media pool.
- Do not mix Data Domain device types with AFTD or tape devices in the same media pool.
- From a high availability perspective, to ensure your backups continue without failures, it is important that your target devices are accessible and online. Depending on your requirements and environment, you may wish to have a backup failover device in case the Data Domain system becomes unavailable. If this is the case, be aware that routing backups to a non-deduplicating device will impact your deduplication ratios.

![Media pool configuration](image)

**Figure 2. Media pool configuration**

**Network planning**

There are varying network considerations associated with any given NetWorker deployment. At a minimum, a single Data Domain system that is configured as a Data Domain device type may be network-connected to a NetWorker storage node. This following section discusses examples of network topologies in various NetWorker and Data Domain configurations.
Example network topologies

Figure 3 shows an example of a NetWorker storage node network connected to two Data Domain systems that share a common LAN configured for clone-controlled replication. In this use case, both backup and clone-controlled replication traffic use the same network interface card (NIC) on a given Data Domain system.

![Diagram]

Figure 3. Clone-controlled replication with a common network
Figure 4 shows three NetWorker storage nodes and two Data Domain systems. Each NetWorker storage node has a LAN connection (blue) to both Data Domain systems. A dedicated GbE (green) also connects the Data Domain systems to each other. Backup traffic uses the LAN connection between a given NetWorker storage node and a Data Domain system. Clone-controlled replication traffic uses a separate NIC on each Data Domain system. This configuration may be preferred in cases where the additional bandwidth between Data Domain systems of a 1 GbE network provides the ability to accommodate clone-controlled replication traffic.

Figure 4. Clone-controlled replication with a separate network

A NetWorker client can be backed up through a number of different NetWorker storage nodes by leveraging the storage node affinity feature, as shown in the next figure. Multiple storage nodes can be added to the list in the Globals (2 of 2) tab in the Client resource properties. In this example, the Data Domain system has been configured so that each NetWorker storage node has access to the NetWorker Client resource. This configuration allows NetWorker to bypass an offline storage node when fulfilling a backup or restore request.
Figure 5. NetWorker storage node affinity with the Data Domain device type

The EMC NetWorker 7.6 SP1 Administration Guide provides details on storage node affinity.

**Recommended network configuration**

Depending on customer requirements and the existing network infrastructure and deployments, it is recommended to use a dedicated backup network for NetWorker storage nodes and NetWorker Data Domain devices.

The following recommendation outlines configurations that have shown the most beneficial results with consideration for simplicity and enhanced performance:

- Use a dedicated backup area network for NetWorker storage nodes and Data Domain systems. NetWorker clients and NetWorker storage nodes should ideally be interconnected through a separate front-end network (Figure 6).
  - Segregate NetWorker storage node and Data Domain device type traffic from other network traffic so contention issues are limited. Knowing the available
bandwidth to be managed helps achieve aggressive data protection and recovery service levels.

- Establish a scalable infrastructure in case data protection network bandwidth requirements change over time.

While not always possible, based on customer requirements and pre-existing NetWorker storage node and network infrastructure deployments, the use of a dedicated backup network is preferred when compared to mixed-use network configurations.

![Figure 6. Dedicated backup network](image)

**Data Domain Boost distributed segment processing**

DD Boost distributed segment processing distributes parts of the deduplication process to NetWorker storage nodes, generally increasing the backup data transfer rate and decreasing both NetWorker storage node and Data Domain aggregate CPU utilization. Additionally, LAN traffic is reduced between NetWorker storage nodes and the Data Domain system as only unique data is transferred over the backup network.

DD Boost distributed segment processing is enabled by default on Data Domain 4.8 and later systems. Data Domain systems upgraded to DD OS 4.8 from prior DD OS versions will not have the distributed segment processing option enabled by default so the user must enable this manually from the Data Domain storage systems interface.

- Use distributed segment processing in NetWorker environments. Enable this option by using the command line interface `ost option set boost enabled` command on the EMC Data Domain system (Figure 7) or from the EMC Data Domain Enterprise Manager GUI.
Note: It is important to ensure that NFS is also enabled. If NFS is not enabled, Boost will not be active.

![dd690a.se.local - PuTTY](image)

```
sysadmin@dd690a# ost option show
Option    Value
boost     disabled
```

```
sysadmin@dd690a# ost option set boost enabled
```

```
sysadmin@dd690a# ost option show
Option    Value
boost     enabled
```

Figure 7. Enabling the Boost option

DD Boost option status is queried with the `ost option show` command on the Data Domain system. The results of the command indicate that the Boost option was disabled. The subsequent command, `ost option set boost enabled`, is used to enable the DD Boost option.

Specifically not recommended are configurations where a NetWorker storage node and Data Domain system are connected via WAN.

To get the most of the Boost software with NetWorker ensure you use at least eight backup streams. If the total number of streams from all NetWorker storage nodes is guaranteed to be lower than four at all times, disabling Boost may result in better performance.

For additional information, please refer to the *EMC NetWorker Data Domain Deduplication Devices 7.6 SP1 Integration Guide*.

**Authentication based on reserved ports**

Before a NetWorker storage node can access the desired Logical Storage Unit (LSU) path on the Data Domain system, the path needs to be exported to the host for NFS access. Failure to do so will manifest in to an NMC Device Configuration Wizard error that says “Failed to contact the Data Domain system. Host could be unreachable, or username/password could be incorrect. Do you wish to configure manually?” If the manual configuration step is followed then it will lead to the error "... the user has insufficient privilege". Figure 8 shows an example:
Figure 8. A user has insufficient privilege in the Device Configuration Wizard

To fix this error, the Data Domain system must provide NFS access to the NetWorker storage node. This can be done with the following command on the ddsh console for the Data Domain system.

```
nfs add /backup/ost <hostname> (rw,no_root_squash,no_all_squash,insecure)
```

Pay special attention to the "insecure" flag as the default export setting uses the secure option. Failure to set this flag to insecure will still result in an access issue with a different error message: "... user does not have sufficient access rights". Figure 9 shows an example.

A word about the insecure flag: The "insecure" option is not inferior to the "secure" option. It just allows the clients using port numbers > 1024 to connect to the Data Domain system. Robust security is ensured through authentication based on the Data Domain system’s username/password.

With DDOS 4.9 and later, the default is changed to "insecure" for ease of configuration.
Throttling clone-controlled replication

Throttling is controlled at a global level and is configured on each Data Domain system. Limiting the rate of network bandwidth used by the replication process can be based on various criteria such as a scheduled or temporary rate. Caution should be exercised as throttling back network bandwidth consumption may extend replication runtimes.

The net effect of network bandwidth throttling may impact recovery point objectives for disaster recovery. Other effects might include the queuing of jobs, as clone jobs contribute to the destination storage unit’s or Data Domain system’s concurrent jobs. Once a storage unit’s maximum concurrent sessions parameter has been reached, new jobs requiring the use of the storage unit will be queued as they await storage unit resource availability. Based on service-level requirements, it might be possible to limit the quantity of network bandwidth required by limiting the amount of data that needs to be replicated. One possibility worth considering is the cloning of only full backups, where incremental backups are not duplicated. As well, considering this with the scheduled cloning feature is important. Be sure that the throttling needs are
considered when designing schedules to avoid conflicts and a negative impact on the success of clone jobs.

**Enable SNMP to allow NetWorker monitoring of Data Domain devices**

NMC provides the ability to monitor details of current Data Domain backup operations as well as display events that require user intervention. Ensure that SNMP is enabled on the Data Domain system and that the Data Domain system is configured to send traps to NMC server.

**Deployment considerations**

This section gives deployment considerations that you should be aware of as you deploy NetWorker into your environment and take advantage of NetWorker's ability to manage your backups with Data Domain via the DD Boost interface.

**NetWorker server considerations**

- **Retention policies**
  - Data Domain retention locks are not honored within NetWorker. If NetWorker has differing retention policies configured it will honor those but the actual data on the Data Domain system will not be deleted if the retention lock option is on. It is a best practice to avoid using the retention locking on the Data Domain system and let NetWorker manage this operation.

- **Parallelism**
  - Each Data Domain system model supports a maximum number of simultaneous connections. To NetWorker this translates to the number of streams as part of the parallelism settings. Do not exceed server parallelism for the allowable Data Domain model's max streams. If this number is exceeded, performance degradation can occur.
  - If using one Data Domain system between two or more datazones the total number of streams is limited to the max stream count of the Data Domain system. Be sure to partition the streams accordingly between the datazones.

- **NetWorker client-side encryption via `encryptasm`** is not supported when writing to a Data Domain system using the Data Domain device type.

- **NetWorker client-side compression via `compressasm`** is not supported when writing to a Data Domain system using the Data Domain device type.

**Storage node considerations**

The following are the options to consider when determining if an existing storage node in your deployment is sufficient for your backup operations or whether you should consider adding a new storage node.
Memory considerations

The Target sessions attribute refers to the optimal number of backup sessions accepted by any active NetWorker device. If a device has a Target session setting of 4, NetWorker will look for another available device when, say, a fifth stream is started. If no device is available the target sessions setting will be ignored and the extra stream will get added to the initial four on the original device. The Max session setting determines the absolute maximum number of streams a device will allow. The Max session is a hard limit and the Target session a soft limit. The default setting for the Data Domain device Target session is 1 and the Default Max session is set to 4, with an overall hard limit of 10.

Each read-write device (active nsrmmd process) that takes one save stream requires 48 MB of RAM on the storage node. Each additional save stream requires 16 MB. The corresponding read-only device requires 16 MB, regardless of the number of save streams. Thus one save stream requires 64 MB, four save streams require 128 MB, and the maximum of 10 save streams requires 224 MB.

It is recommended that no more than 16 Data Domain device types reside on a storage node that is running the maximum number of 10 save streams per device. This scenario would require \((16 \times 224) = 3.6\) GB of physical memory on the storage node.

It is recommended that a storage node have at least 8 GB of RAM when hosting NetWorker Data Domain devices. The use of two interface connections (1 GbE or 10 GbE) is also highly recommended.

To determine the amount of memory needed for your particular deployment, use the following formula as a guideline:

\[
m = n((16 \times s) + 64)
\]

Where \(m\) = Memory requirements in MB, \(n\) = Number of DD devices, \(s\) = Max sessions

Processing power/CPU

One of the powerful benefits to using NetWorker with the DD Boost software is the ability of the storage node to process the data to be backed up and only send the unique data to the Data Domain device. This processing, known as distributed segment processing (DSP), means that the storage node will have a higher CPU utilization on the first backup, but subsequent backups should see smaller CPU utilization due to less data needing to be processed over time. In addition to the lower CPU utilization, LAN traffic is reduced between the NetWorker storage nodes and the Data Domain system as only unique data is transferred over the backup network. This means that typically the same CPU being used in an existing environment will suffice when moving to an integrated NetWorker and DD Boost solution.
Media pools

Each NetWorker client stores data to a media or “target” pool. A pool is used to direct the data from backup clients or the data from storage volumes in clone operations, to the storage devices that are members of the pool.

Each NetWorker Data Domain device is associated with a storage volume (Volume Name). Each volume is named by a label template when the volume is mounted on the device. The Volume Name implicitly associates the volume and the device with the specified pool.

Figure 10. The relationship between volumes, devices, and pools

When creating pools, as a best practice, it is recommended that you create your pools for the Data Domain devices with only devices from a single Data Domain system in a pool. This will ensure that the deduplication ratios are generated with information from a single Data Domain storage system, ensuring more exact figures. As well, not mixing device types in a pool will ensure that the Managed File replication of the Data Domain systems will be leveraged.

Client considerations

There is a new attribute in the Client resource where the administrator can assign a client to a specific pool.

As a best practice it is recommended that you not separate groups from target pools at the client level. It is further recommended that you not have mixed device types in your pool. However, if you do choose to have mixed device types in your pools, make sure that you understand the constraints and that you select Data Domain Backups in the Deduplication attributes group. This ensures that the client data will be backed up only to the NetWorker Data Domain devices, your data will be sure to be deduplicated as planned, and the ratios you expect will be achieved.
The Client resource setting for Data Domain backup applies only to the **save sets** (the data portion of backups). Selecting this option does not back up bootstrap and index files to a Data Domain device.

Figure 11. Separate source and destination NetWorker storage nodes

Figure 11 shows replication between two Data Domain storage units. The NetWorker storage node initiating a replication job – also known as clone-controlled replication – needs to have credentials to access both the source and destination storage units.

NetWorker 7.6 SP1 provides a GUI-driven device creation wizard that allows credentials to be specified.

In cases where clone-controlled replication uses a destination storage unit that may be geographically distant from the NetWorker storage node initiating the replication job, the storage unit definition should not allow the geographically distant NetWorker storage node to use the storage unit for backup jobs.
Network considerations

Replicating backup images under the control of NetWorker clone-controlled replication includes the ability to use the same network that is used for backup and restore operations or to use a different network. In cases where clone-controlled replication traffic flows between geographically different locations, some customers have chosen to use a separate dedicated network connection for replication. This connection links source and destination Data Domain systems specifically for the purpose of replication controlled by NetWorker-initiated clone-controlled replication. Users needing to track WAN link usage may also prefer this approach.

The network used for clone-controlled replication is based on network name resolution on the source Data Domain system. The destination Data Domain system is known to the source Data Domain system based on the IP address supplied by DNS, or by a local hosts file entry. Populating the source Data Domain systems hosts file with the desired IP address of the destination Data Domain system is all that is required to use a specific NIC and network. If this value is not present, NetWorker will perform clone-controlled replication using the same network it uses to access the source and destination Data Domain systems for backup and restore jobs.

DD Replicator low-bandwidth optimization

Supported with DD OS version 4.8 and later, DD Replicator low-bandwidth optimization provides increased efficiency when performing replication across WAN links of less than 6 MB/s. Enabling this DD Replicator feature requires no changes on NetWorker storage nodes as NetWorker is unaware of this Data Domain system configuration option.

Seeding remote Data Domain systems

New deployments looking to utilize clone-controlled replication may have network bandwidth limitations between sites that could cause the first week of clone-controlled replication jobs to run for extended time periods. One solution to this dilemma is to seed the remote Data Domain system locally, and then after a week or so relocate it to the intended remote site.

Multiple NetWorker datazone configurations

By definition a “NetWorker datazone” refers to a unique NetWorker server and the associated storage nodes and clients. This section examines the use of Data Domain systems that are configured for use in multiple NetWorker datazones, where two or more NetWorker servers have configurations that relate to a single physical Data Domain system.

Deploying a single Data Domain system in multiple NetWorker datazones provides cost advantages when compared to the alternative of using different Data Domain systems for each NetWorker datazone. Take, for example, the scenario where there are two NetWorker datazones. Each NetWorker datazone is in a different data center.
Data center 1 performs local backups to a Data Domain system and then uses clone-controlled replication to create a copy of the backup images at data center 2. Likewise, at data center 2 local backups are performed to a Data Domain system and clones using clone-controlled replication are created at data center 1. Using non-shared Data Domain systems would require four separate systems. The shared model is cost-effective in that it only requires two separate systems.

![Image of Data Domain system](image)

**Figure 12. Deploying a single Data Domain system in multiple datazones**

Note: It is important to understand that there is no ability to share backups or NetWorker index data between NetWorker datazones with this implementation.

Data Domain recommends using naming conventions that assist in simplifying the shared use of Data Domain systems in multiple NetWorker datazones. Naming the storage units such that they include the Data Domain system name, the NetWorker server name, and the “-su” suffix assists in properly configuring each storage unit as a Data Domain device within NetWorker.

**Conclusion**

Together, NetWorker and Data Domain improve backup and recovery operations by increasing performance, allowing you to store more data on disk, simplifying the creation of backup copies, and simplifying management.

NetWorker with DD Boost improves backup throughput, reduces LAN bandwidth requirements, and decreases storage node loads. With DD Boost, NetWorker is able to control Data Domain network-efficient replication to create and manage backup copies at remote sites, enabling advanced disaster recovery strategies. Disaster recovery copies of backup images are created faster, and are available at the disaster recovery location sooner when compared to tape-based solutions. Greater visibility into the entire backup operation including the Data Domain system from a central NetWorker interface helps administrators be more effective and productive.

With proper planning and consideration, an integrated software and hardware deduplication solution from EMC will provide unmatched reliability and efficiency for your backup and recovery operations.
References

The following documents can be found on the EMC Data Domain secure access customer support site at [https://my.datadomain.com/](https://my.datadomain.com/) and EMC Powerlink®:

- *EMC Data Domain Operating System (DD OS) Administration Guide*
- *EMC NetWorker 7.6 SP1 Documentation Portfolio*
- *EMC NetWorker 7.6 SP1 Release Notes*
- *EMC NetWorker Data Domain Deduplication Devices 7.6 SP1 Integration Guide*