Dell™ PowerVault™ DL Backup-to-Disk Appliance Powered by CommVault

Remote Office Backup Strategy

CommVault DL Team

Dell | CommVault

delltechcenter.com

commvault.com/dell

Table of Contents

Dell™ PowerVault™ DL2000 i
Executive Summary 2
Introduction 2
What is Discrete Data Replicator (DDR)? 2
Remote Single Instance Backup Use Cases 3
Efficient Storage Consolidation of Remote Offices 3
Create Copies for Disaster Recovery at a Secondary Site 3
Use Cases Not Addressed by DDR 4
Automatic Failover between Primary and Secondary Media Agents 4
Creating Point-In-Time Images of Mount Paths 4
Replacing AuxCopy and Synthetic Full 4
Fan-Out with DDR 4
DDR Configuration 5
Prerequisites 5
Configure the Magnetic Library on Replicated Disks 6
Other Considerations 8
Out of Band Initialization 8
Fan-In Configuration 9
Conclusion 9

Table of Figures

FIGURE 1: REMOTE SINGLE INSTANCE BACKUP USING THE REPLICATED DISK LIBRARY 5
FIGURE 2: ENABLE REPLICATION ON THE REPLICATED DISK LIBRARY 6
FIGURE 3: SDDR REPLICATION PAIR 7
FIGURE 4: MEDIA MANAGEMENT PARAMETER FOR DDR REPLICATION INTERVAL 8

Executive Summary

The Remote Office Backup Solution (SDR) is designed to seamlessly replicate data from remote office DL2000 media agents—which are ideal systems for disaster tolerance and recovery protection—to a centralized data center, where data is stored on magnetic media for ready access. Consolidating all backup data at the centralized data center provides redundancy for disaster recovery, as well as an alternate source for normal data recovery operations.

This innovative solution is easily deployed, without any disruption of current data protection operations, and with no impact to end users at remote offices; no additional software or configuration is required on client computers, as only DL2000 media agents are involved. Even more importantly, the data from every CommVault® *i*DataAgent deployed at a remote office can be replicated to the centralized data center, no matter the source of the data—this includes all supported file systems, databases, and data types. By utilizing a pair of DL2000s, you effectively have a self-contained remote office management solution that is cost effective and requires a short, one-time setup process.

Introduction

Remote office data can be consolidated in two ways for data protection. The traditional method is to backup servers in remote offices to a local media agent, and then Auxilary Copy (AuxCopy) the data to a central site. While simple, this method requires a significant amount of network bandwidth to transfer the backups. Since AuxCopy operations are performed outside the backup window, this consumes network bandwidth during business hours, unless a dedicated network is available for backups. A second method is to use Continuous Data Replicator (CDR) technology to replicate remote office data to a central location. Once the initial replica is created, only changed bytes are continuously transferred, thus using less network bandwidth. While highly effective, neither of these consolidation methods leverages the Object Level Single Instancing (SIS) functionality introduced in CommVault® Simpana® 7.0 software. This feature prevents duplicate objects from being transmitted over the network and thus reduces both the network bandwidth usage and the amount of space required to backup the data.

Simpana 7.0 Service Pack 2 (SP2) introduces Discrete Data Replicator (DDR), an extension to the CDR model which, in conjunction with a replicated disk library, transfers only the unique backed-up objects from remote offices to the central data center. This functionality, also called Remote Single Instance Backup, enables network efficient remote replication and consolidation of data protection copies in the data center, yet retains the ability to perform local backups and restores. This document describes how DDR helps prevent duplicate objects from being transmitted over the network, thus reducing bandwidth utilization and lowering disk costs.

What is Discrete Data Replicator (DDR)?

Simpana 7.0 software introduces a special configuration of CDR that eliminates the use of a continuous byte-level capture driver; this configuration is called Discrete Data Replicator (DDR). Unlike CDR, DDR does not continuously monitor changes on the source file system. Instead, it initiates a SmartSync operation between the source and target file areas on a periodic basis, replicating new file changes and

deletes (orphan handling) based on changes that have occurred on the source since the last DDR operation.

This sequence is currently triggered by a timed update service running on the DL2000 media agent that causes files in the Disk Library data path to be replicated to the destination file area. This model is referred to as “discrete” replication, since the files are replicated at fixed discrete intervals. This form of replication will only copy the changes relative to those discrete time periods, as compared to continuous replication which continuously captures and replicates all changes.

Remote Single Instance Backup with DDR is essentially a replicated disk library consisting of a replication pair set up between the two mount paths of the replicated disk library (see Figure 1). All backups are written to the local read/write mount path, which is then replicated by DDR on a periodic basis to the secondary read-only mount path. On enabling Single Instancing on the Storage Policy, only unique objects are written to the primary mount path, and only these unique objects are replicated to the secondary mount path. Encryption can be enabled on the replication set to secure data during transfer.

Note: DDR uses the same installation package and set of binaries as CDR; however, these binaries can operate only in one of two modes: CDR or DDR. Enabling DDR on the media agent disables the ability to perform continuous replication operations.

Remote Single Instance Backup Use Cases

Remote Single Instance Backup with DDR is primarily used for replication and/or centralizing data protection operations when there is limited network bandwidth between remote sites and the data center.

Efficient Storage Consolidation of Remote Offices

The recommended CommVault solution for remote office consolidation is CDR. However, CDR’s lack of de-duplication capabilities can be perceived as a disadvantage when replicating unstructured data. DDR’s Remote Single Instance Backup capabilities neutralize that perceived disadvantage. Backup (or archive) of remote clients consolidates data at the remote media agent; DDR is then used to replicate this data to the central data center. Single instancing at the remote media agent ensures that only unique objects are transmitted over the network, thus ensuring minimal network impact. The copy at the central site is also available for secondary operations.

Create Copies for Disaster Recovery at a Secondary Site

Traditionally, AuxCopy is used to create a disaster recovery copy at the secondary site. Primary backup sets are created during the backup window; once the backup window is complete, these eligible backup sets are copied to the secondary copy.

Having this kind of policy has two main side effects. First, the network between primary and secondary sites is occupied for data protection activities, even outside the backup window. Second, since all the data from the latest backup set is transmitted simultaneously, accomplishing this operation in a timely manner requires higher network bandwidth.

Implementing a Replicated Single Instance Disk Library eliminates both of these side effects. DDR ensures that data is copied to the second mount path of the replicated disk library soon after it is written to the first mount path. Thus, data is available at the secondary site within minutes of backup completion. Also, since the data is single instanced, only unique objects are copied over, thus reducing bandwidth requirements considerably. The secondary mount path can then be used as a source for various

purposes, such as creating additional copies of data, remote restore, tape vaulting (with or without encryption), content indexing, and enterprise search.

Data on the primary mount path of the replicated disk library continues to be available for local recovery operations.

Use Cases Not Addressed by DDR

DDR does not replace any existing CommVault functionality or address some other specific use cases.

Automatic Failover between Primary and Secondary Media Agents

DDR in itself does not support the ability to failover between primary and secondary media agents—replication with DDR is unidirectional. Moreover, the Single-Instance Database (SIDB) is not replicated. Also, the secondary media agent and mount path is available strictly as a read-only copy for recovery or for secondary operations. If the secondary mount path is converted to a writeable copy (via mount path properties), subsequent backups will not be replicated to the first mount path.

Creating Point-In-Time Images of Mount Paths

DDR does not support creating point-in-time snapshots of the disk library, and it does not use drivers to track and log file system updates. It also does not have the ability to create recovery points. Periodically, the media agent invokes DDR (default 30 minutes), which then checks the USN on the primary mount path and determines the files that need to be copied over.

Replacing AuxCopy and Synthetic Full

Remote Single Instance Backup with DDR does not replace AuxCopy. Even though there are two copies of the data, they have the same retention period. When data aging criteria for the storage policy copy is met, data from both mount paths is deleted.

Remote Single Instance Backup with DDR does not replace Synthetic Full operations, either. However, depending on the customer requirements, the frequency of Synthetic Full operations may be reduced, since there are two copies of the data.

Fan-Out with DDR

DDR does not yet support the ability to replicate the mount paths to multiple destination media agents, which may not be desirable. DDR copies all have the same retention criteria, and the secondary copies are read only. If data from the primary mount path is aged, it is also removed from the secondary mount paths.

The suggested method for creating multiple copies of data is to use DDR to replicate the primary copy to create two copies of data, and then use AuxCopy to create as many copies as required—each with different retention criteria. AuxCopy can use the secondary mount path as the source.

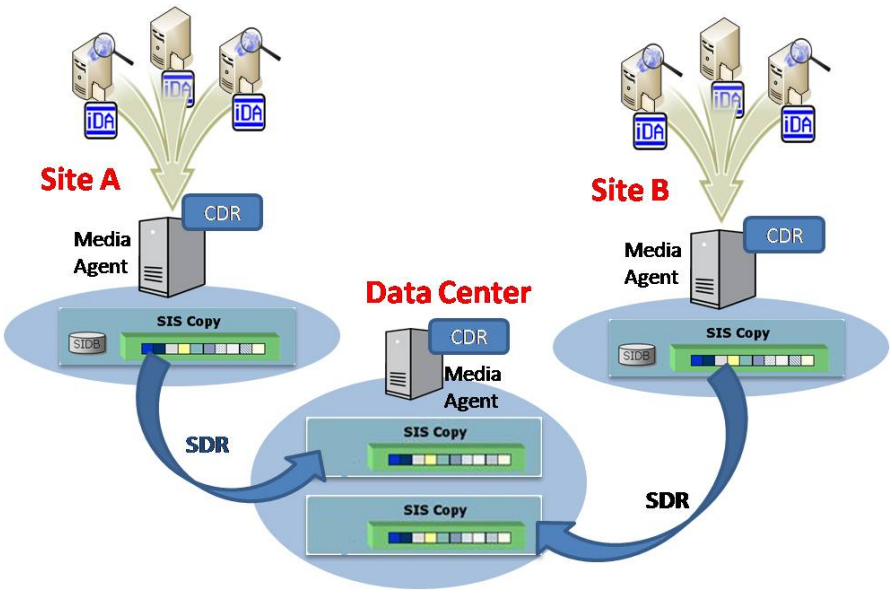

FIGURE 1: REMOTE SINGLE INSTANCE BACKUP USING THE REPLICATED DISK LIBRARY

Note that the SIDB itself is not replicated; only the single-instanced data path content is replicated. The SIDB is required for checking the uniqueness of an object only when writing it to the first mount path. The SIDB is not required when reading the object for recovery; hence, there is no need to replicate the SIDB to provide read access from the replicate dataset. Dell recommends backing up the SIDB using File System iDA.

Service Pack 2 also includes updates for automatic configuration of a replicated disk library with DDR replication pairs. Single instancing needs to be enabled on the Storage Policy copy separately. Additionally, data transfer options like encryption and compression are also available. The section entitled “DDR Configuration” provides more details.

Note: DDR uses the same installation package and set of binaries as CDR; however, these binaries can operate only in one of two modes: CDR or DDR. Enabling DDR on the media agent disables the ability to perform continuous replication operations.

DDR Configuration

Prerequisites

1. Make sure the system requirements for CDR are met on both the media agents that will be part of DDR.
2. Install media agents with operating system compatibility on the two servers.
3. Install CDR agents on each of the two servers.
4. Do not create any replication pairs.
5. Apply the latest Service Pack for Simpana.

Configure the Magnetic Library on Replicated Disks

Step-by-step instructions are available in **Books Online** at **http://documentation.commvault.com/commvault/release_7_0_0/books_online_1/english_us/features/magnetic_library/replicated_mag_lib_config.htm**.

IMPORTANT: Always check the **Enable replication** option, as shown in Figure 2, to automatically create a DDR replication pair.

NOTE: DDR does not support Replicated Disk Libraries where the primary disk target (read/write path) is a network share. The primary disk target has to be a dedicated local volume.

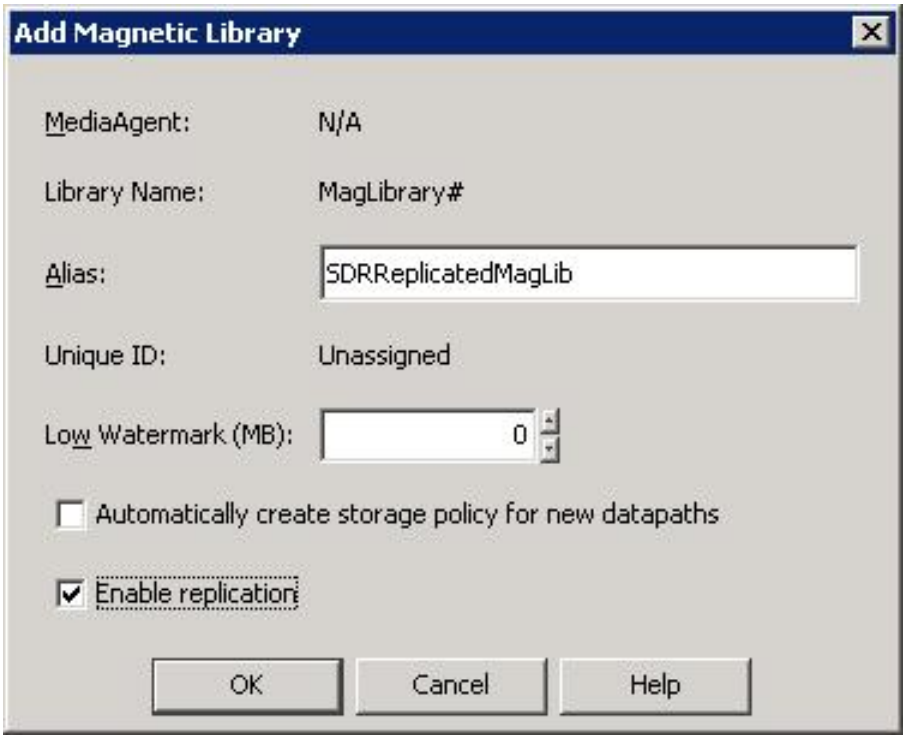

FIGURE 2: ENABLE REPLICATION ON THE REPLICATED DISK LIBRARY

The DDR Simple Replication Pair shown in **Figure 3** has the primary mount path of the replicated disk library as its source and the secondary mount path as its target.

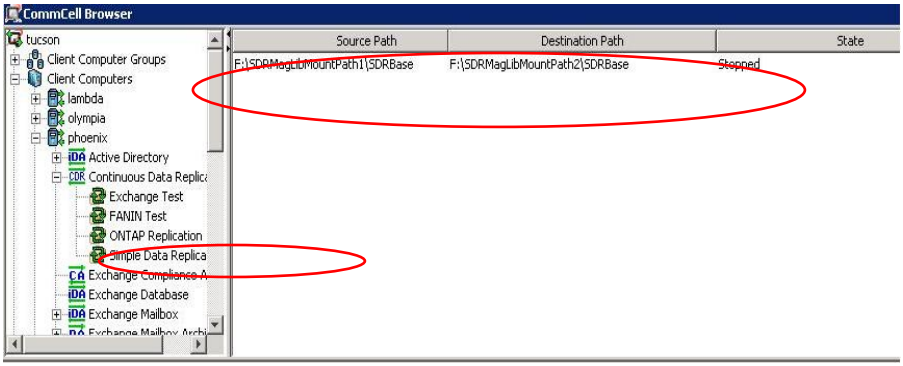

FIGURE 3: SDDR REPLICATION PAIR

The state of the replication pair is initially **Stopped**. When the media agent initiates the DDR process, the state changes to **Replicating,** and then returns to **Stopped** once the two mount paths are synched.

Storage Policy

Create a new Storage Policy and a Storage Policy Copy, and point them to the Replicated Disk Library. Enable Single Instancing on the Storage Policy Copy.

Media Management Configuration

DDR replication is controlled by the media agent service. On a periodic basis, the media agent service initiates the DDR process, which determines changed and deleted files on the primary mount path and replicates these changes to the secondary mount path. The frequency of these checks can be configured from the **Media Management Configuration** settings, which are accessible from the **Control Panel**. The parameter **Interval (Minutes) between magnetic space updates** (default 30 minutes) controls the frequency of DDR. For directions on changing the default frequency, refer to the CommVault documentation.

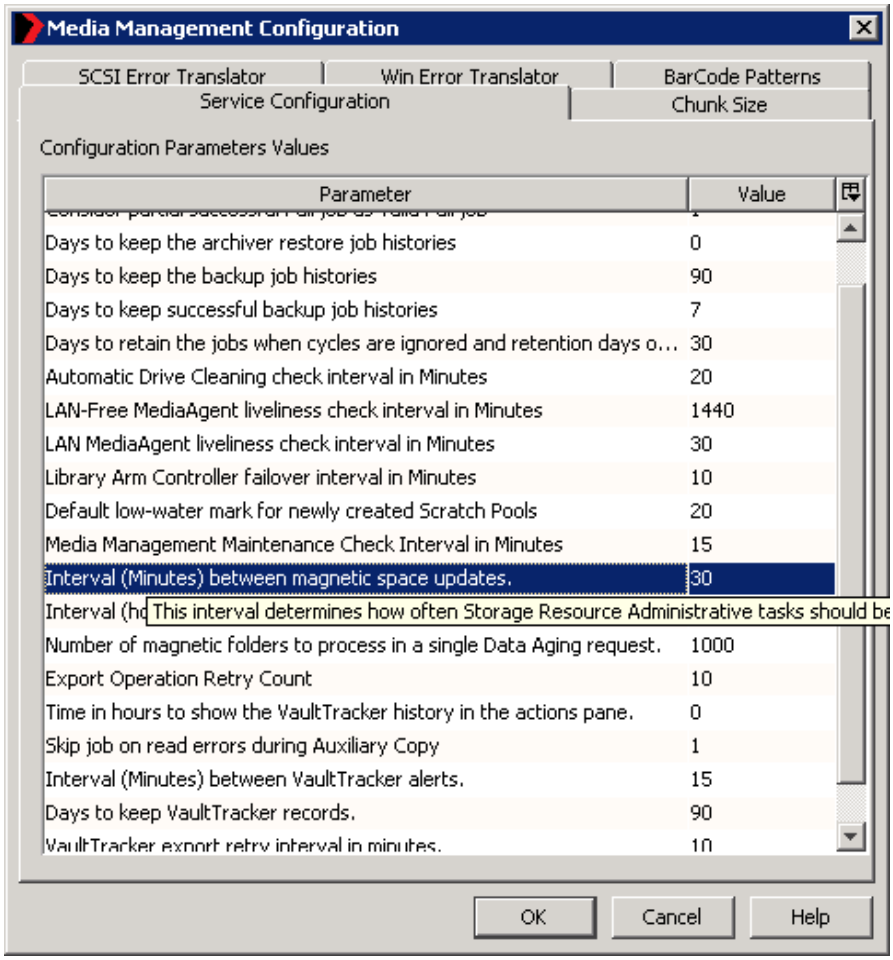

FIGURE 4: MEDIA MANAGEMENT PARAMETER FOR DDR REPLICATION INTERVAL

Other Considerations

Out-of-Band Initialization

It is possible to perform the initial sync using the Out-of-Band Sync process as described in the online documentation available at:
http://documentation.commvault.com/commvault/release_7_0_0/books_online_1/english_us/prod_info/flr.htm

This is useful when the first backup is too large to be replicated in a reasonable amount of time over the network. This does not imply that an existing magnetic library can be converted into a replicated disk library with SIS—the replicated disk library must be configured first.

The steps to configure a replicated disk library are:

1. Install media agents and CDR nodes.
2. Create a replicated disk library with replication enabled.

3. Disable DDR replication using the registry key.
4. Perform the Out-of-Band Sync process.
5. Enable DDR replication using the registry key.

The next DDR operation will continue to sync from that point onward.

Fan-In Configuration

It is possible to have multiple remote nodes replicating to the same target server in the data center. To accomplish this, a separate replicated disk library per pair is required. For each of these replicated disk libraries, the secondary mount path must be a unique location.

Conclusion

DDR Remote Single Instance Backup capability provides an efficient and effective option to protect remote offices by providing local backup and restore capabilities along with a secondary disaster recovery copy at a central site. By combining a Replicated Disk Library and Single Instancing, DDR provides the ability to create secondary disaster recovery copies while achieving maximum network efficiencies. Available enterprise-class features—like data encryption and compression—provide additional tools to implement a secure, robust, and optimal remote office protection strategy.