Disaster Recovery Solutions Guide

Dell FS8600 Network-Attached Storage (NAS)

FluidFS System Engineering

January 2015

Revisions

Revision	Date	Description
A	October 2013	Initial Release
B	January 2015	Updated for FluidFS v4

Table of contents

1 Preface

1.1 Audience

The audience for this document is intended to be systems, networking, storage or backup administrators who are responsible for the day-to-day management responsibilities of a Dell Compellent FS8600 FluidFS NAS solution.

Proper management of an FS8600 requires administrators (or teams of administrators) capable of managing and configuring enterprise-class Fibre Channel SAN and Ethernet networks, any enterprise-grade backup software intended to be used, the Dell Compellent Storage Center, as well as general purpose NAS administration

1.2 Purpose

The purpose of this document is to help the storage administrator understand the FS8600 disaster recovery mechanism. This document is not intended to be a primer or Dell Compellent FS8600 introductory resource for any of the subject matters involved, and it assumes at least introductory knowledge of many of the subjects covered in this document.

This document should be used in conjunction with other Dell Compellent resources as listed in Appendix B – Additional Resources.

1.3 Disclaimer

The information contained within this best practices document is intended to provide general recommendations only. Actual configurations in customer environments may need to vary due to individual circumstances, budget constraints, service level agreements, applicable industry-specific regulations, or other factors. Configurations should be tested before implementing them a production environment.

1.4 Customer support

Dell Compellent provides live support at 1-866-EZSTORE (866.397.8673), 24 hours a day, 7 days a week, 365 days a year. For additional support, email Dell Compellent at support@compellent.com. Dell Compellent responds to emails during normal business hours.

2 Introduction

The Dell FS8600 scale-out NAS solution running Fluid File System (FluidFS) version 4 has several built-in features that help businesses protect against unplanned outages and data loss. Downtime and loss of data, even if temporary, can be very costly for an organization. Extended downtime can even be fatal to an organization.

The best practice for protecting data against disasters is typically to create and keep copies of important data in a different location so it can always be recovered. FluidFS v4 running on an FS8600 NAS appliance offers powerful, flexible and easy-to-manage asynchronous replication for disaster recovery.

Below is an example of a possible DR solution. This image shows a DR configuration between four production sites and one DR site. All production sites are replicating volumes to one DR site using FluidFS replication. In case of a production site failure, the storage administrator can recover from the DR site.

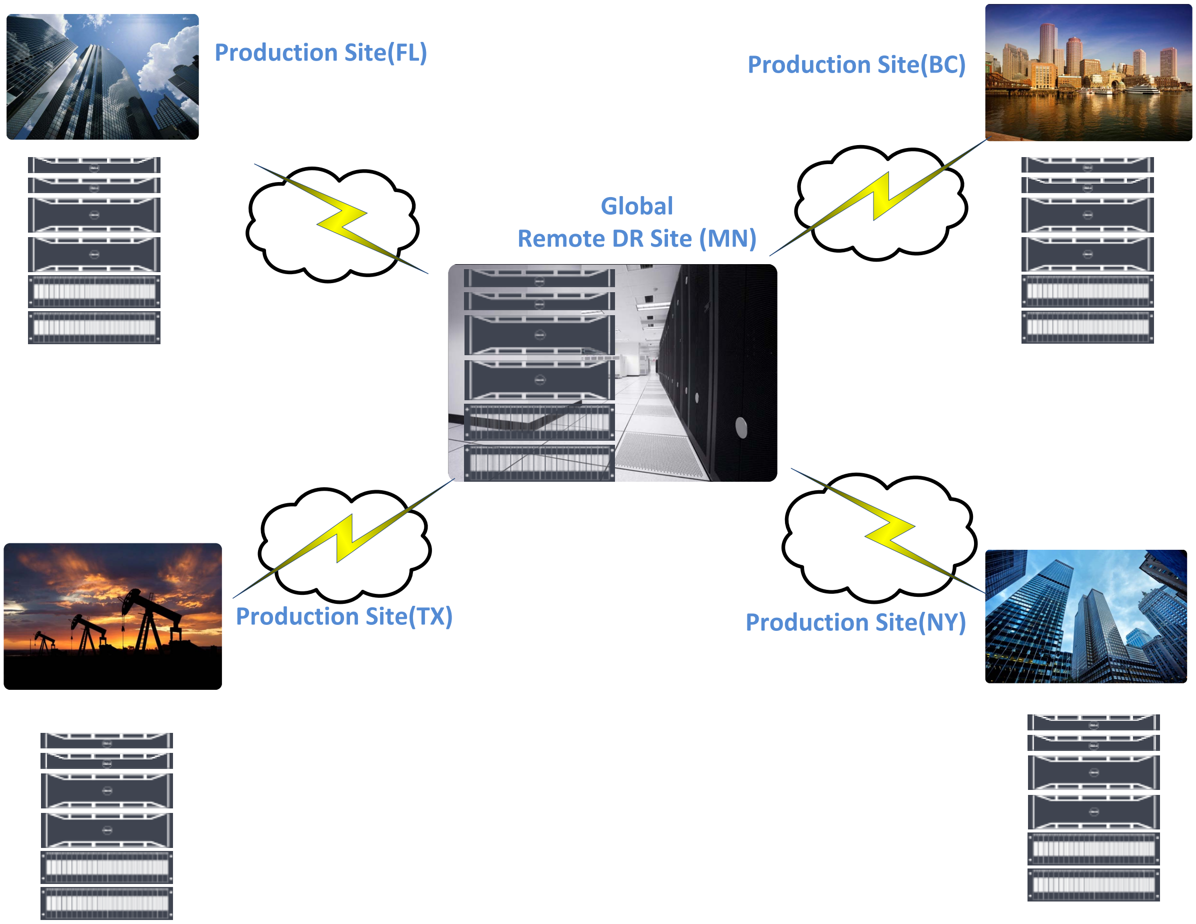


Figure 1 Possible DR solution between four production sites and one centralized DR site.

2.1 FluidFS replication

FS8600 appliances run Dell Fluid File System (FluidFS) to provide file services to clients. Up to four FS8600 NAS appliances can be clustered together to build a single scale-out NAS cluster. FluidFS includes snapshot-based asynchronous replication between two FS8600 NAS clusters. FluidFS does not support utilizing SAN-based snapshots or replication to protect data. While all NAS data is stored on a backing SAN via a FluidFS cluster, all snapshotting and replication of NAS data takes place between FluidFS clusters and not the backing SANs.

FluidFS replication operates at the NAS volume level. A NAS volume may contain one or more SMB shares or NFS exports. An entire volume is replicated to another volume, either on the same NAS cluster or on another remote NAS cluster. The destination volume is an exact replica of the source volumes folders and files, and it can be configured to be an exact replica of the source volumes configuration items, including snapshots, shares, exports, snapshot schedules, and quotas. The exceptions to this rule include data that has been deduplicated and/or the deduplication policy, or if different snapshot retention policies are in use. The destination volume can be configured as a snapshot archive so as to keep snapshots for a time period different than on the primary site. Alternatively, the solution provides flexibility to not keep any snapshots at the DR site at all.

The storage administrator can configure a schedule to run replication, which is independent of any snapshot schedules. When replication runs, it replicates every snapshot consecutively, one after the other, along with the active data, until the destination volume is an exact mirror of the source at the time the replication began.

When a replication is initiated between FluidFS clusters, a temporary snapshot of the current active data is created. This temporary snapshot is used to maintain a consistent data set throughout the replication process and to encapsulate any active data that is not already managed by a previous snapshot. The replication will leverage the delta changes maintained by any interval snapshots that were taken since the previous replication so that only the data changed between replications is moved. FluidFS optimizes block size for data, as low as 4K, depending on data ingest pattern. This optimizes replication due to being able to track changed blocks on a very granular level, reducing replication time and network overhead. Snapshots on the source volume will be replicated in such a way as to be accessible on the destination volume.

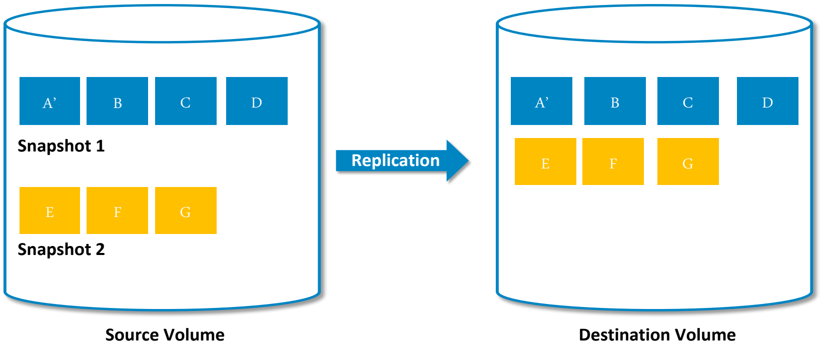


Figure 2 Snapshot based Replication

2.2 FluidFS Replication prerequisites

The following parameters are prerequisites to be used in an FS8600 DR environment. Please make sure that these prerequisites are in place before initiating replication setup.

Firewall ports

FluidFS uses TCP ports 10550-10551 and 10560-10568 for the replication process.The storage administrator needs to make sure these ports are open, to allow traffic between the primary and secondary clusters in a firewalled environment. Additionally, ICMP echo requests (ping) should be allowed between the two clusters.

Networking Connectivity

In order to improve replication performance, FluidFS replicates using all NAS controllers in parallel. Hence, all NAS controllers need to be able to communicate with one another to be able to start the replication process. When configuring the networking on the local/remote site, if possible, make sure that the NAS clusters can ping to the appropriate FluidFS NAS **controller IP's and Virtual IPs (VIPs)** at the remote site. Replication uses the client-facing physical controller IP addresses, as well as the Virtual IPs. As a troubleshooting step, one can also verify ports are open and network connectivity is established by doing a telnet from one FluidFS NAS cluster to the other, using the ports specified above in the "Firewall ports" section.

FluidFS NAS Cluster Sizing (appliance count and disks)

As long as all NAS clusters are at FluidFS v4 or later, the source and destination cluster can have a different number of appliances. However, if the source NAS clusters are at FluidFS v3 or previous, the source and destination cluster must have the same number of controllers/appliances. There is no requirement or limitation around physical capacity deployment or disk enclosure configuration at each site. However, the destination system should have at least as much free space as the source system, and should take into account deduplicated data that will be rehydrated when replicated.

Replication performance and hence the RPO is highly dependent on the overall change rate of the data set, which dictates the amount of data that needs to be replicated at a given time. Replication performance is also highly dependent on the performance of the back-end Storage Centers, the available performance capacity of the NAS clusters themselves and, most importantly, the available bandwidth between replication sources and destinations. FluidFS allows the administrator to configure a NAS volume to be replicated as often as once every minute. Please consult the FluidFS Support Matrix document for the maximum number of supported replications per hour per FluidFS cluster. The FluidFS Support Matrix can be found on the FluidFS TechCenter website, which is referenced in Appendix B – Additional Resources.

Authentication/access control configuration

It is important that the authentication configuration of the systems be identical, including user identity databases (Active Directory, NIS, LDAP), local users, Windows/UNIX user mappings, etc. This is important because if the DR system goes into production in a disaster scenario, the same users must be able to authenticate so they can access data. Otherwise, valuable time can be lost manually configuring authentication on the DR system during the downtime incurred during a disaster.

However, if the secondary system is only intended to be used as a backup, and not for Disaster Recovery, the authentication configuration on FluidFS is optional.

There should be no conflicting Windows-to-Unix user mapping configurations between the source and destination NAS clusters.

Quota configuration

Additionally, to avoid quota conflicts, it is suggested that no user quotas be configured on the destination NAS volume. If user quotas must be enabled at the destination NAS volume, there must be no conflicting quota rules on the source NAS volume. If the NAS volume configuration of the source volume is restored, and there are conflicting quota rules on the destionation volume, this can cause problems. Furthermore, since the replication destination volume is read-only, there is not really a valid use case to put quotas on the replication destination volume. The Dell recommendation is to not use quota rules on replication destination NAS volumes, but if they must be used, they should be kept identical to the quota rules on the source NAS volume.

Fluid Data Reduction and FluidFS NAS Replication

Fluid Data Reduction is a valuable tool that "gives" capacity back to customers through data deduplication as well as compression. Fluid Data Reduction can be used on both the primary and DR cluster, in order to reduce the overall capacity used by NAS data. When NAS data that has been deduplicated or compressed is replicated to a DR cluster, that data is rehydrated before it is sent to the DR cluster. Consequently, the data must be deduplicated/compressed again after it has been replicated to the DR cluster.

3 Replication topologies

FluidFS supports several topologies for replication, both single- and multiple- NAS cluster replication. The following images show all five supported topologies.

3.1 Single NAS cluster replication topologies

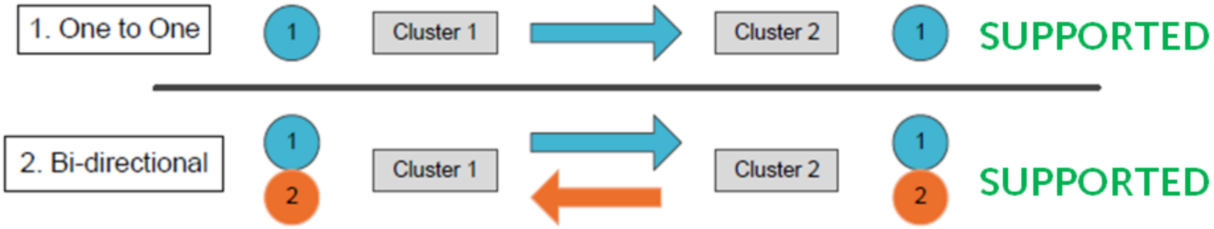


Figure 3 Single NAS cluster replication scenarios

Scenario 1 shows the typical one-to-one disaster recovery (DR) solution. In this scenario the storage administrator chooses Volume 1 in the primary site for replication to the secondary site.

Scenario 2 shows typical bi-directional DR solution when both FS8600 clusters serve as DR clusters. In this scenario the storage administrator chooses to replicate Volume 1 on the primary cluster to the secondary site, and he also chooses to replicate Volume 2 from the secondary site to the primary site. This is a great solution in the event that the data at the secondary site needs to be protected as well as the data on the primary site. However, it is important to note that a single FluidFS cluster cannot contain two sets of SMB home shares, so keep this in mind if bi-directional replication is in use.

3.2 Multiple NAS cluster replication topologies

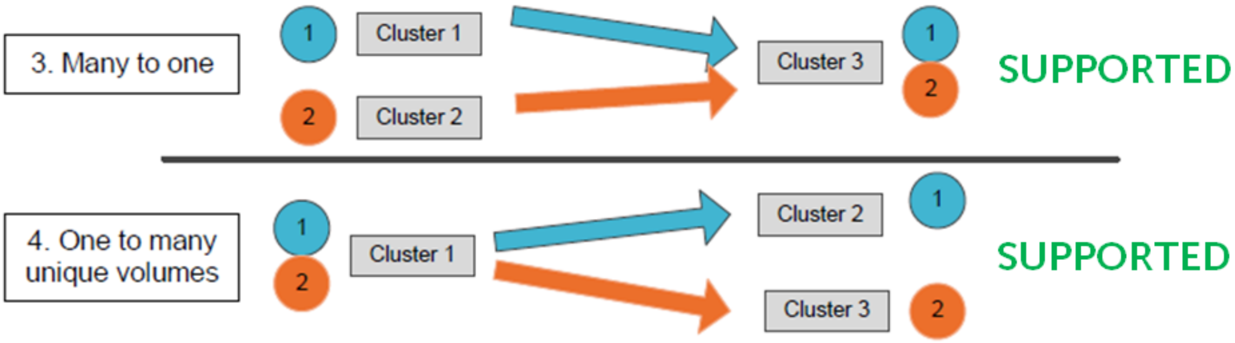


Figure 4 Multiple NAS cluster replication scenarios

Scenario 3 shows a DR solution where two separate NAS clusters replicate to one NAS cluster in a remote site. In this scenario the storage administrator chooses Volume 1 in the NAS Cluster 1 at the primary site to replicate to the secondary site, and the target volume in the remote site is Volume 1.

The storage administrator also chooses Volume 2 in NAS Cluster 2 in the primary site to replicate to the secondary site, and the target volume in the remote site is Volume 2.

Scenario 4 shows a DR solution where one NAS cluster replicates two NAS volumes, one to each of two separate NAS clusters (Volume 1 and Volume 2) at two different DR sites. In this scenario the storage administrator chooses Volume 1 at the primary site to replicate to the FS8600 Cluster 2 at secondary site, and the target volume in the secondary site is Volume 1.

The storage administrator also chooses Volume 2 at the primary site to replicate to the FS8600 Cluster 3 at different a DR site, and the target volume in the secondary site is Volume 2.

Note: The FS8600 DOES NOT support replicating a single volume to multiple target volumes.

3.3 Disaster Recovery Plan and Best Practices

Replicating data may be part of a complete data center disaster recovery plan, but file-based replication alone is not sufficient. A comprehensive DR plan also needs to address aspects of network recovery, application recovery and any other elements of the IT infrastructure that are not under the control of the NAS system. However, making sure an up-to-date copy of the actual data is available is a key element in any disaster recovery plan.

RPO and RTO A typical disaster recovery plan attempts to achieve two goals: a recovery point objective (RPO) and a recovery time objective (RTO).

The RPO is determined by the amount of time the destination system lags behind the source system (note: for the purpose of a disaster recovery plan it is assumed that the source and destination volumes reside on separate systems). For example, if the replication policy is set to run every 5 minutes, then the gap between the volumes will be at best 5-10 minutes. The gap could be more than 5 minutes if the data change rate on primary site is more than what can be transferred to the secondary site in 5 minutes.

The RTO determines the length of time in which the destination system can be activated once it is required. This objective depends on several components outside the FS8600 NAS system — for example, the site DNS — so the RTO is dependent upon the overall disaster recovery plan.

4 Configuring replication

The following section will demonstrate step by step configuration of NAS replication for an FS8600 FluidFS NAS cluster.

4.1 Setup prerequisites

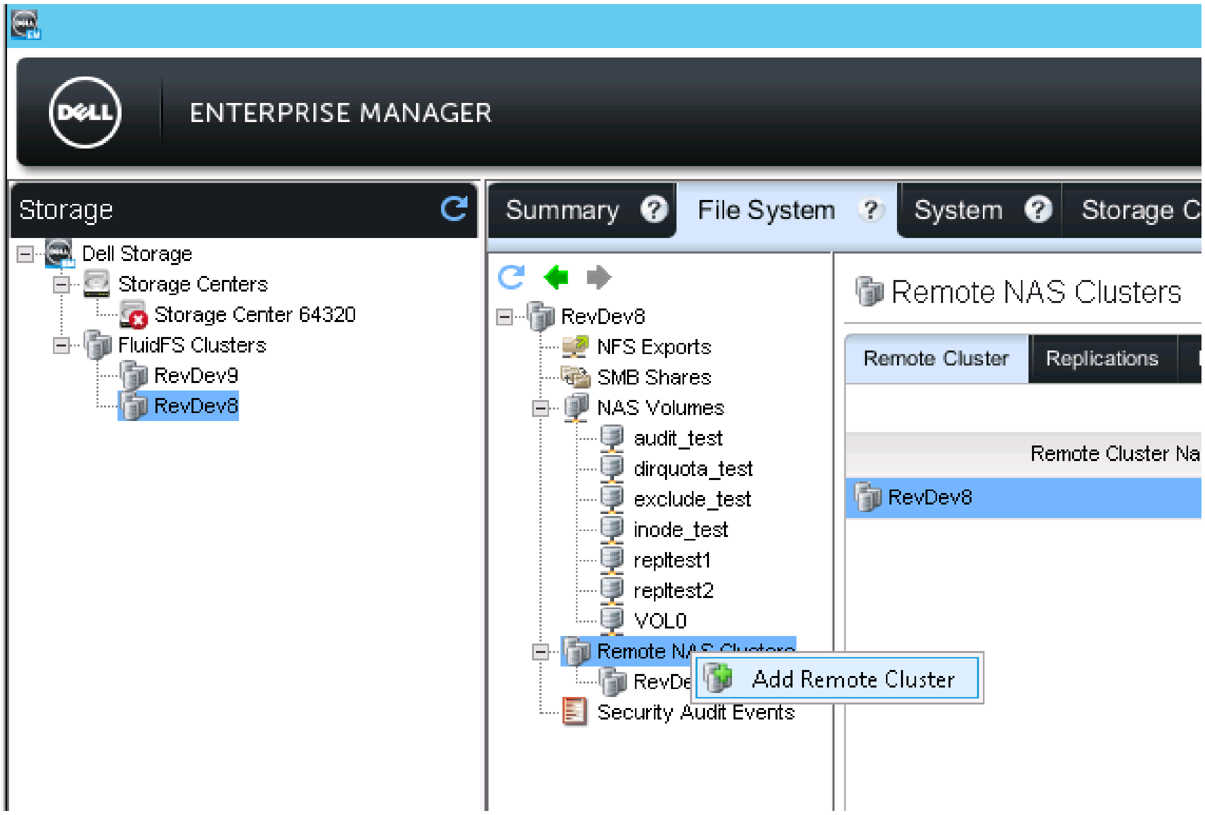
- Network connection between the primary site and the secondary site is already configured and verified to send packets in both directions. See Section 2.2 – FluidFS Replication prerequisites
- The primary and secondary clusters must both be managed by the same EM Data Collector.

4.2 Configuring replication – One-to-one configuration

The following are step-by-step instructions for configuring NAS volume to replicate to a remote site. This example demonstrates replication between two FS8600 NAS clusters. The primary cluster is RevDev8 and the primary volume is VOL0, the secondary cluster is RevDev9 and the secondary volume will be VOL0_REP.

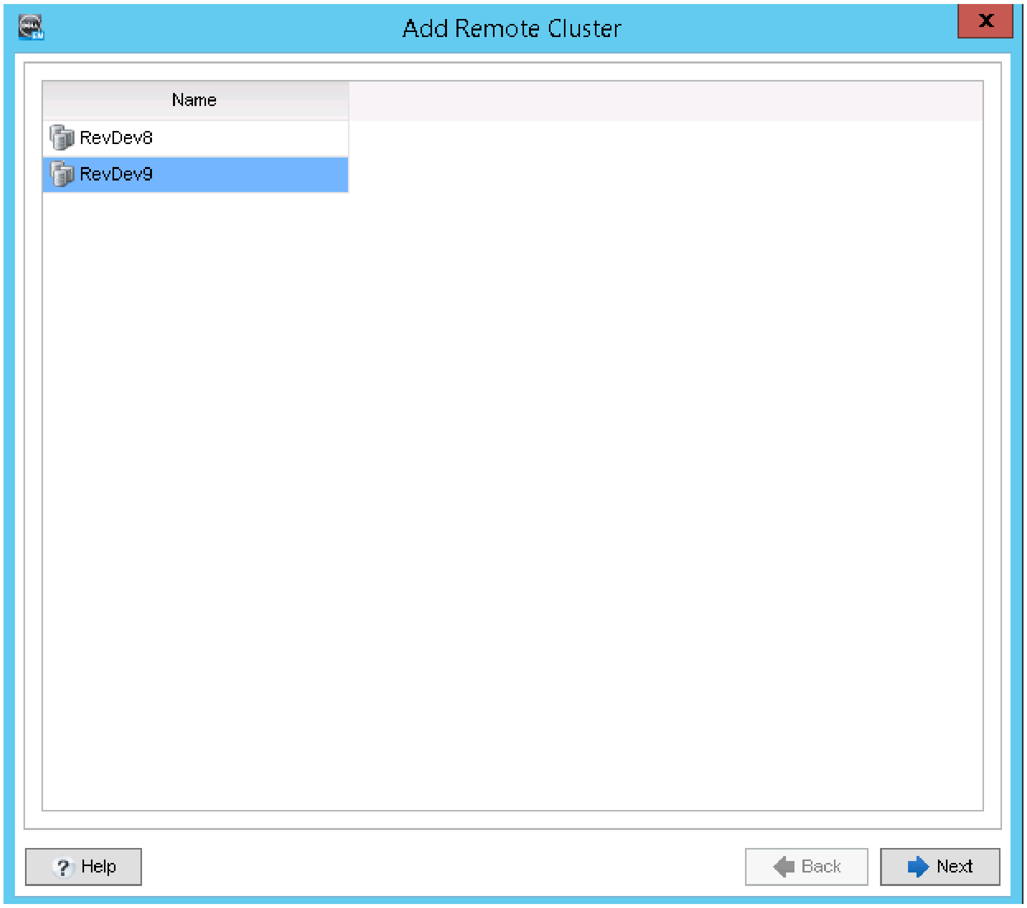
Step 1: Add Remote Cluster

This step creates a partnership between the primary and secondary cluster. From the File System tab, right-click "Remote NAS Clusters" and select "Add Remote Cluster".



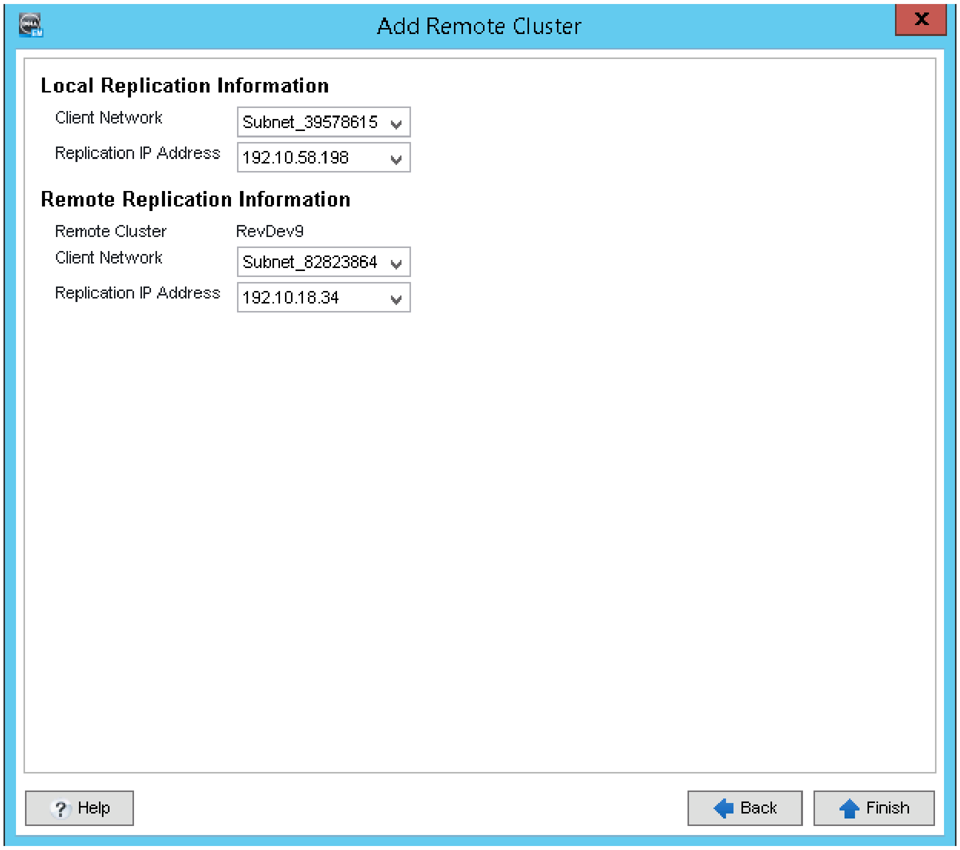
Step 2: Choose Remote Cluster

Choose the remote/secondary cluster from this list, and click next. This list only displays the other FS8600 clusters that are registered under your Enterprise Manager view.



Step 3: Select network to use for replication

Select the client network to use on the primary (local) and secondary (remote) cluster. Some customers prefer to create a subnet (and routes) specifically for replication (only) and this is the screen in which you specify which subnet to use for replication traffic.



Local Replication Information

Client Network: Subnet_39578615

Replication IP Address: 192.10.58.198

Remote Replication Information

Remote Cluster: RevDev9

Client Network: Subnet_82823864

Replication IP Address: 192.10.18.34

Step 4: Verify Partnership on Secondary Cluster

Replication trusts are two way trusts. Once a trust is created from the primary FluidFS cluster, replication policies can be set up from the primary NAS cluster to the secondary NAS cluster, or from the secondary NAS cluster to the primary NAS cluster. It is a one time operation, which applies for both directions of replication.

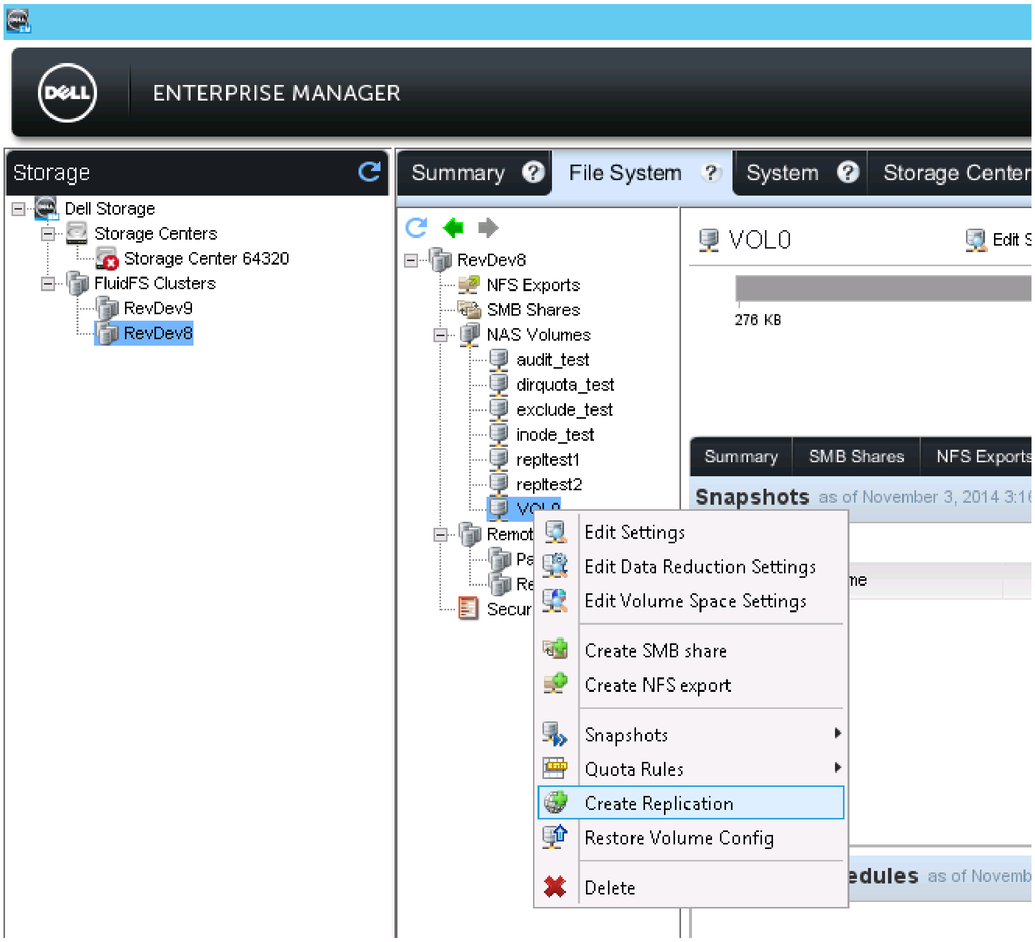
To verify the trust on the secondary cluster, look at the remote (secondary) cluster in EM, and verify the primary cluster is listed as a partner. You can do this by clicking on the secondary cluster in EM, and then look at the list of clusters under “Remote NAS Clusters”.

Step 5: Create Secondary NAS Volume

Create a NAS volume on the secondary cluster (remote cluster) at least as big as the source volume. (on the primary cluster) Alternatively, the volume creation can also be done while setting up the replication in the replication wizard.

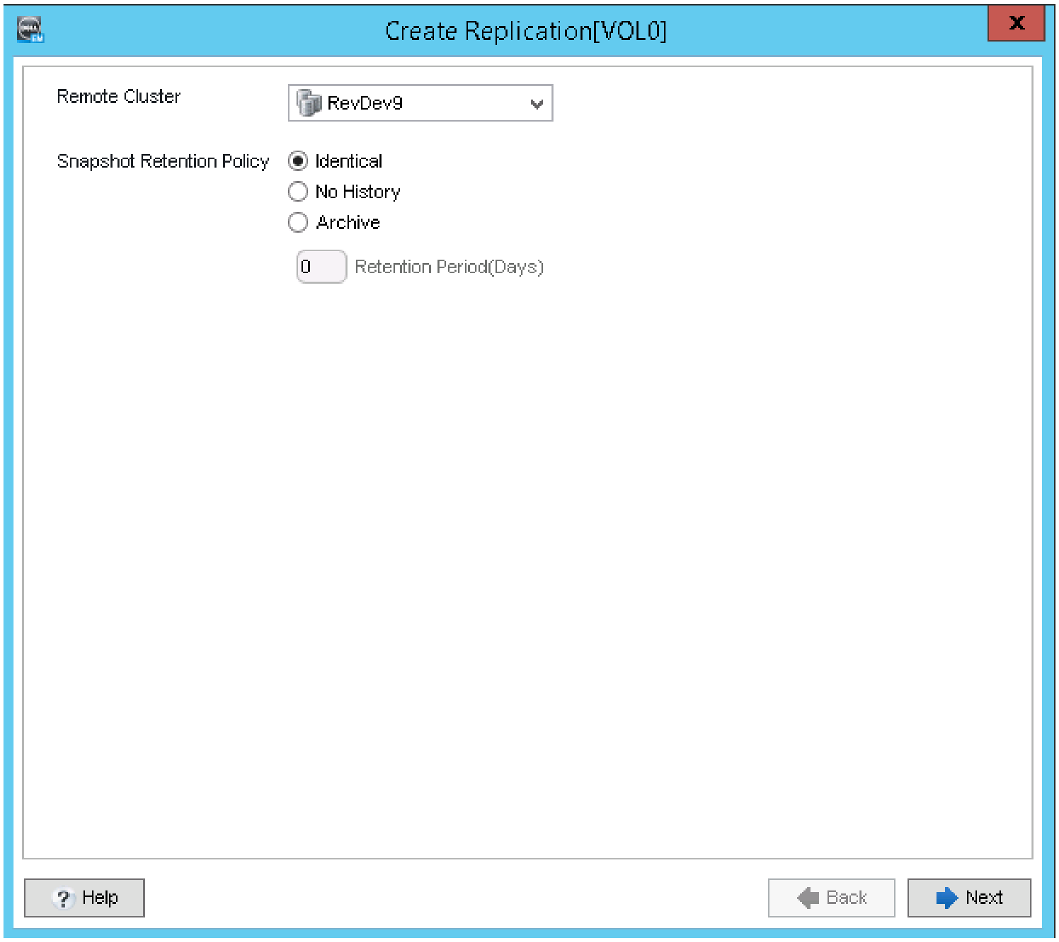
Step 6: Create Replication Policy

On the primary cluster, right-click the volume you would like to replicated to the secondary system, and select "Create Replication"



Step 7: Choose Remote Cluster and Snapshot Retention Policy

Choose the remote NAS cluster (secondary cluster) from the drop-down list, and decide if you would like a different snapshot retention policy on the destination volume (on the secondary cluster), and click Next



Create Replication[VOL0]

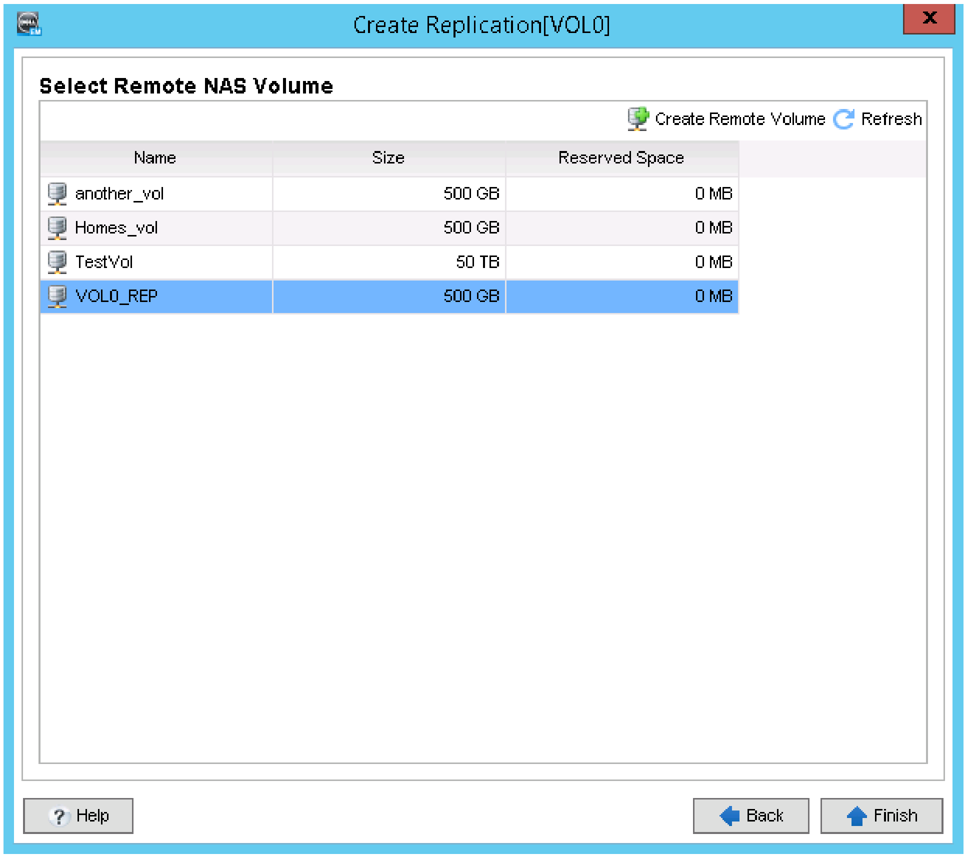
Remote Cluster: RevDev9

Snapshot Retention Policy: ● Identical ○ No History ○ Archive

0 Retention Period(Days)

Step 8: Select Volume to replicate to on secondary (remote) system

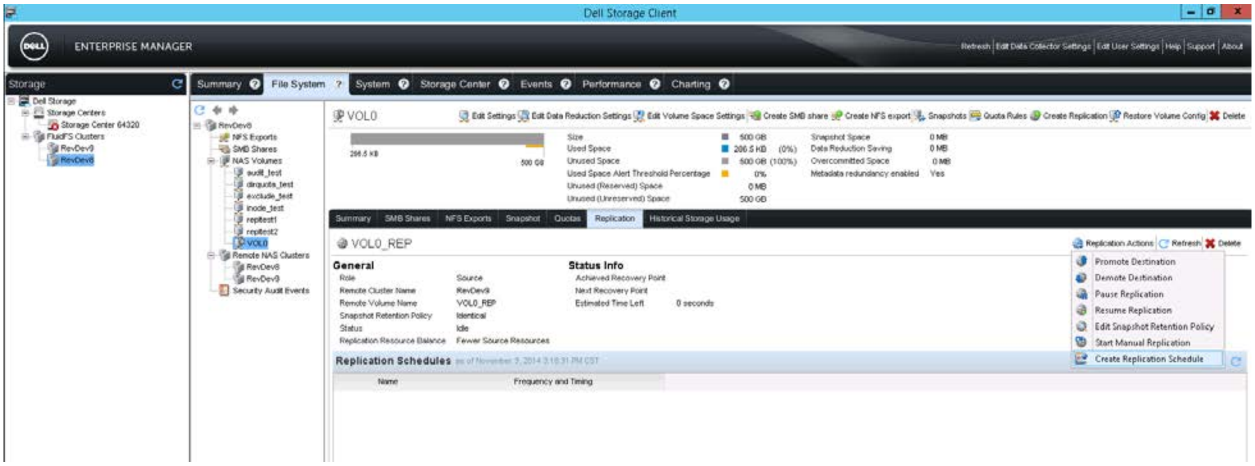
Select the volume previously created on the secondary (remote) NAS cluster and click Finish. If a volume on the secondary cluster has not been created yet, one can be created by clicking the “Create Remote Volume ” button.



Name	Size	Reserved Space
another_vol	500 GB	0 MB
Homes_vol	500 GB	0 MB
TestVol	50 TB	0 MB
VOL0_REP	500 GB	0 MB

Step 9: Create a schedule for the replication

On the primary cluster, click on the NAS volume, and then on the Replication tab. Click the Replication Actions button, and create a replication schedule as shown below. What is being scheduled is a replication "pulse". At each scheduled point in time that is specified in the schedule, FluidFS will replicate all data that is new/changed since the last replication, to the secondary volume. In the case there has never been a replication between these two volumes, or there is not a common "base replica snapshot" present on both systems, all of the data on the primary volume will be replicated to the secondary volume.



Size	500 GB	Snapshot Space	0 MB
Used Space	206.5 KB (0%)	Data Reduction Saving	0 MB
Unused Space	500 GB (100%)	Overcommitted Space	0 MB
Used Space Alert Threshold Percentage	0%	Metadata redundancy enabled	Yes
Unused (Reserved) Space	0 MB		
Unused (Unreserved) Space	500 GB		

General

Role	Source
Remote Cluster Name	RevDev9
Remote Volume Name	VOL0_REP
Snapshot Retention Policy	Identical
Status	Idle
Replication Resource Balance	Fewer Source Resources

Status Info

Achieved Recovery Point	
Next Recovery Point	
Estimated Time Left	0 seconds

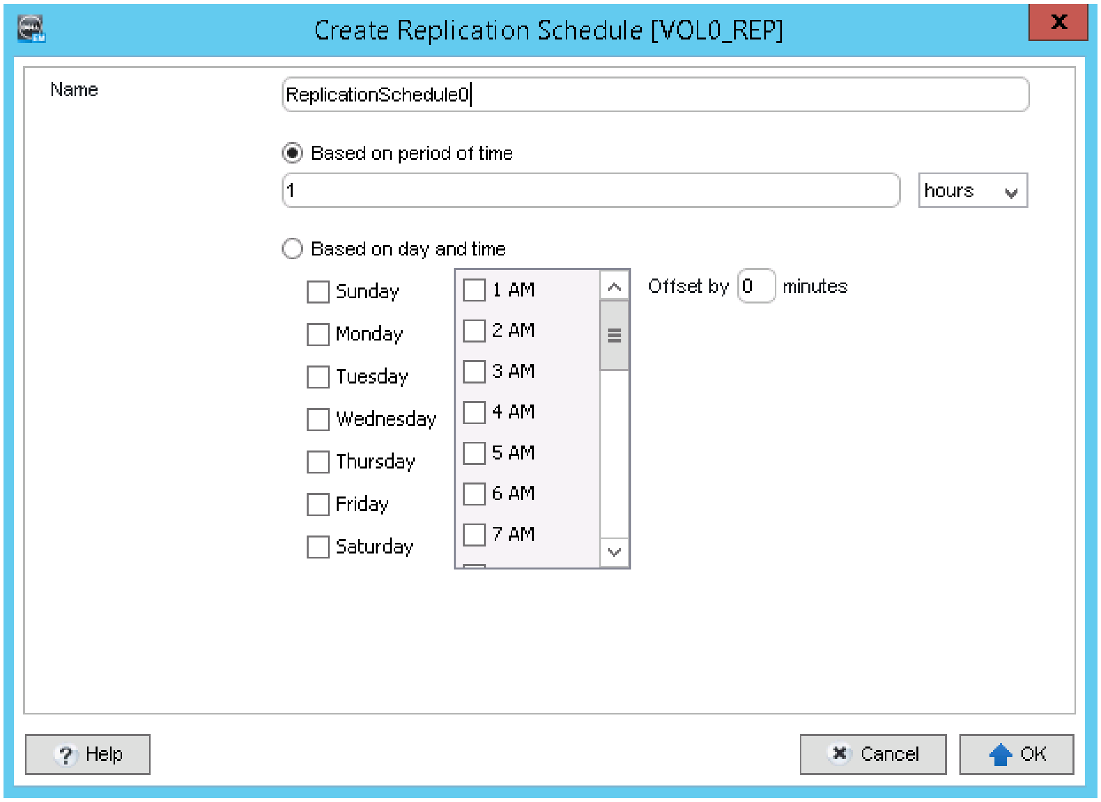
Replication Schedules as of November 3, 2014 3:18:31 PM CST

Name	Frequency and Timing

Step 10: Create the replication schedule and click OK.

The schedule can be configured for once per minute, hour, day, or week.

Alternatively, specific days and times can be set. The “Offset by” field will result in the replications occurring at the selected times, plus the amount of minutes specified. For example, if the administrator wishes not to replicate on the hour, but wishes to replicate 15 minutes past the hour, such as at 1:15AM and 5:15AM and 9:15AM, they would enter “15” into this field, and select 1AM, 5AM, and 9AM. This would result in replicating at 1:15AM, 5:15AM, 9:15AM etc… on the days specified.



Create Replication Schedule [VOL0_REP]

Name: ReplicationSchedule0

● Based on period of time

1 hours

○ Based on day and time

□ Sunday
□ Monday
□ Tuesday
□ Wednesday
□ Thursday
□ Friday
□ Saturday

□ 1 AM
□ 2 AM
□ 3 AM
□ 4 AM
□ 5 AM
□ 6 AM
□ 7 AM

Offset by 0 minutes

4.3 Managing replication policies

Once a replication policy is set, it can always be changed or modified as follows:

- To manage a replication policy (either local or remote), right-click on the replication schedule and select Edit Settings.
- Replicate on-demand lets the storage administrator initiate a replication pulse manually at any given time. This function is primarily used for the initial replication.

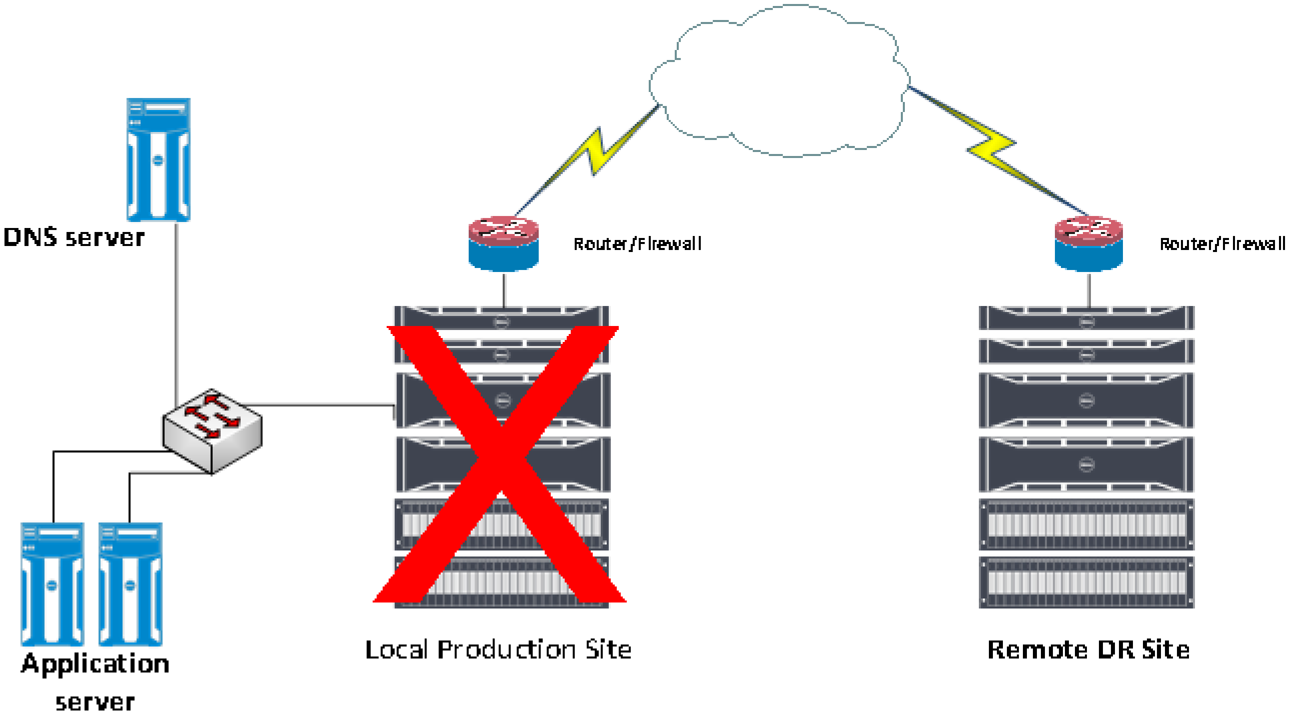
Additional replication tasks:

- The storage administrator can pause/resume a replication, which temporarily disables the replication pulses scheduled by the administrator.
- The storage administrator can promote/demote the destination volume.
  - Promoting a destination volume makes it read+write instead of read-only, and also stop the scheduled replication pulses.
  - Demoting a volume makes it read-only, and allows replication pulses to continue on schedule (or manually initiated).
- The storage administrator can edit the Snapshot Retention Policy for the destination volume.
  - In some cases, the storage administrator wishes for the snapshot retention policy on the destination volume to be different than the source volume.
  - In some cases, there will be less storage space available in the NAS pool of the destination cluster than there is on the source cluster. In such case, space can be conserved by reducing the snapshot overhead on the destination volume. This is accomplished by using either the "No History" or "Archive" mode. Using "No History" mode only replicates the active view of the data to the destination cluster, without any snapshots. Using the "Archive" mode, snapshots can have their expiration reduced so they are expired sooner than they would naturally be expired based on their expiration time set in the snapshot schedule. For example, some administrators configure weekly snapshots to expire after 1 year, but on the destination volume this can consume space un-necessarily. If the administrator wishes to conserve space, the snapshot retention policy can be set to some low value such as 3 days. If there is a failure at the primary site, and the destination volume needs to be promoted and taken into production, as long as this is done within the 3 day window, the snapshot will still be present.
  - In some cases, the administrator will wish to retain snapshots on the destination cluster for a longer period of time than on the source, to maintain large archive of snapshots, to offer more history and granularity in restore points. In this case, "Archive" mode can be used, and set to an extended period of time, such as 90 days or 365 days. Of course, the administrator should verify that ample space is available on the destionation FluidFS cluster to be able to store all of the snapshots.

5 Disaster Recovery Procedure

So far we've discussed the NAS replication feature and how to setup replication. Now we will discuss the necessary steps to recover from a site failure.

The diagram below shows a failed site scenario.



5.1 Failover to secondary site

When the primary site fails, the storage administrator should perform the following steps to failover to the secondary site.

It is important to note that this assumes the secondary cluster is following the best practices outlined in Section 3.3 – Disaster Recovery Plan Best Practices. Namely, the secondary cluster should be configured to use the same DNS environment, NTP, Active Directory, and LDAP/NIS as the primary/production cluster. This is to avoid having to perform these timely configuration tasks while trying to recover from an outage.

Note: This procedure must be performed for each NAS volume that the administrator wishes to fail over. In cases where there is a high number of volumes in use, this procedure can be scripted against the FluidFS Command Line Interface. This is covered in Appendix A of this document.

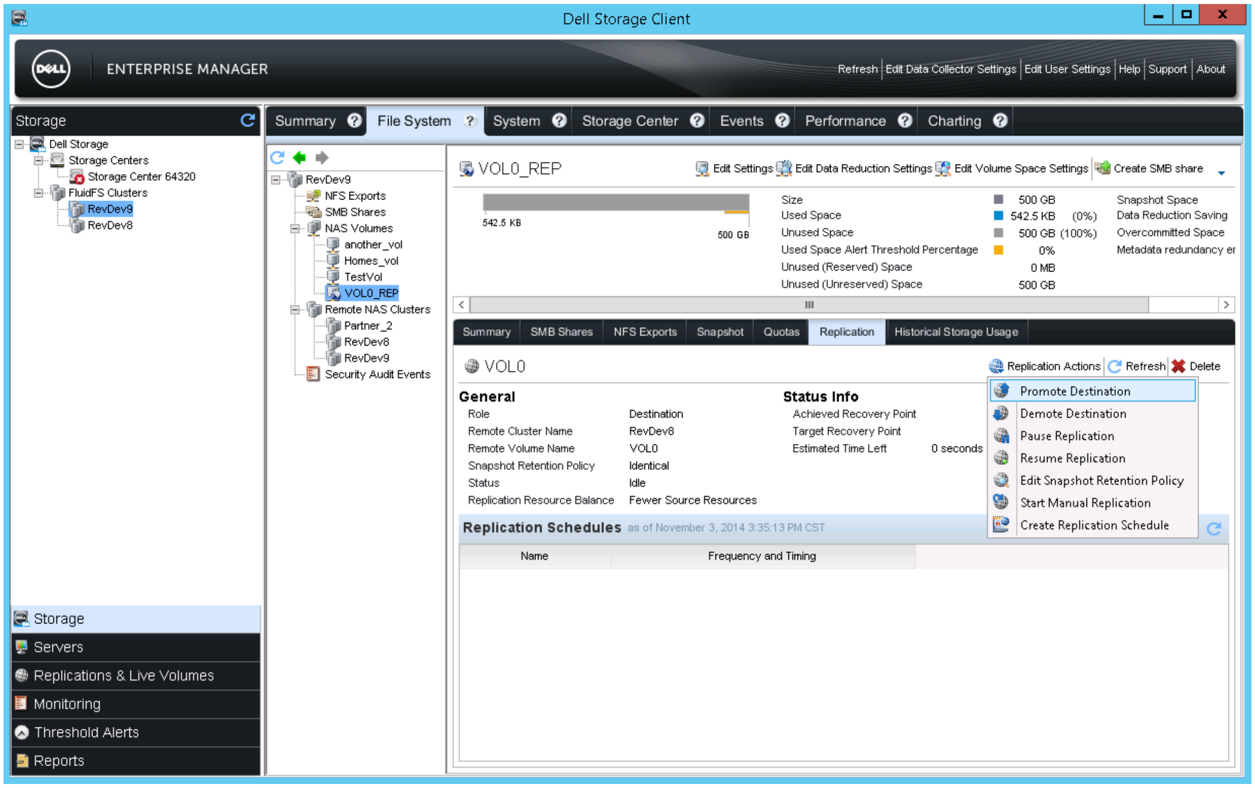
The high level steps of this process are as follows:

1. Promote the destination NAS volume on the secondary cluster. This will make the destination volume Read+Write and will present the latest stable replica, which is the snapshot taken during the last replication. This snapshot name always starts with "rep"
2. Restore the source NAS volume configuration on the secondary cluster. This will recreate all the SMB shares, NFS exports, snapshot schedule, and quota rules that were present on the source NAS volume.
3. Restore/Verify cluster wide configuration items
4. Redirect client systems to the secondary cluster
5. Delete the replication on the destination NAS volume on the secondary cluster.

The detailed steps of how to perform this procedure are as follows:

Step 1: Promote the secondary volume (on secondary NAS cluster)

Promote the secondary (remote) site volume by clicking on the volume and then selecting “Promote Destination”. This will make the secondary volume read/write instead of read-only.



Size	▪ 500 GB	Snapshot Space
Used Space	■ 542.5 KB (0%)	Data Reduction Saving
Unused Space	▪ 500 GB (100%)	Overcommitted Space
Used Space Alert Threshold Percentage	■ 0%	Metadata redundancy er
Unused (Reserved) Space	0 MB	
Unused (Unreserved) Space	500 GB	

General

Role	Destination
Remote Cluster Name	RevDev8
Remote Volume Name	VOL0
Snapshot Retention Policy	Identical
Status	Idle
Replication Resource Balance	Fewer Source Resources

Status Info

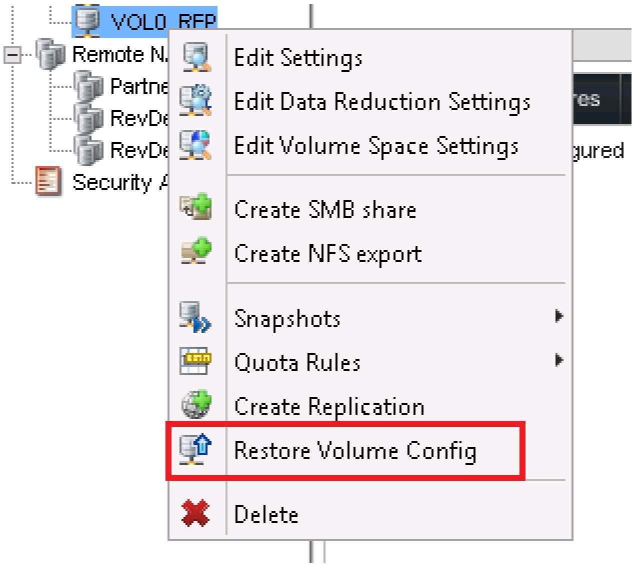
Achieved Recovery Point	
Target Recovery Point	
Estimated Time Left	0 seconds

Replication Schedules as of November 3, 2014 3:35:13 PM CST

Name	Frequency and Timing

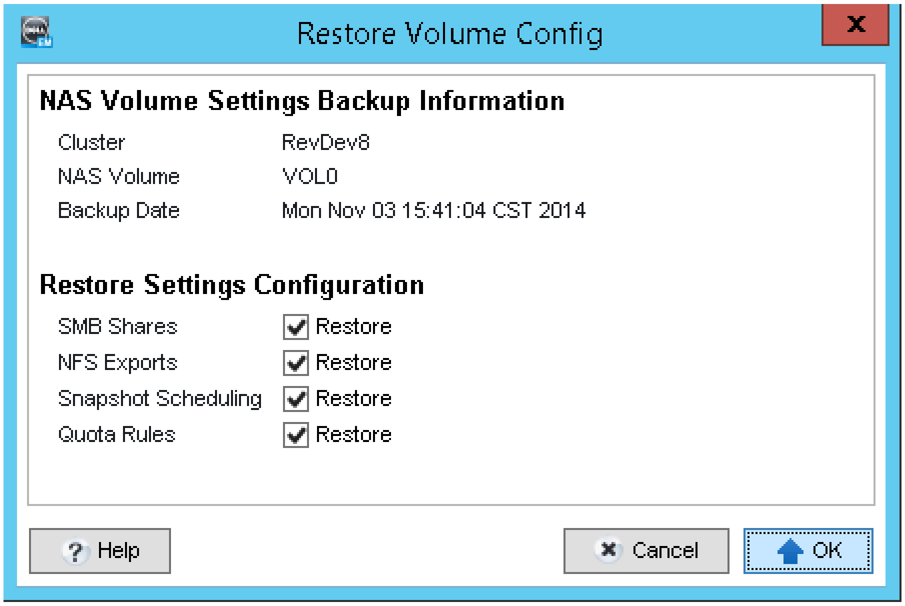
Step 2: Right-click on the remote site NAS volume and click “Restore Volume Config“ (on secondary NAS cluster)

This will restore the shares, exports, snapshot schedules, and quota rules that were present on the primary site volume.



Step 3: Restore NAS Volume Configuration (on secondary NAS cluster)

By default, all available configuration files are checked. Click “OK”



NAS Volume Settings Backup Information

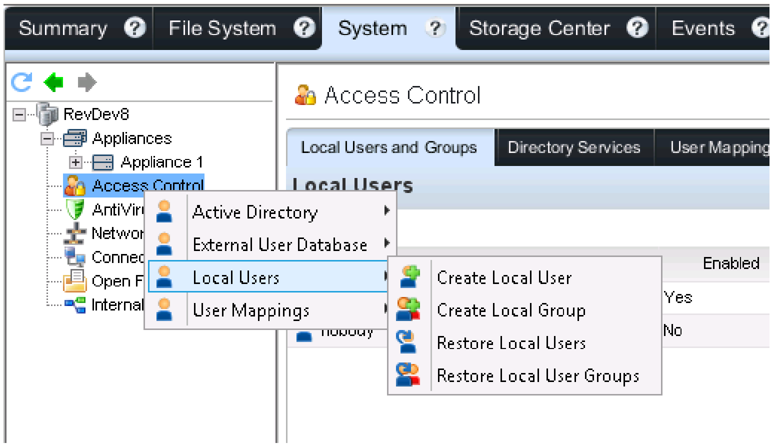
Cluster	RevDev8
NAS Volume	VOL0
Backup Date	Mon Nov 03 15:41:04 CST 2014

Restore Settings Configuration

SMB Shares	☑ Restore
NFS Exports	☑ Restore
Snapshot Scheduling	☑ Restore
Quota Rules	☑ Restore

Step 4: Restore other cluster-wide configuration items (on secondary NAS cluster)

If best practices are being followed, the secondary cluster will already be configured with all environmental services such as DNS, NTP, Active Directory, LDAP/NIS, etc… But in the case that it is not, these must be configured at this stage. Additionally, if local users/groups, or manual user mappings are in place, those must be manually restored. Dell recommends keeping the local users on both FluidFS clusters identical at all times. Local users and groups can be restored through Enterprise Manager:



Any manual user mapping rules that are defined must be restored using the FluidFS Command Line Interface using the command "access-control mapping manual restore". Any networking configuration changes will have to be made manually. FluidFS does not have any automated method to restore network settings (such as IP's or static routes).

NOTE: Dell recommends as a best practice to document all network settings (IP addresses, default gateway, static routes) so that if needed they can be restored onto the secondary site.

Step 5: Repoint clients to the secondary NAS cluster/NAS volume

At this point, the storage admin has two options (use one of the two options below) of how to point clients to the secondary FluidFS cluster:

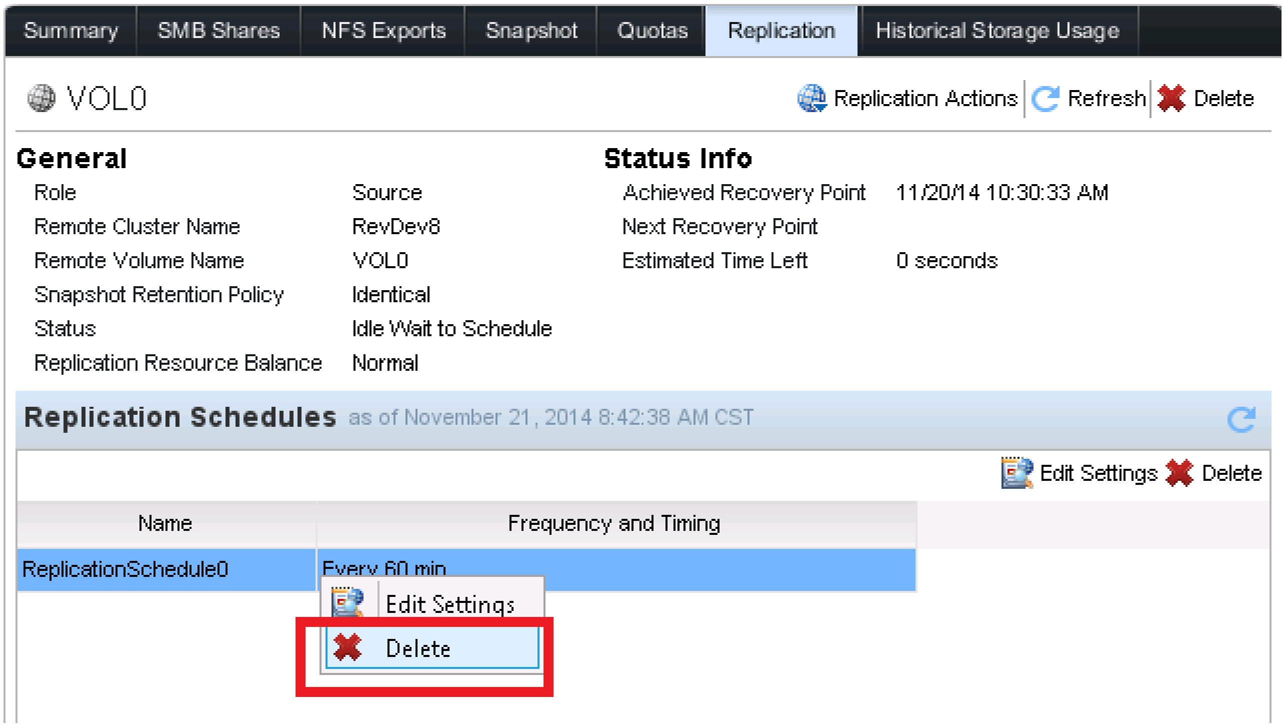
1. Change the DNS entry that originally pointed to the Virtual IPs of the primary site to point to the secondary cluster's Virtual IPs. Ensure that the DNS server(s) that the secondary NAS cluster is using is the same DNS server(s) (or in the same DNS farm) as the DNS server(s) the primary NAS cluster is using.Exising client connections will break and need to be re-established. All NFS exports must be unmounted and mounted back on every client system.

   Note: Lowering the TTL for these DNS entries could help with DNS cache issues when failing back.
2. Change/Add the IP's on the secondary cluster to be identical to the primary NAS clusters IPs. Choosing this option allows for NFS hosts to not have to unmount/remount all of the NFS exports.

Now, users and applications are working off the secondary NAS cluster. However, this is not a permanent state, and eventually the storage administrator will want to fail back to the primary site. But first, the replication schedule and policy must be deleted, so it can be recreated in the opposite direction, replicating the secondary site to the primary.

Step 6: Delete the Replication Schedule (on the secondary NAS cluster)

FluidFS requires that the replication schedule is deleted before the replication policy can be deleted.



General

Role	Source
Remote Cluster Name	RevDev8
Remote Volume Name	VOL0
Snapshot Retention Policy	Identical
Status	Idle Wait to Schedule
Replication Resource Balance	Normal

Status Info

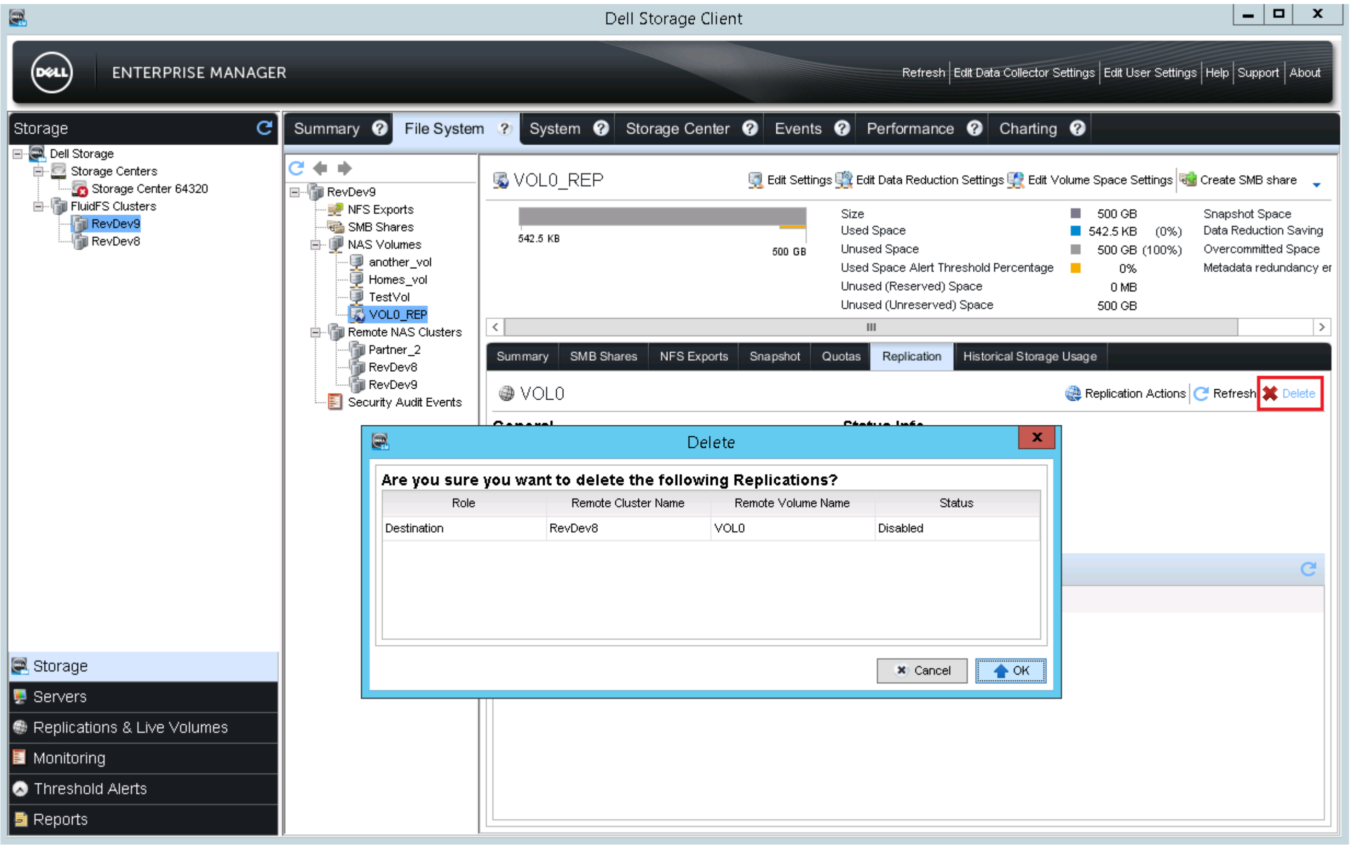
Achieved Recovery Point	11/20/14 10:30:33 AM
Next Recovery Point	
Estimated Time Left	0 seconds

Replication Schedules as of November 21, 2014 8:42:38 AM CST

Name	Frequency and Timing
ReplicationSchedule0	Every 60 min

Step 7: Delete the Replication Policy (on secondary NAS cluster)

The replication policy must be deleted, because the direction needs to be reversed to replicate the secondary site to the primary site. You can “Disconnect” the replication policy between the volumes by doing a **Delete** from Enterprise Manager.



VOL0_REP

Size	500 GB		Snapshot Space
Used Space	542.5 KB	(0%)	Data Reduction Saving
Unused Space	500 GB	(100%)	Overcommitted Space
Used Space Alert Threshold Percentage	0%		Metadata redundancy er
Unused (Reserved) Space	0 MB		
Unused (Unreserved) Space	500 GB		

VOL0

Delete

Are you sure you want to delete the following Replications?

Role	Remote Cluster Name	Remote Volume Name	Status
Destination	RevDev8	VOL0	Disabled

Cancel OK

5.2 Failback to primary site

This procedure, at a high level, is what is covered in the previous section (Section 5.1 – Failover to Secondary Site), in reverse.

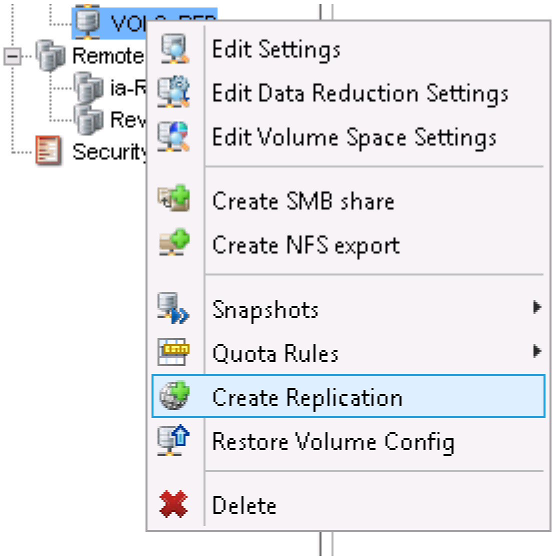
The failback from the secondary site to the primary site should occur only after all issues are fixed, and the primary site is ready to take the full production workload back.

The procedure to fail back to the primary site follows these high level steps:

1. Choose a period of low activity to fail back to the primary site.
2. Set up replication to replicate the NAS volume on the secondary NAS cluster back to the primary NAS cluster
3. Halt IO to the NAS volume on the secondary cluster by deleting shares or moving to NoService mode
4. Promote the NAS volume on the primary cluster, and delete the replication policy
5. Restore the NAS volume configuration on the NAS volume on the primary NAS cluster to recreate all shares, exports, snapshot schedules, and quota rules
6. Redirect clients to the primary NAS cluster
7. Re-establish replication from the primary site back to the secondary site

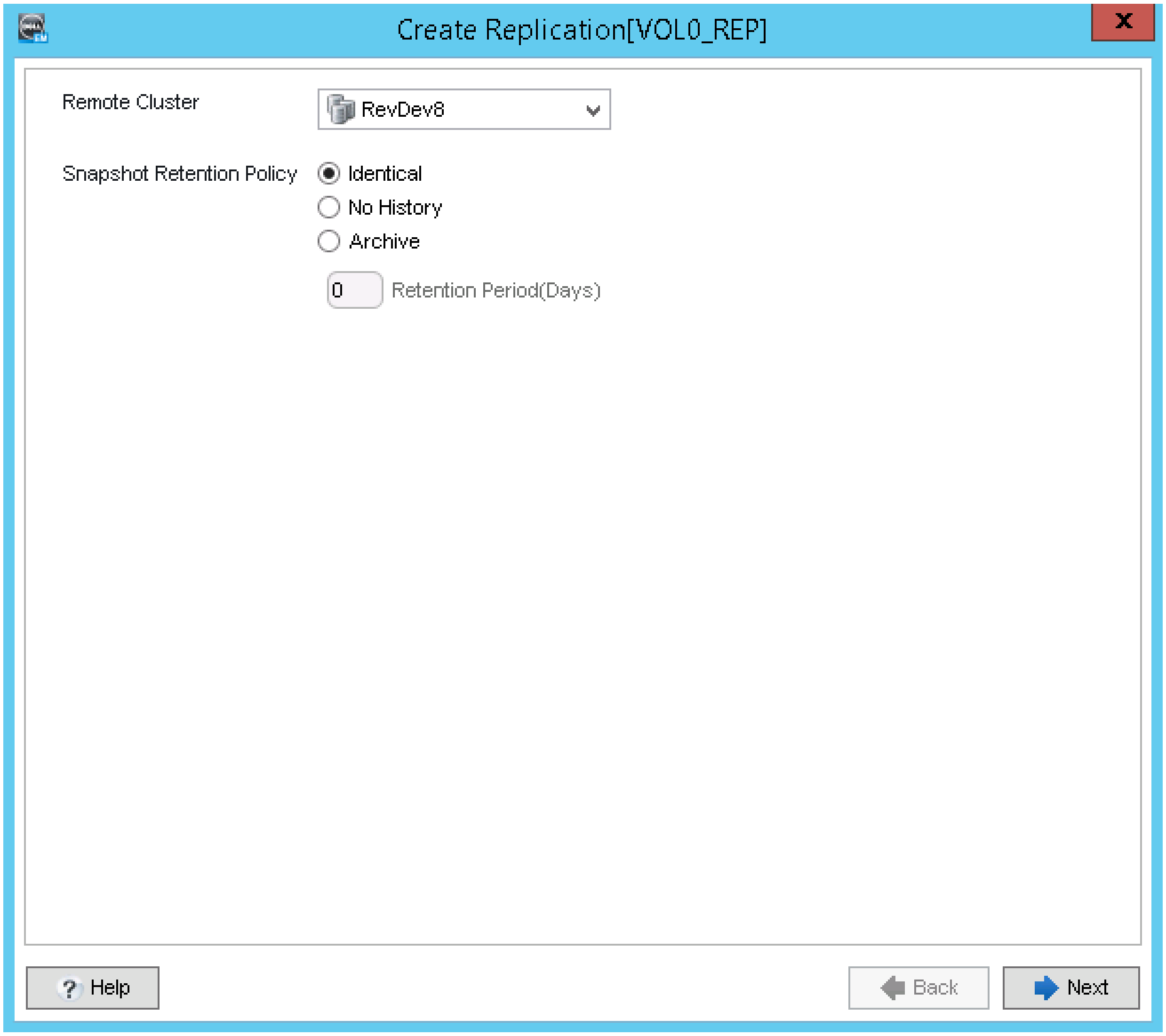
Step 1: Create Replication (on secondary NAS cluster)

On the secondary NAS cluster, right-click the secondary NAS volume and select "Create Replication"



Step 2: Choose Remote Cluster and Snapshot Retention Policy (on secondary NAS cluster)

Choose the remote NAS cluster (primary site NAS cluster) from the drop-down list, and click Next. Typically in a failback scenario the best choice is “Identical”. If any uneeded snapshots are replicated this will only provide more protection of data. Snapshots can always be deleted from the source prior to replication, or the destination NAS volume after replication finishes.



Create Replication[VOL0_REP]

Remote Cluster: RevDev8

Snapshot Retention Policy:
● Identical
○ No History
○ Archive

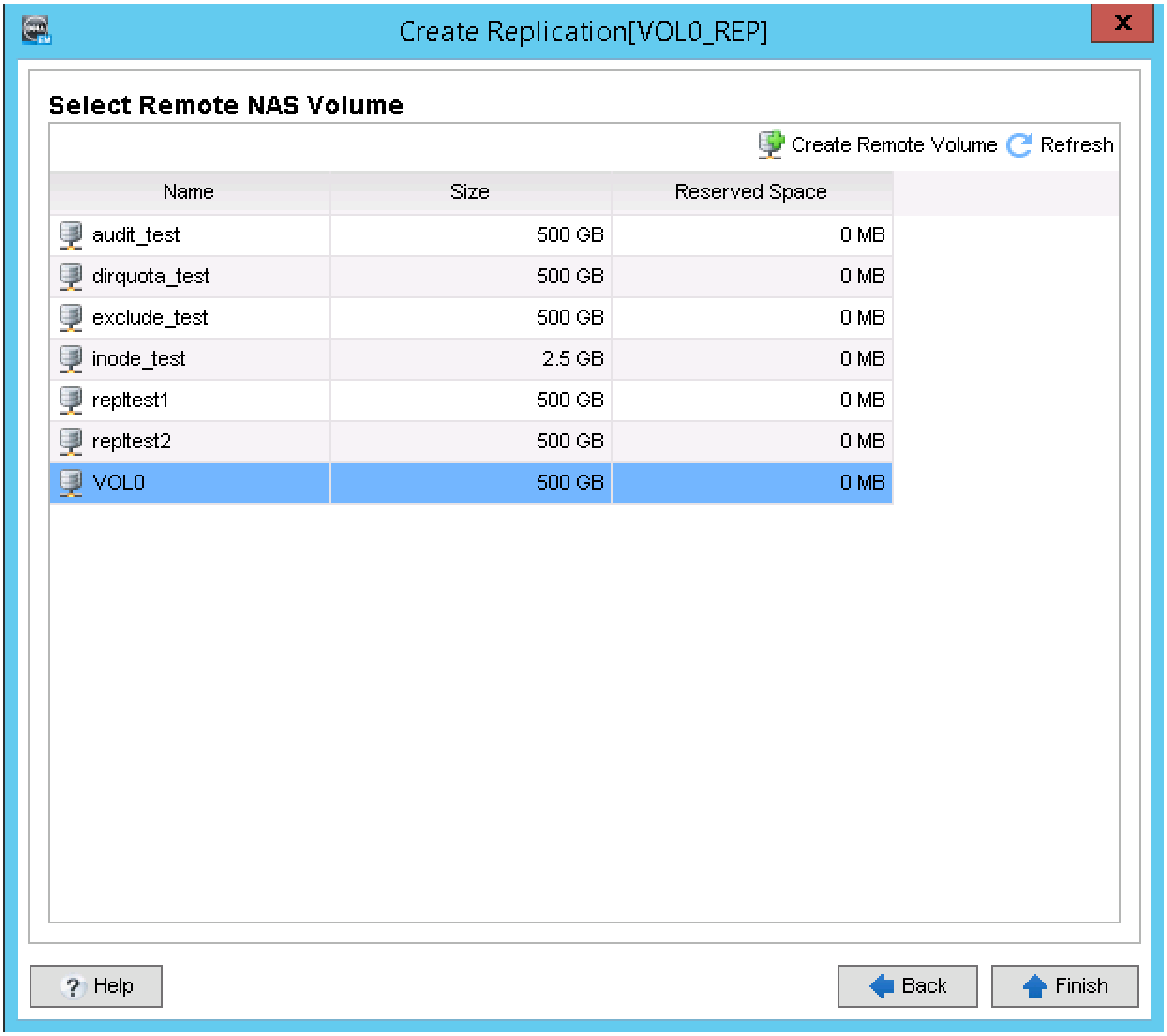
0 Retention Period(Days)

Step 3: Select Volume (on secondary NAS cluster)

Select the volumes previous source NAS volume, and click Finish.

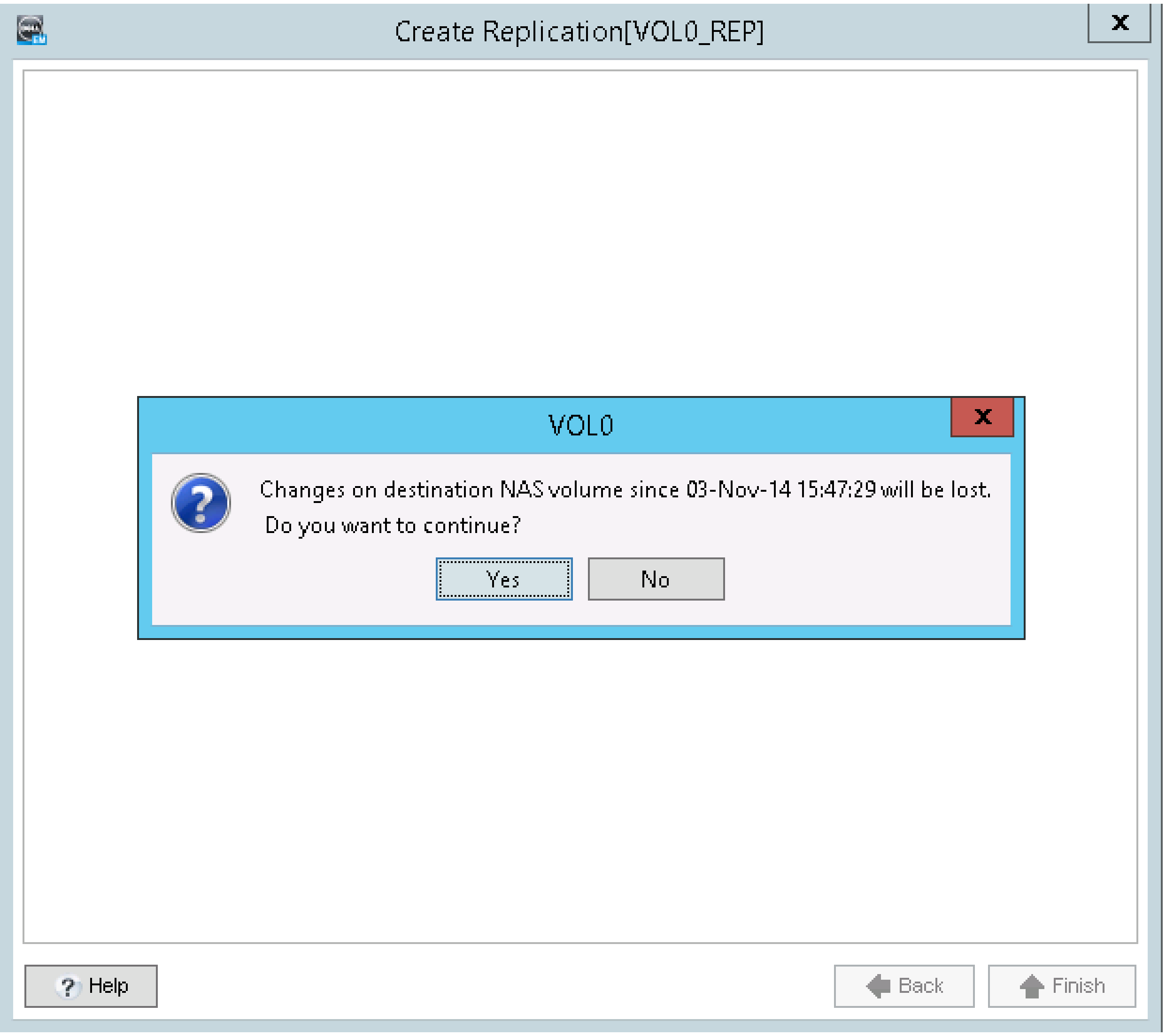
If the original primary NAS volume is still present, FluidFS will only replicate the changed/new data from the secondary site back to the primary site. However, any new or changed data on the primary site, that has been changed or written after the last replication between the two NAS volumes, **will be lost.** The secondary NAS volume uses the "base replication snapshot" that is kept from the last replication, which is identical on the primary and secondary NAS volume, in order to establish a baseline between the two NAS volumes.

If the original primary NAS volume is no longer present, FluidFS will replicate the entire NAS volume back to the primary site. Please note that this initial replication can take a long time, depending on the amount of data that needs to be replicated from the secondary site back to the primary,



Name	Size	Reserved Space
audit_test	500 GB	0 MB
dirquota_test	500 GB	0 MB
exclude_test	500 GB	0 MB
inode_test	2.5 GB	0 MB
repltest1	500 GB	0 MB
repltest2	500 GB	0 MB
VOL0	500 GB	0 MB

When failing back, when the original source NAS volume is selected, this message will appear below. Any changes that took place on the original source NAS volume since the last replication will be lost.



Changes on destination NAS volume since 03-Nov-14 15:47:29 will be lost.
Do you want to continue?

Step 4: Initiate a Manual replication from the secondary NAS volume to the primary NAS volume (on secondary NAS cluster)

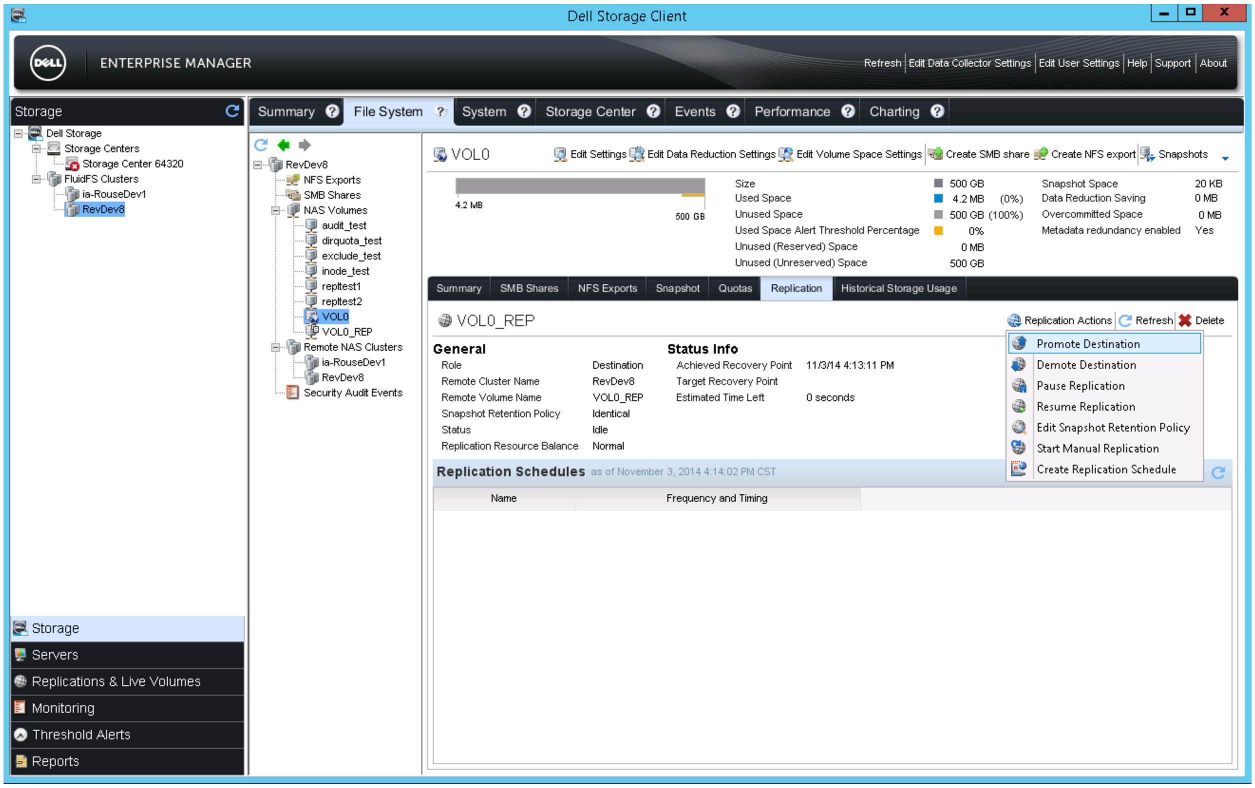
Step 5: Wait for the replication to finish

Ideally, the failback procedure will take place during off hours when there is minimal I/O to the secondary NAS cluster. Manual replication can be repeatedly triggered to synchronize changes, or a replication schedule can be set up if desired.

Step 6: Promote the primary NAS volume (on primary or secondary NAS cluster)

> NOTE: Once the primary site NAS volume is promoted, it will become read+write, and replications will stop. In order to avoid data loss, its important to halt all I/O to the volume on the secondary NAS cluster, so the data matches on the secondary NAS volume and the primary NAS volume.
>
> I/O can be halted on the secondary NAS cluster by deleting shares/exports, or changing the filesystem into NoService mode.

Promote the primary site NAS volume by clicking on it and then selecting “Promote Destination”. This will make the primary NAS volume read/write instead of read-only.


Dell Storage Client
ENTERPRISE MANAGER
Refresh | Edit Data Collector Settings | Edit User Settings | Help | Support | About
Storage
Dell Storage
Storage Centers
Storage Center 64320
FluidFS Clusters
ia-RouseDev1
RevDev8
Summary | File System | System | Storage Center | Events | Performance | Charting
RevDev8
NFS Exports
SMB Shares
NAS Volumes
audit_test
dirquota_test
exclude_test
inode_test
repltest1
repltest2
VOL0
VOL0_REP
Remote NAS Clusters
ia-RouseDev1
RevDev8
Security Audit Events
Edit Settings | Edit Data Reduction Settings | Edit Volume Space Settings | Create SMB share | Create NFS export | Snapshots
4.2 MB
500 GB
Summary | SMB Shares | NFS Exports | Snapshot | Quotas | Replication | Historical Storage Usage
Replication Actions | Refresh | Delete
Promote Destination
Demote Destination
Pause Replication
Resume Replication
Edit Snapshot Retention Policy
Start Manual Replication
Create Replication Schedule
Storage
Servers
Replications & Live Volumes
Monitoring
Threshold Alerts
Reports

VOL0

Size	500 GB	Snapshot Space	20 KB
Used Space	4.2 MB (0%)	Data Reduction Saving	0 MB
Unused Space	500 GB (100%)	Overcommitted Space	0 MB
Used Space Alert Threshold Percentage	0%	Metadata redundancy enabled	Yes
Unused (Reserved) Space	0 MB		
Unused (Unreserved) Space	500 GB		

VOL0_REP

General

Role	Destination
Remote Cluster Name	RevDev8
Remote Volume Name	VOL0_REP
Snapshot Retention Policy	Identical
Status	Idle
Replication Resource Balance	Normal

Status Info

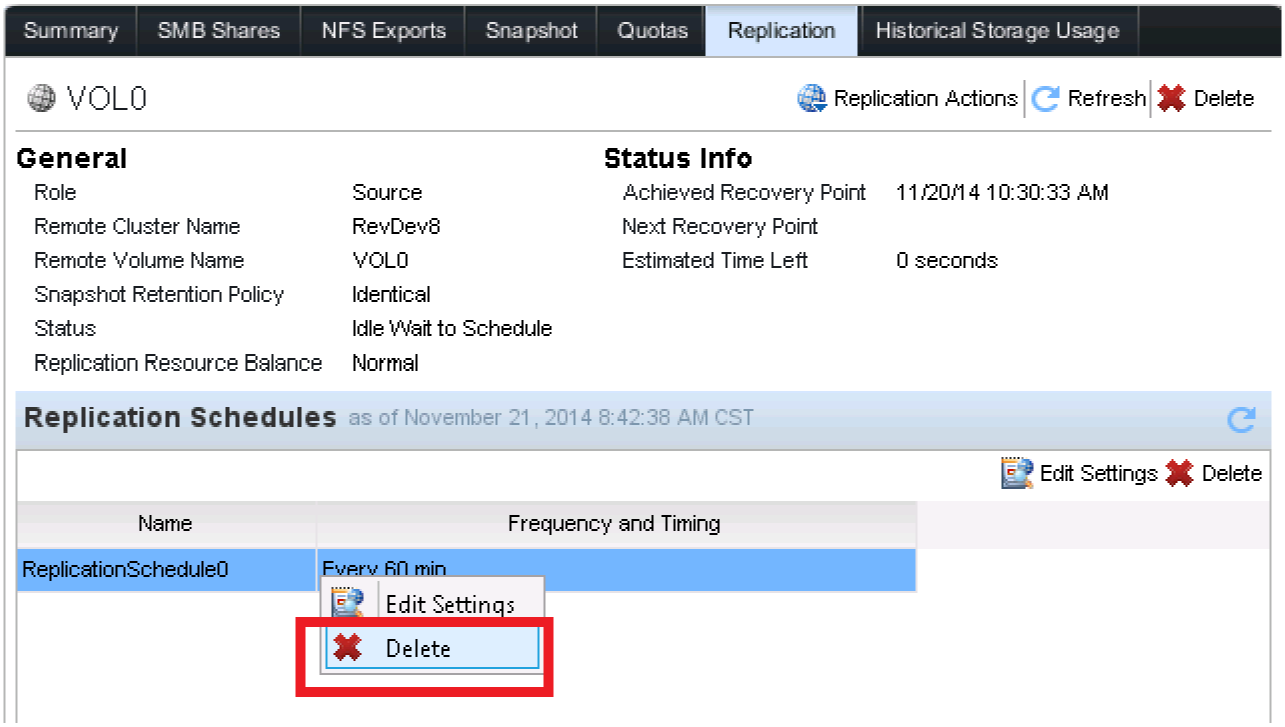
Achieved Recovery Point	11/3/14 4:13:11 PM
Target Recovery Point	
Estimated Time Left	0 seconds

Replication Schedules as of November 3, 2014 4:14:02 PM CST

Name	Frequency and Timing

Step 7: Delete the Replication Schedule (on secondary NAS cluster)

FluidFS requires that the replication schedule is first deleted, before the Replication Policy can be deleted.



VOL0

General

Role	Source
Remote Cluster Name	RevDev8
Remote Volume Name	VOL0
Snapshot Retention Policy	Identical
Status	Idle Wait to Schedule
Replication Resource Balance	Normal

Status Info

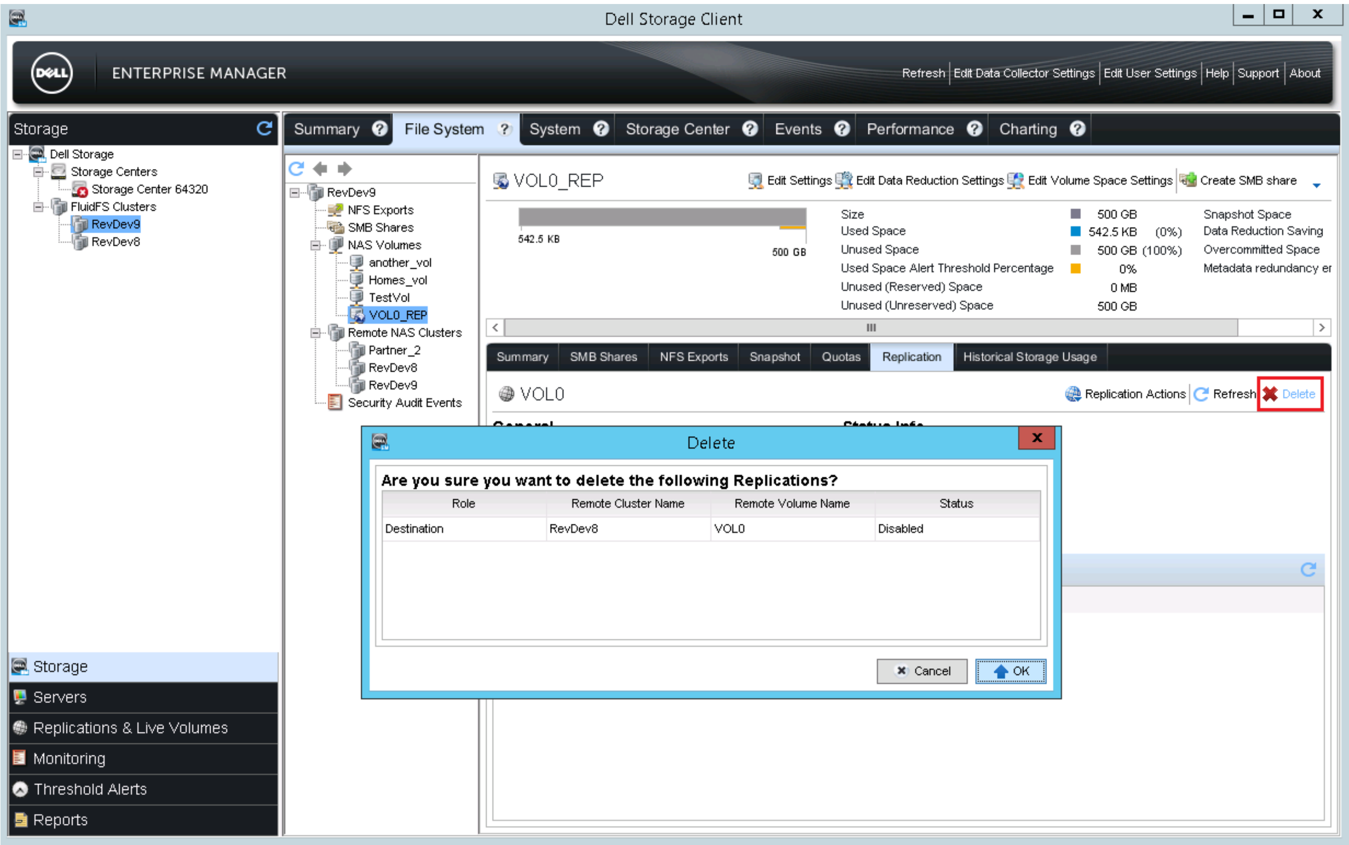
Achieved Recovery Point	11/20/14 10:30:33 AM
Next Recovery Point	
Estimated Time Left	0 seconds

Replication Schedules as of November 21, 2014 8:42:38 AM CST

Name	Frequency and Timing
ReplicationSchedule0	Every 60 min

Step 8: Delete the Replication Policy (on secondary NAS cluster)

After promoting the primary site NAS volume, you can “Disconnect” the replication policy between the NAS volumes by doing a **Delete** from Enterprise Manager.

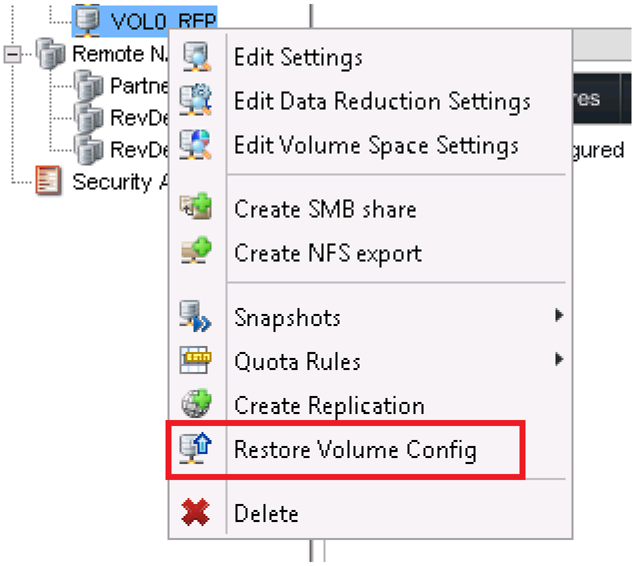


Are you sure you want to delete the following Replications?

Role	Remote Cluster Name	Remote Volume Name	Status
Destination	RevDev8	VOL0	Disabled

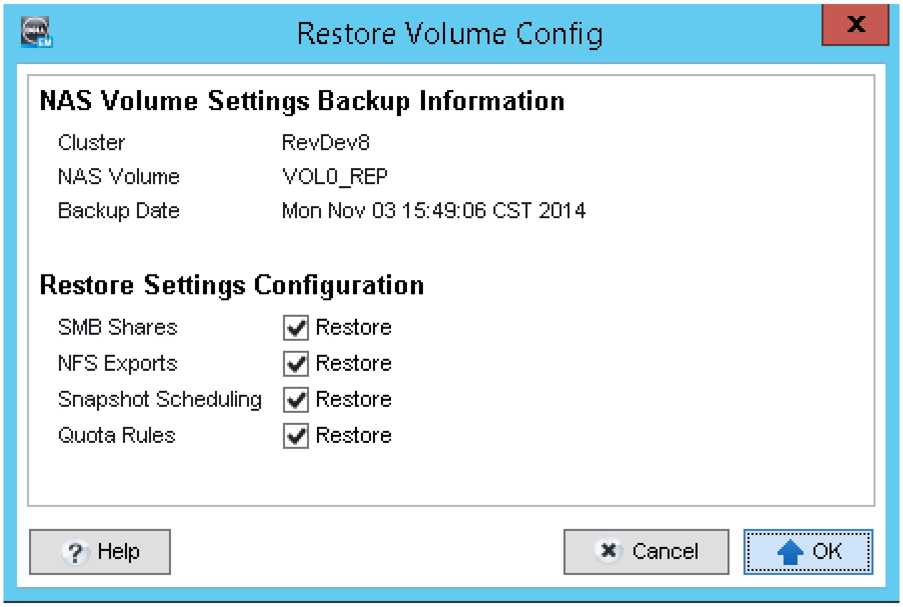
Step 9: Right-click on the primary site NAS volume and click “Restore Volume Config“ (on primary NAS cluster)

This will restore the shares, exports, snapshot schedules, and quota rules.



Step 10: Restore NAS Volume Configuration (on primary NAS cluster)

By default, all available configuration files are checked. Click “OK”


Restore Volume Config

NAS Volume Settings Backup Information

Cluster	RevDev8
NAS Volume	VOL0_REP
Backup Date	Mon Nov 03 15:49:06 CST 2014

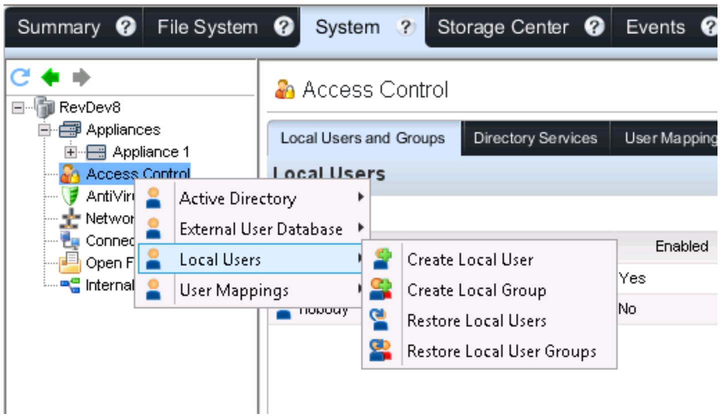
Restore Settings Configuration

SMB Shares	☑ Restore
NFS Exports	☑ Restore
Snapshot Scheduling	☑ Restore
Quota Rules	☑ Restore

Step 11: Restore other cluster-wide configuration items (on primary NAS cluster)

If best practices are being followed, the primary NAS cluster will already be configured with all environmental services such as DNS, NTP, Active Directory, LDAP/NIS, etc… But in the case that it is not, these must be configured.

Additionally, if local users/groups, or manual user mappings are in place, those must be manually restored. Local users and groups can be restored through Enterprise Manager:



Any manual user mapping rules that are defined must be restored using the FluidFS Command Line Interface using the command "access-control mapping manual restore".

Any networking configuration changes will have to be made manually. FluidFS does not have any automated method to restore network settings (such as IP's or static routes).

NOTE: Dell recommends as a best practice to document all network settings (IP addresses, default gateway, static routes) so that if needed they can be restored onto the secondary site.

Step 12: Repoint clients to the primary NAS cluster/NAS volume

At this point, the storage admin has 2 options of how to point clients to the primary NAS cluster:

1. Change the DNS entry that originally pointed to the Virtual IPs of the secondary site to point to the primary NAS cluster's Virtual IPs. Ensure that the DNS server(s) that the primary NAS cluster is using is the same DNS server(s) (or in the same DNS farm) as the DNS server(s) the secondary NAS cluster is using.Exising client connections will break and need to be re-established. All NFS exports must be unmounted and mounted back on every client system.
2. Change the IP's on the primary NAS cluster to be identical to the secondary NAS clusters IPs. Choosing this option allows for NFS hosts to not have to unmount/remount all of the NFS exports.

Now, users and applications are working off the primary NAS cluster.

Step 13: Re-establish replication from the primary NAS cluster to the secondary NAS cluster for all NAS volumes

6 Planned Failover Procedure

Often times, IT organizations wish to test their DR plan by conducting a planned failover of the entire NAS cluster, or maybe just a few NAS volumes (for a select number of applications). Additionally, planned maintenance may result in power outages, network outages, etc… and require a planned failover from the primary site to the secondary (DR) site. This section covers how to conduct a planned failover and failback of a full NAS cluster, as well as a planned failover and failback of individual NAS volumes.

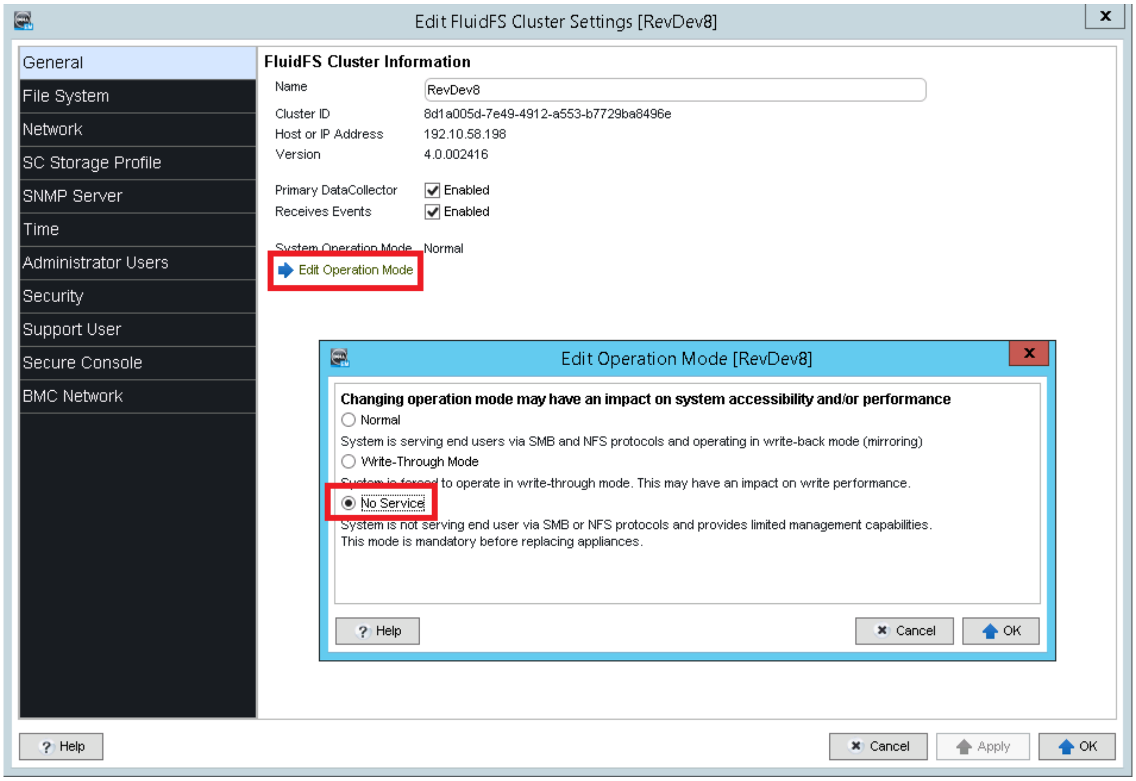
It is important to note however, that the best way to test out data on a secondary/DR volume is to use volume clones, which is detailed in Section 7 – Testing Disaster Recovery using Volume Clones.

6.1 Full Cluster Planned Failover and Failback

The full cluster planned failover and failback procedure is nearly identical to the steps outlined in Section 5 – Disaster Recovery Procedure. The main difference is that the primary NAS cluster must be "forced" to take all shares offline. This is accomplished by putting the primary NAS cluster into "NoService" mode. This can be done in the FluidFS CLI using the command:

```
system internal file-system service-mode set NoService
```

Alternatively, this can be done in Enterprise Manager by right clicking on the NAS cluster, and then clicking "Edit Settings":



FluidFS Cluster Information

Name: RevDev8
Cluster ID: 8d1a005d-7e49-4912-a553-b7729ba8496e
Host or IP Address: 192.10.58.198
Version: 4.0.002416

Primary DataCollector: ☑ Enabled
Receives Events: ☑ Enabled

System Operation Mode: Normal

Changing operation mode may have an impact on system accessibility and/or performance

○ Normal
System is serving end users via SMB and NFS protocols and operating in write-back mode (mirroring)
○ Write-Through Mode
System is forced to operate in write-through mode. This may have an impact on write performance.
● No Service
System is not serving end user via SMB or NFS protocols and provides limited management capabilities.
This mode is mandatory before replacing appliances.

The full NAS cluster **planned failover** procedure follows these high level steps:

1. (On Primary NAS Cluster) Promote the secondary NAS volume
2. (On Secondary NAS Cluster) Right-click on the remote site NAS volume and click "Restore Volume Config".
3. (On Secondary NAS Cluster) Restore NAS Volume Configuration
4. (On Secondary NAS Cluster) Restore other cluster-wide configuration items
5. (On Primary NAS Cluster) Change to "NoService" mode to bring down all shares and exports
6. (Using DNS or change IPs on secondary NAS cluster) Repoint clients to the secondary NAS cluster/NAS volume
7. (On Primary NAS Cluster) Delete the Replication which is using the primary as the source and the secondary as the destination

The full NAS cluster **planned failback** procedure follows these high level steps:

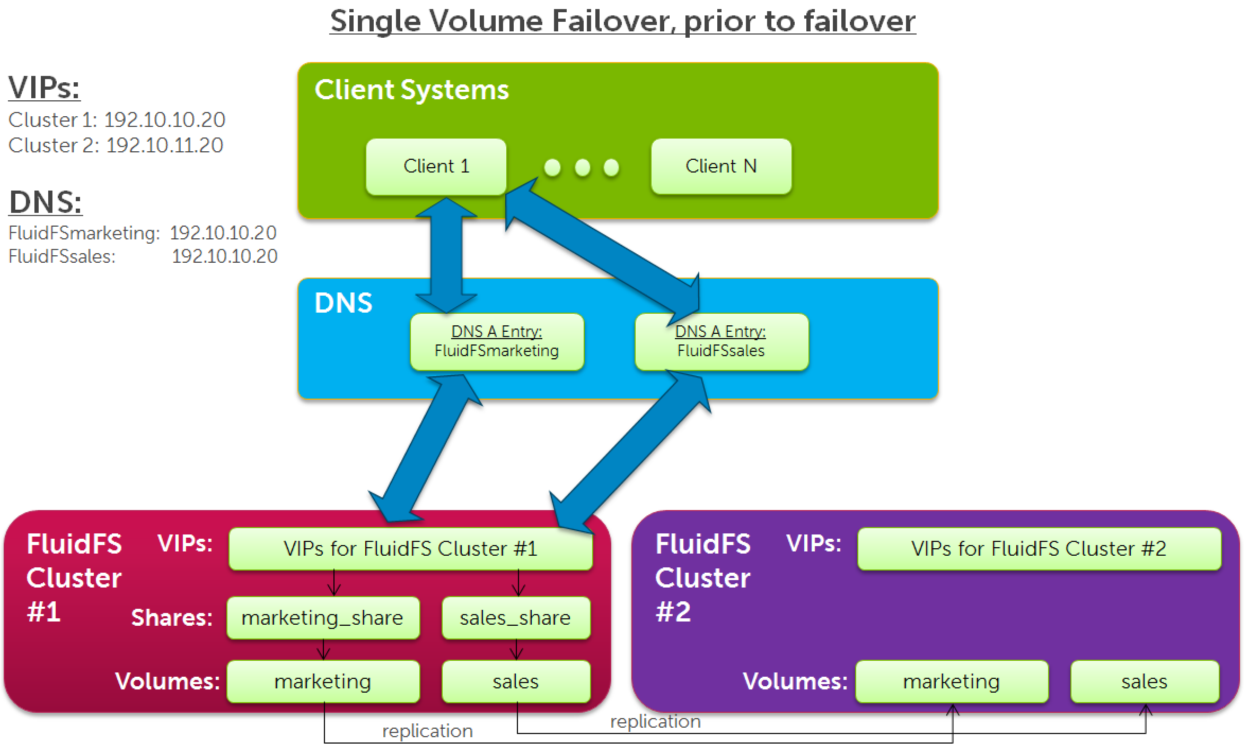
1. (On Secondary NAS Cluster) Create Replication using the secondary NAS volume as the source and the primary NAS volume as the destination
2. (On Secondary NAS Cluster) Choose Remote Cluster and Snapshot Retention Policy
3. (On Secondary NAS Cluster) Select Volume on primary NAS cluster to replicate to
4. (On Secondary NAS Cluster) Create a schedule for the replication
5. (On Secondary NAS Cluster) Create the replication schedule and click OK.
6. (On Secondary NAS Cluster) Wait for the replication to finish
7. (On Secondary NAS Cluster) Promote the primary NAS volume
8. (On Primary NAS Cluster) Right-click on the primary site NAS volume and click "Restore Volume Config".
9. (On Primary NAS Cluster) Restore NAS Volume Configuration
10. (On Primary NAS Cluster) Restore other cluster-wide configuration items
11. (On Secondary NAS Cluster) Change to "NoService" mode to bring down all shares and exports
12. (Using DNS or change IPs on primary NAS cluster) Repoint clients to the primary NAS cluster/NAS volume
13. (On Secondary NAS Cluster) Delete the Replication which is using the
14. (On Primary NAS Cluster) Recreate replication using the primary NAS volume as the source and the secondary NAS volume as the destination

7 Single NAS Volume Planned Failover and Failback

For single NAS volume failover, it is important that the environment is set up to properly migrate clients of the NAS volumes you are failing over, without disrupting the clients of other NAS volumes you are not failing over.

When a NAS volume is failed over from one NAS cluster to another, the IP addresses that are used to access it change from NAS Cluster A's IP addresses to NAS Cluster B's IP addresses. Dell recommends facilitating this change using DNS. It is recommended to set up a DNS entry to correlate to each NAS volume, and change the DNS entry for single NAS volumes when they are failed over.

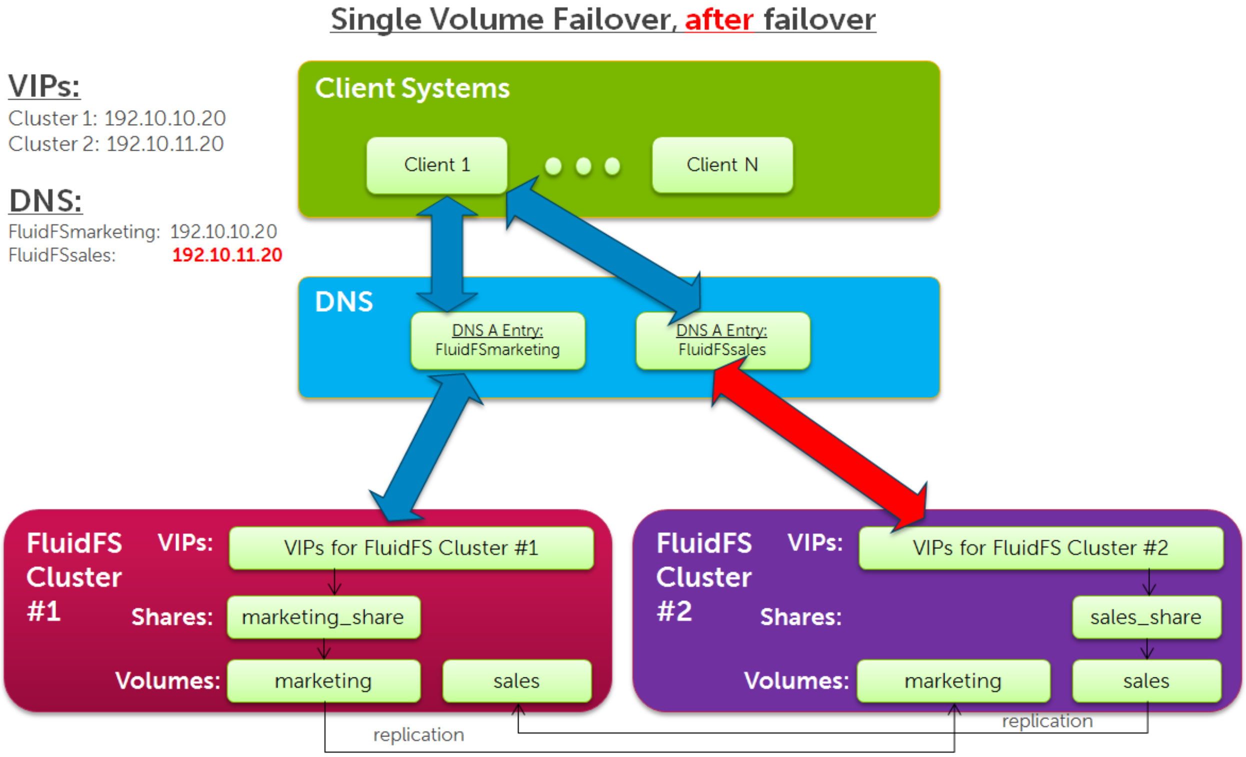
For example, suppose Marketing and Sales have their own NAS volumes, with an SMB share on the NAS volume named **marketing_share** and **sales_share** respectively. A DNS A (host) entry named **FluidFSmarketing**, is created for Marketing and another DNS A (host) entry for Sales named **FluidFSsales** is created. Both NAS volumes point to the same set of VIPs on source NAS Cluster A. Marketing can access the Marketing SMB share using **\\FluidFSmarketing\marketing**, and Sales can access the Sales SMB share using **\\FluidFSsales\sales**. The topology looks like this diagram below. For simplicity, only one VIP was used, but in reality there are usually more than one.



Initially, both DNS entries **FluidFSmarketing** and **FluidFSsales** point to the same set of VIPs. At this point, both the **marketing** and **sales** SMB shares can be accessed from either one of the DNS names, since FluidFS serves all shares through all VIP's by default. When you want to fail over a single NAS volume (for example **sales**) change the DNS entries for **FluidFSsales** to resolve to the VIPs on Cluster B. Client systems may need to refresh their DNS cache. DNS entries can also be created with lower TTLs to make failing back quicker.

To force SMB and NFS clients to NAS Cluster B, you must delete the SMB shares and NFS exports on NAS Cluster A. This forces the SMB and NFS clients to reconnect, at such time they are connected to NAS Cluster B.

After restoring the source volume's configuration on NAS Cluster B, all of the SMB shares and NFS exports will be present on the target NAS volume (on NAS Cluster B), so no SMB share/NFS export configuration information is lost. The failed over NAS volume can now be accessed using the exact same DNS name and SMB share/NFS export name as it was when hosted on NAS Cluster A, except now it is hosted on NAS Cluster B. After the single NAS volume failover for the **sales** volume, the topology will look like this:



To fail back from NAS Cluster B to NAS Cluster A, the same procedure is run in reverse.

Dell recommends that you maintain accurate and detailed records/tables to track which DNS entries are used to access each NAS volume. This helps when performing failover and setting up group policies.

8 Testing Disaster Recovery using Volume Clones

FluidFS has the ability to create Thin Volume Clones of a NAS volume utilizing its snapshots. This feature is a powerful tool for testing disaster recovery, without affecting production volumes. Often times, administrators will have application servers that reside in a test environment, and wish to test out their DR plan, but they don’t want to affect production. FluidFS Thin Volume Clones are a perfect tool for facilitating this.

As has been discussed earlier in this document, when a NAS volume is a replication secondary NAS volume, it is in a read-only state. In order to change the secondary NAS volume from read-only to read/write, it must be promoted. However, that will result in replication pausing for the time period that the secondary NAS volume is in the “promoted” state. Using Thin Volume Clones allows the replication to continue on its normal schedule, while testing DR.

The procedure to test failover and failback using Thin Volume Clones is below:

1. Select the secondary NAS cluster in Enterprise Manager
2. On the secondary NAS cluster, select the secondary NAS volume in Enterprise Manager
3. Navigate to the “Snapshots” tab
4. To create a thin volume clone from the last replication snapshot, right click on the snapshot that starts with “rep_”. This snapshot represents the state of the source NAS volume at the time point of the last scheduled replication. Any other snapshot can be chosen as well.
5. Select “Create NAS Volume Clone” and create a NAS Volume Clone
6. Create an SMB share or NFS export on that NAS volume clone
7. Repoint the application to the DNS name of the secondary NAS cluster, and point it to the NFS export or SMB share on the clone of the secondary NAS volume.
8. If the administrator wishes to test failback, the clone of the secondary NAS volume can be replicated to a new, empty NAS volume on the primary NAS cluster. This assumes that adequate space is available on the primary NAS cluster to duplicate data. Keep in mind that if there is a large amount of data, replicating the entire dataset can take some time.
9. After the clone of the secondary NAS volume finishes replicating to the new NAS volume on the primary NAS cluster, an SMB share or NFS export will need to be created on the new NAS volume on the primary NAS cluster.
10. Repoint the application to the DNS name of the primary NAS cluster, and point it to the NFS export or SMB share on the new NAS volume on the primary NAS cluster.
11. After testing is completed, the application can be pointed back to the main/production NFS export or SMB share, if the administrator so chooses.

A Appendix A: Scripting failover/failback using the FluidFS CLI

One way to optimize the amount of time taken to perform the failover/failback procedure is to script against the FluidFS Command Line Interface. The FS8600 Administrators Guide details how to set up passwordless SSH from a Linux host. For the examples given in this Appendix, we will assume that passwordless SSH was set up using the "Administrator" user. The commands given here will follow the failover and failback procedures detailed in this document.

However, as a reminder, the key parameters are in the table below:

Parameter	Value
Source NAS Cluster	RevDev8
Source NAS Volume	VOL0
Source NAS Cluster Virtual IP	192.168.10.50/24 – DNS Name PrimaryVIP
Destination NAS Cluster	RevDev9
Destination NAS Volume	VOL0_REP
Destination NAS Cluster Virtual IP	192.168.11.50/24 – DNS Name SecondaryVIP

A.1 Replication Setup

The script to perform the failover procedure follows these high level steps:

- Create replication partnership
- Create secondary NAS volume
- Connect source and destination NAS volume
- Create a schedule to replicate once per hour

For example, the script from the Linux client would look like this:

```
ssh Administrator@PrimaryVIP system data-protection cluster-partnerships add SecondaryVIP
ssh Administrator@SecondaryVIP NAS-volumes add VOL0_REP
ssh Administrator@PrimaryVIP NAS-volumes replication connect VOL0 RevDev9 VOL0_REP
ssh Administrator@PrimaryVIP NAS-volumes replication schedules add VOL0 RevDev9 VOL0_REP Rep0
Periodic –Period 60
```

A.2 Failover CLI Scripting

The high level steps, along with the CLI commands to perform these steps, **run against the secondary cluster**, within the context of this example, are as follows:

- promote secondary NAS volume
  - NAS-volumes replication promote VOL0_REP RevDev8 VOL0
- restore NAS volume config

    - NAS-volumes configuration-backups restore-configuration VOL0_REP CifsShare,NfsExport,QuotaRule,SnapshotSchedule
- (optional)
    - restore local users
        - access-control local-users restore RevDev8
        - access-control local-groups restore RevDev8
    - restore manual user mapping
        - access-control mapping manual restore RevDev8
- Update DNS or change IPs
- delete replication
    - NAS-volumes replication disconnect VOL0_REP RevDev8 VOL0

For example, the script from the Linux client would look like this:

```
ssh Administrator@SecondaryVIP NAS-volumes replication promote VOL0_REP RevDev8 VOL0

ssh Administrator@SecondaryVIP NAS-volumes configuration-backups restore-configuration
VOL0_REP CifsShare,NfsExport,QuotaRule,SnapshotSchedule

ssh Administrator@SecondaryVIP access-control local-users restore RevDev8

ssh Administrator@SecondaryVIP access-control local-groups restore RevDev8

ssh Administrator@SecondaryVIP access-control mapping manual restore RevDev8

ssh Administrator@SecondaryVIP NAS-volumes replication disconnect VOL0_REP RevDev8 VOL0
```

After the script is run, either DNS should be updated, or the IP's on the secondary NAS cluster changed to match those of the primary site.

A.3 Failback CLI Scripting

The high level steps, along with the CLI commands to perform these steps, **run against the primary and secondary NAS cluster**, within the context of this example, are as follows:

- Create replication from secondary site to primary site (Run from Secondary NAS Cluster)
    - NAS-volumes replication connect VOL0_REP RevDev8 VOL0
- Trigger manual replication (Run from Secondary NAS Cluster)
    - NAS-volumes replication start VOL0_REP RevDev8 VOL0
- promote primary volume (Run from Primary NAS Cluster)
    - NAS-volumes replication promote VOL0 RevDev9 VOL0_REPREP
- restore volume config (Run from Primary NAS Cluster)
    - NAS-volumes configuration-backups restore-configuration VOL0 CifsShare,NfsExport,QuotaRule,SnapshotSchedule
- (optional)
    - restore local users (Run from Primary NAS Cluster)

        - access-control local-users restore RevDev9
        - access-control local-groups restore RevDev9
    - restore manual user mapping (Run from Primary NAS Cluster)
        - access-control mapping manual restore RevDev9
- Update DNS or change IPs
- delete replication (Run from Primary NAS Cluster)
    - NAS-volumes replication disconnect VOL0 RevDev9 VOL0_REP

For example, the script from the Linux client would look like this:

```
ssh Administrator@SecondaryVIP NAS-volumes replication connect VOL0_REP RevDev8 VOL0

ssh Administrator@SecondaryVIP NAS-volumes replication start VOL0_REP RevDev8 VOL0
```

Wait for the replication to finish…

```
ssh Administrator@PrimaryVIP NAS-volumes replication promote VOL0 RevDev9 VOL0_REP

ssh Administrator@PrimaryVIP NAS-volumes configuration-backups restore-configuration VOL0
CifsShare,NfsExport,QuotaRule,SnapshotSchedule

ssh Administrator@PrimaryVIP access-control local-users restore RevDev9

ssh Administrator@PrimaryVIP access-control local-groups restore RevDev9

ssh Administrator@PrimaryVIP access-control mapping manual restore RevDev9

ssh Administrator@PrimaryVIP NAS-volumes replication disconnect VOL0 RevDev9 VOL0_REP
```

After the script is run, either DNS should be updated, or the IP's on the primary NAS cluster changed to match those of the secondary site.

B Appendix B: Additional resources

FluidFS on Dell TechCenter

Below are some links to additional resources located at http://kc.compellent.com

- Dell Compellent Documentation
- Dell Compellent Replay Manager 7 Administrator's Guide
- Dell Compellent Enterprise Manager 6.5 Users Guide
- Dell Compellent Storage Center 6.5 Users Guide