Technical Report

NetApp SnapProtect Management Software Overview and Design Considerations

Chris Blackwood, Larry Touchette, NetApp
May 2011 | TR-3920 Version 1.2

ABSTRACT

NetApp® SnapProtect™ management software is changing today's backup and recovery landscape. SnapProtect software combines simplified manageability, power, and flexibility for virtual environments with full support for enterprise database applications while providing virtually seamless integration with NetApp Snapshot® technology for fast and efficient backup operations. In addition, SnapProtect software integrates with NetApp SnapVault® and SnapMirror® software with support for content-based cataloging and movement to tape-based media. This document is an introduction to the SnapProtect solution. It provides an overview of the technology and describes some of the basic configuration steps required to get started.

TABLE OF CONTENTS

LIST OF TABLES

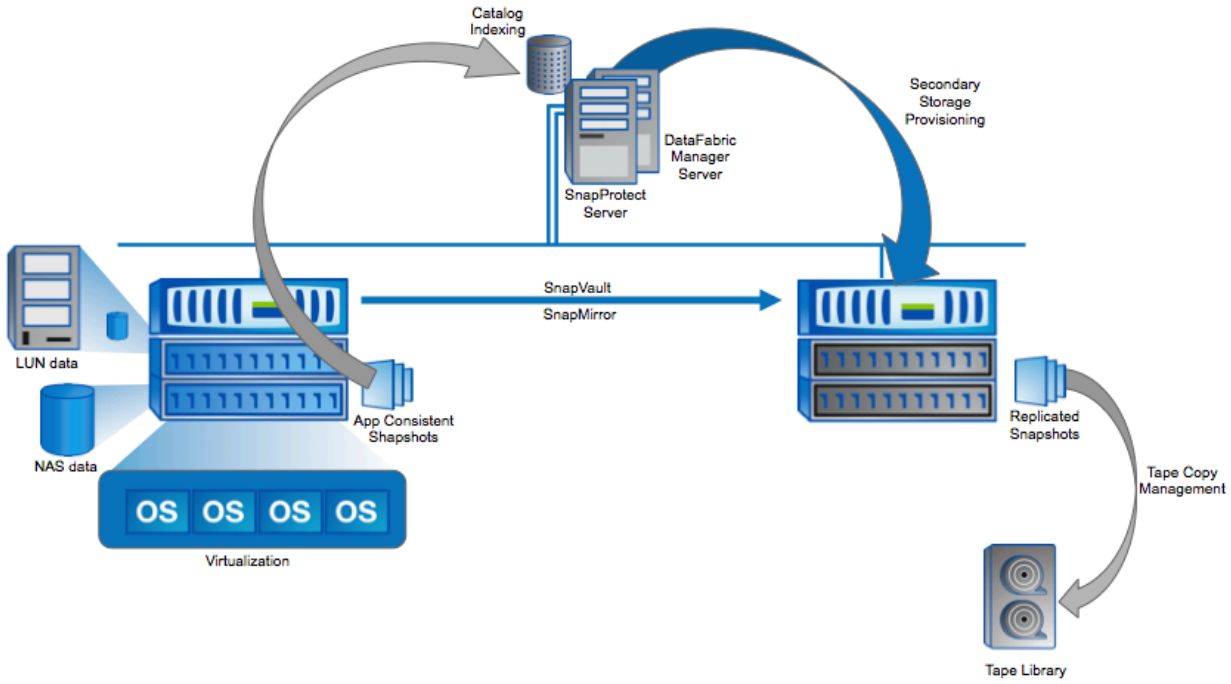
LIST OF FIGURES

1 INTRODUCTION

NetApp is an industry leader in array-based data protection. The efficiencies of Snapshot copy technology and data replication have changed the way we look at backup and recovery and disaster recovery strategies. The need to achieve higher SLAs and to meet backup windows is a constant challenge, given the explosion of data that enterprises are dealing with today. Data center consolidation through virtualization has also created challenges around data protection. Disk-to-disk data protection solutions are becoming more widely accepted for both backup and recovery and disaster recovery strategies. NetApp data protection solutions offer speed and flexibility while reducing storage capacity requirements through the use of efficient array-based technologies. The result is a simplified approach that reduces costs and administrative effort.

NetApp SnapProtect management software offers enterprise-class management for backup and recovery in the data center. The SnapProtect software manages Snapshot copies on NetApp primary storage and replication to secondary and tertiary storage, as well as tape creation. Whether you are protecting NetApp application data, file data for NAS, file data in LUNs, or data in virtualized environments, the SnapProtect solution provides the management, the storage provisioning, the cataloging, and the granular recoverability required for seamless operation. Figure 1 shows this basic flow.

Figure 1) SnapProtect software overview.



SnapProtect software can be used to protect the following applications hosted on NetApp primary storage:

- Microsoft® Active Directory® (Windows®)
- Microsoft Exchange (Windows)
- Microsoft SQL Server® (Windows)
- Microsoft Office SharePoint® Server (Windows)
- Oracle® (UNIX® and Linux®)
- DB2 (UNIX and Linux)

- SAP® for Oracle (UNIX and Linux)

In addition, the SnapProtect solution supports the following virtualization product:

- VMware® vSphere™

For a complete list of supported platforms and product versions, refer to the NetApp SnapProtect administration information.

Note: The SnapProtect solution requires NetApp Data ONTAP® 7.3.5, as well as DataFabric® Manager 4.0.2.

1.1 TERMINOLOGY

The following SnapProtect components work together to create a full solution.

- **CommCell.** A single instance of a SnapProtect environment.
- **CommServe.** The master server in a SnapProtect environment. This server uses a Microsoft SQL Server database and therefore must be a Microsoft Windows system (Windows Server® 2003 or 2008).
- **Media agent.** A media server in a SnapProtect environment. Media agents have broad operating system support, including Windows, Linux, and UNIX options.
- **CommCell Console.** The SnapProtect management interface.
- **iDataAgent (iDA).** Agents that control data consistency during backup operations.
- **Clients.** Hosts running iDataAgents for which data is protected.
- **Backup Set.** A layer of management within iDataAgents for grouping subclients.
- **Subclient.** A layer of management within a Backup Set. A client can have multiple subclients, each of which can be associated with different source data.
- **Disk Library.** A storage resource with an associated mount path that is used in the SnapProtect solution to store index information backups.
- **Storage Policy.** A logical object through which a subclient is protected. The storage policy defines how data is backed up and replicated as well as retention requirements.
- **DataFabric Manager server.** A server running NetApp DataFabric Manager server software. The DataFabric Manager server and the CommServe server should typically be separate systems. DataFabric Manager 4.0.2 or later is required.
- **NetApp Management Console (NMC).** The NetApp Management Console is an interface used for creating resource pools and provisioning policies within the DataFabric Manager framework. The NMC should be installed on a separate system from the DataFabric Manager server.
- **NetApp primary.** The production NetApp storage array.
- **NetApp secondary.** The secondary NetApp storage array used as a destination for replication.
- **NetApp tertiary.** A third NetApp storage array used for replicating previously replicated data.
- **Snapshot copy.** A NetApp array-based point-in-time copy used for recovering data.
- **SnapVault.** A NetApp replication technology used for backup and recovery. In the SnapProtect solution, a “vault” copy uses SnapVault.

- **SnapMirror.** A NetApp replication technology used for disaster recovery. In the SnapProtect solution, a “mirror” copy uses SnapMirror.

2 TECHNICAL OVERVIEW

This section covers the technical details of the SnapProtect software and how the components work together.

2.1 BASIC FUNCTIONS

The SnapProtect solution delivers several basic functionalities to create a simplified user experience, including the following:

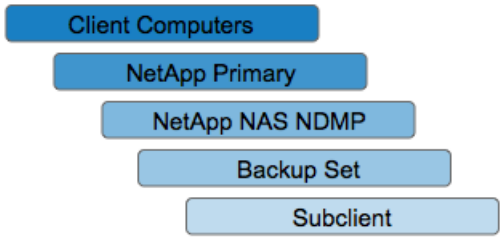
- Snapshot copy creation
- Cataloging and indexing
- Storage provisioning
- Data replication using SnapVault and/or SnapMirror
- Data movement to tape

SNAPSHOT COPY CREATION

The SnapProtect software creates Snapshot copies on the NetApp primary storage as its first backup copy. This is important because Snapshot technology allows backups to complete very quickly. Primary Snapshot copy creation is handled differently for different types of data. For NAS data, the NetApp primary system is treated as a client with an associated iDataAgent (iDA) called “NetApp NAS NDMP.” Subclients within the iDA are configured and associated with the NetApp data that requires protection. When a backup for the subclient runs, NetApp Snapshot copies are created for the volumes in that subclient.

Figure 2 shows how this structure looks for a NetApp primary and its iDA.

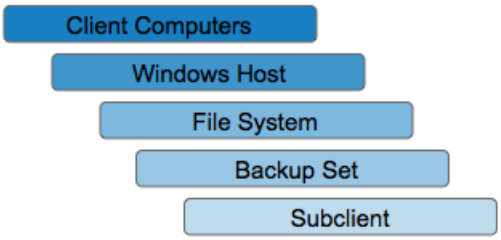
Figure 2) NetApp NAS NDMP iDA.



For LUN data hosted on NetApp primary storage, the host accessing the data is treated as the client. The attached drive on the client is associated with a subclient within the File System iDA or the associated application iDA. On Windows clients, the iDA calls Microsoft Volume Shadow Copy Service (VSS) to make sure that the data within the file system is consistent. With application-integrated Snapshot copies, the application agent calls VSS (Windows) or places the database in hot-backup mode (UNIX or Linux) for backup consistency. Then the Snapshot copy is created on the NetApp primary system for the volume containing the LUN.

Figure 3 shows how this structure looks for a Windows client and its File System iDA.

Figure 3) Windows File System iDA.



CATALOGING AND INDEXING

The ability to index the contents of a backup is a core value of the SnapProtect solution. For basic NAS data, the contents of the Snapshot copies that are created by the SnapProtect software are indexed directly. For LUN data, LUN clones are created and used for indexing. For LUN data, the contents inside the LUN are indexed. Specific proxy servers can also be assigned to handle indexing.

The indexes are stored in disk libraries. NetApp recommends creating disk libraries with paths that point to NetApp primary storage.

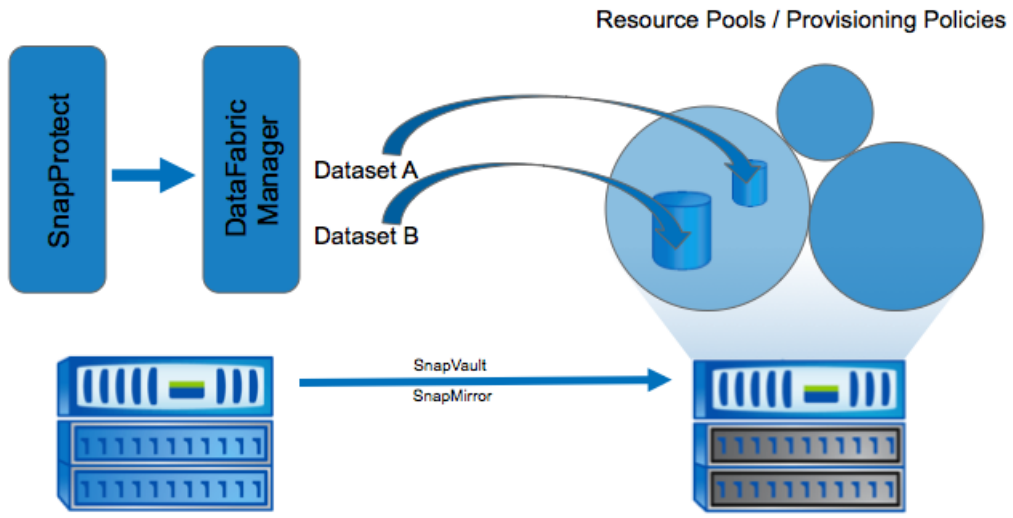
STORAGE PROVISIONING

Storage provisioning is required whenever Snapshot copies on NetApp primary storage need to be replicated to NetApp secondary and tertiary storage. Before replication can be established, the secondary system must have the appropriate volumes in place, along with the correct volume settings. SnapProtect software takes care of this by using the provisioning services of NetApp DataFabric Manager.

DataFabric Manager uses policy-based rules that define how storage should be provisioned under different circumstances. Where the storage is provisioned is also flexible in that it uses pools of storage called *resource pools*—DataFabric Manager containers that point to one or more aggregates within a NetApp system. The storage administrator only needs to provide the backup administration resource pools to use. The backup administrator simply directs the SnapProtect software to use the resource pools and provisioning policies for replication purposes. The appropriate storage is provisioned automatically.

Figure 4 shows how provisioning works. When a new replication relationship is configured in SnapProtect management software, this information is passed to DataFabric Manager, which creates the datasets needed to manage the replication requirements. DataFabric Manager then provisions the necessary volumes, using the resource pools and provisioning policies that were assigned in the SnapProtect configuration.

Figure 4) Storage provisioning.



The resource pools must be created manually by using the NetApp Management Console (NMC). The SnapProtect software then discovers the resource pools and makes them available in the CommCell console for selection.

It is not necessary to create provisioning policies manually unless custom policies are needed; the SnapProtect software has a set of preconfigured provisioning policies (Table 1).

Table 1) Preconfigured provisioning policies.

Policy Name	Availability	Deduplication	Space Thresholds
SnapProtect_RAID-DP	RAID-DP	No	No
SnapProtect_Dedupe	RAID-DP	On-Demand	80%, 90%
SnapProtect_Mirror_Destination	RAID-DP	No	80%, 90%

It is considered a best practice to use NetApp RAID-DP® for resiliency. Therefore all of the preconfigured provisioning policies enable RAID-DP.

Deduplication can be enabled or disabled for vault copies independently of the deduplication setting on the primary data volume. Deduplication for vault copies uses the On-Demand setting, and deduplication runs automatically as vault copy jobs complete. To enable deduplication for vault secondary storage, the SnapProtect_Dedupe provisioning policy can be used. For vaulting that does not require deduplication on secondary storage, the SnapProtect_RAID-DP policy can be used.

Mirror copies inherit the deduplication settings of the primary data volume. Therefore deduplication is disabled in the SnapProtect_Mirror_Destination provisioning policy. If the primary volume has deduplication enabled, the mirror copy volume is also a deduplicated volume.

Space thresholds represent the “nearly full threshold” and “full threshold” properties used when provisioning storage.

Custom provisioning policies will also be available. If there is a need for settings apart from the preconfigured policies, new provisioning policies must be created by using the NetApp Management Console. The SnapProtect software will automatically discover the custom provisioning policies.

DATA REPLICATION USING SNAPVAULT AND/OR SNAPMIRROR

As mentioned previously, there are two types of replication operations that can be configured using the SnapProtect solution – vault copies and mirror copies. It is important to understand the difference between the two. A vault copy uses NetApp SnapVault, and a mirror copy uses NetApp asynchronous volume SnapMirror. SnapVault and SnapMirror replicate data in a similar way in that they both replicate only the blocks that have changed since the last replication operation. A full copy of the data is made only once. After the initial full copy, block-based incremental backups are made. This approach to disk-to-disk data protection offers speed, network efficiency, and storage capacity savings.

A vault copy has a certain independence from the primary data, allowing retention levels between the primary copy and the secondary copy to be different. For example, a vault copy allows longer-term retention than the primary copy.

A mirror copy depends more firmly on the primary data and is an exact mirror of the source volume and its Snapshot copies. A mirror copy cannot have an independent level of retention. Mirror copies are traditionally used in disaster recovery solutions because they allow failover to the secondary copy. Another use for mirror copies is to duplicate a vault copy to a remote location.

SnapProtect allows several combinations of vaulting and mirroring, satisfying many disk-to-disk-to-tape requirements.

Note: SnapVault and SnapMirror licenses must be enabled on the NetApp systems.

DATA MOVEMENT TO TAPE

The SnapProtect solution allows several ways for backup copies to be replicated to tape, including NDMP dump and tape streaming through media agents. Tape backups are enabled in the storage policy properties under the Snapshot tab; select Enable Backup Copy.

Tape backups use the Primary(Classic) copy in the storage policy. They can be configured so that any one of the replication destinations can be used as the source for this tape copy. Select the source for the backup copy in the properties for the storage policy; under the Snapshot tab, modify Source Snap Copy.

The NetApp NAS NDMP iDA is the only agent that allows NDMP dump. Tape copies from other iDAs stream through a media agent.

When using the NetApp NAS NDMP iDataAgent to create primary Snapshot copies that will be moved to tape by using NDMP dump, the indexing function for the SnapProtect job can be disabled. This allows the NDMP dump operation to run without delay. To skip indexing the primary Snapshot copy, make sure that the Skip Catalog Phase For Snap Backup checkbox is selected when running the SnapProtect backup.

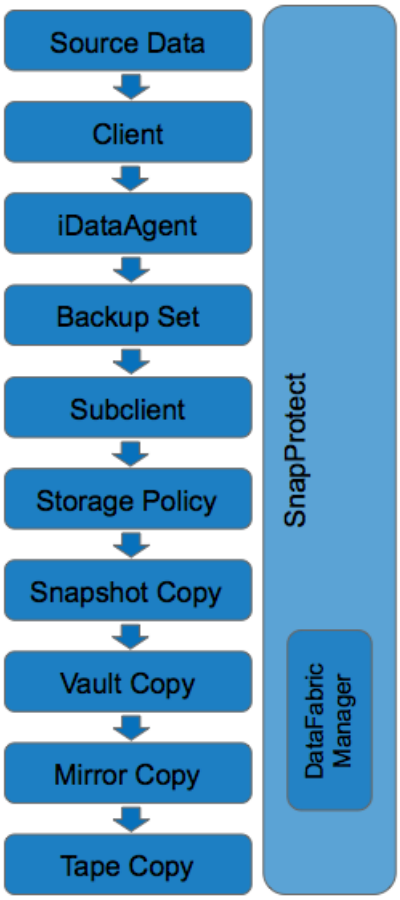
Creating full or incremental copies to tape depends on how the SnapProtect backup was initiated. If it is required to run incremental tape jobs, then incremental schedules must be created for the SnapProtect backup. To select the tape backup job, right-click the storage policy → Properties → Snapshot tab.

2.2 UNDERSTANDING THE BACKUP WORKFLOW

To understand how the SnapProtect software works, it is important to understand the workflow, starting from the source data and working outward.

Clients own source data. Clients have specific iDataAgents, depending on the type of client and the data being protected. Backup sets and subclients are configured within the iDA, and they group the source data to be protected. For example, if the F:\ drive is to be protected for an individual client, a subclient would contain an entry for the F:\ drive.

Figure 5) SnapProtect workflow: vault then mirror then tape.



The storage policy determines the behavior of the data protection operations as well as the retention properties. Each subclient is associated with a storage policy, which contains entries for the various copies in the data protection layout. In the example in Figure 5, client data is protected by NetApp Snapshot copies. Vaulting is performed for longer-term retention. The vaulted data is then mirrored for redundancy. Tape copies are then created from the mirror copy. The SnapProtect software orchestrates the operations, passing the vaulting and mirroring job control to the DataFabric Manager server.

Figure 5 shows one example. Section 2.4 discusses other layouts for replicating data.

2.3 HANDOFF TO DATAFABRIC MANAGER

When a new replication copy is created in SnapProtect, the task is given to DataFabric Manager to carry out. DataFabric Manager creates the required datasets, provisions the required volumes, and initiates the baseline data transfers. It is important to know how source data gets from a subclient to a dataset to a destination volume during replication. This may differ, depending on the replication type and the data type.

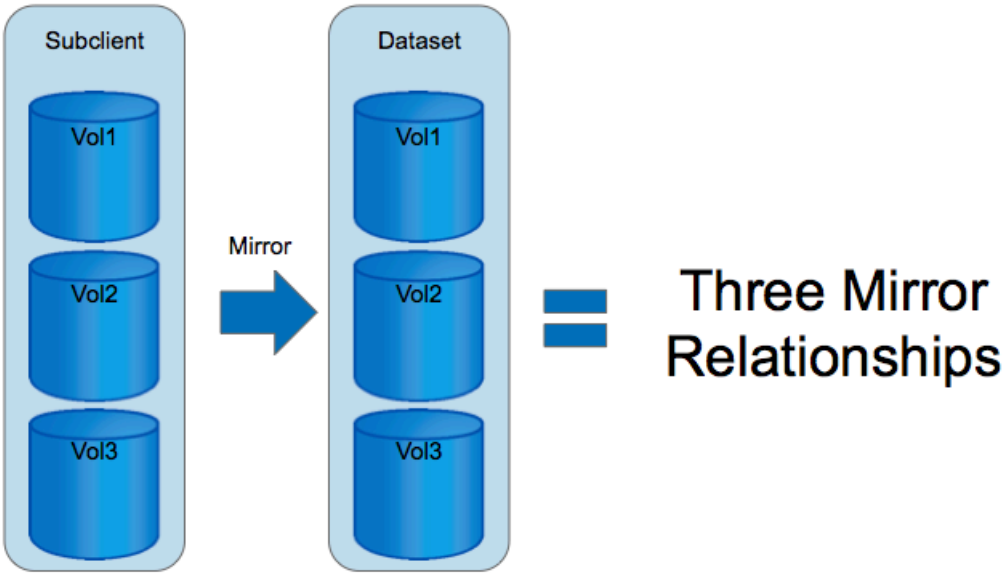
In all cases, a SnapProtect subclient has a 1-to-1 mapping to a DataFabric Manager dataset. In addition, a SnapProtect subclient cannot span clients.

NAS DATA

For example, if a single NetApp primary system is configured as a NAS client, all of the NAS volumes on that primary system could be grouped together by a single subclient. The result would be a single dataset in DataFabric Manager. If the storage policy in the SnapProtect software calls for mirroring this subclient, then the dataset would create mirror relationships for each of the volumes in the subclient.

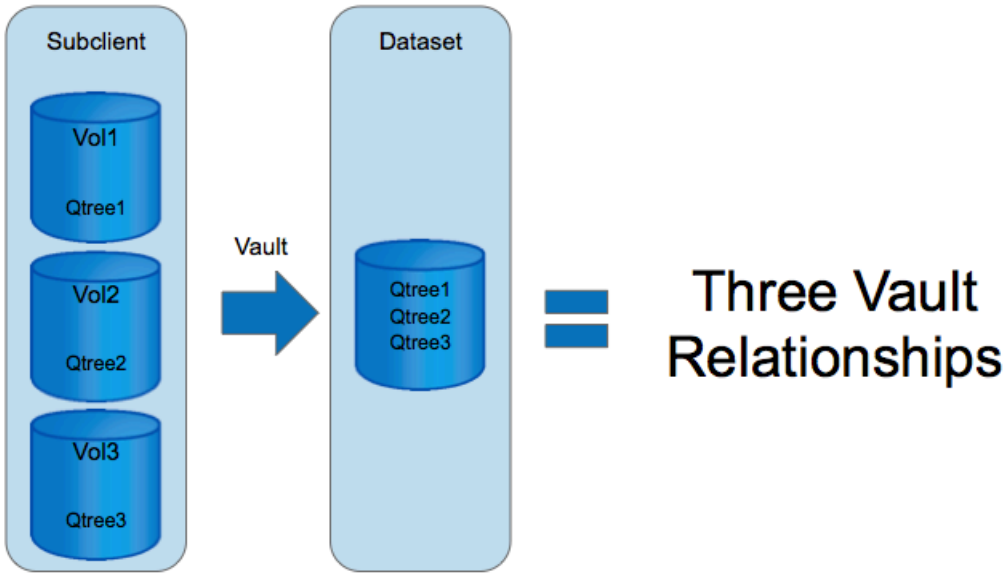
In Figure 6, three NAS volumes are grouped into a single subclient. Creating a mirror copy results in a single dataset in the DataFabric Manager server and three mirror relationships are established.

Figure 6) Mirroring NAS volumes.



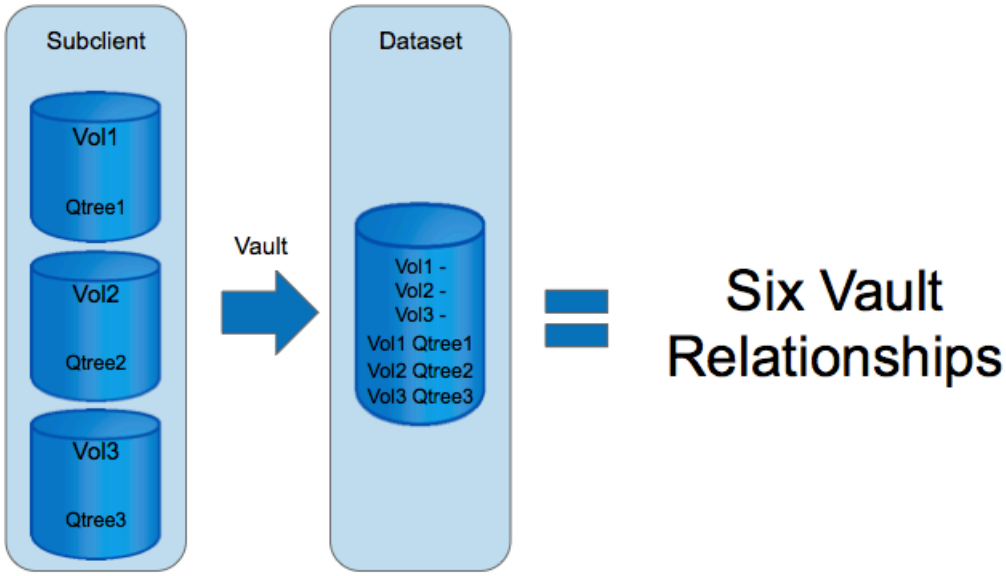
Vaulting NAS data is slightly different because a vault can be more granular in scope. In addition to the entire volume, individual qtrees in a primary volume can be selected for vaulting purposes. In Figure 7, single qtrees from three volumes are grouped into a single subclient. Creating a vault copy results in a single dataset in the DataFabric Manager server and three vault relationships are established, one for each of the qtrees. In this example, DataFabric Manager is configured to allow a fan-in of the vaulted relationships. Enabling fan-in on the DataFabric Manager server requires setting the `dpMaxFanInRatio` parameter on the server. For example, to set the fan-in ratio to 10, run **`dfm options set dpMaxFanInRatio=10`** on the DataFabric Manager server.

Figure 7) Vaulting NAS qtrees.



On the other hand, if the entire volumes were vaulted, six vault relationships would be established. This is because a relationship for each volume's non-qtree data would be created as well as a relationship for each of the volume's qtrees. In Figure 8, the three volumes are grouped into a single subclient and vaulted.

Figure 8) Vaulting NAS volume and qtrees.



LUN DATA

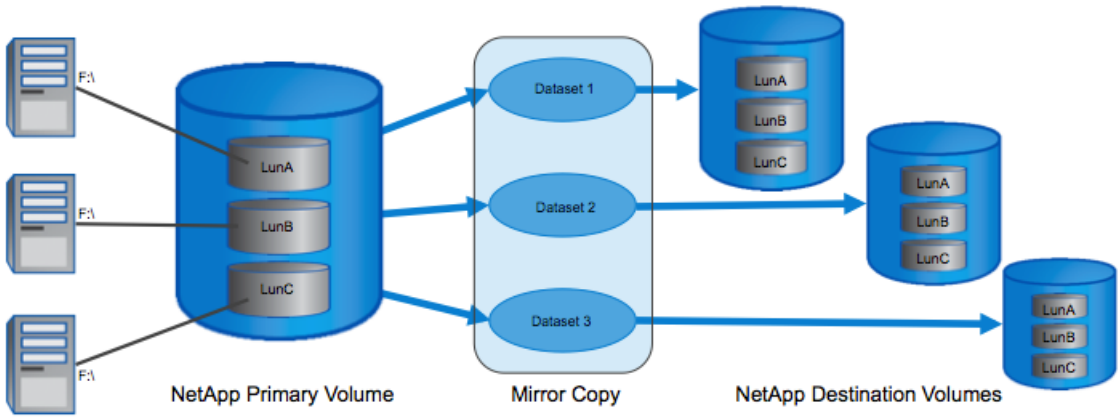
When working with LUN data, whether using an application iDA or the file system iDA (see Figure 3), there are specific guidelines to follow. Because subclients do not span clients, primary data must be laid out with the subclient in mind.

It is best if all LUNs in the primary volume are protected by a single subclient. For this to work, the following must be true:

- All LUN data on the volume must belong to the same client.
- All LUN data on the volume can be protected by the same iDataAgent.

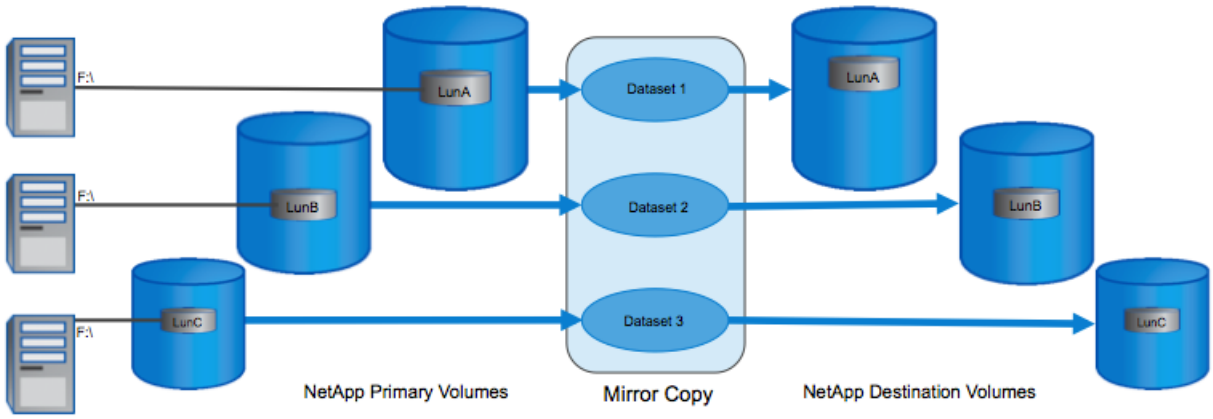
Volumes with LUN data split across multiple subclients can result in increased capacity requirements for replication operations. Consider an example where three clients each map to LUNs in a common volume. If these subclients were mirrored, it would result in three datasets and three baseline copies for the common volume, as shown in Figure 9. Vaulting would result in the same behavior unless each LUN were in its own qtree.

Figure 9) Multiple clients with LUN data on a common volume.



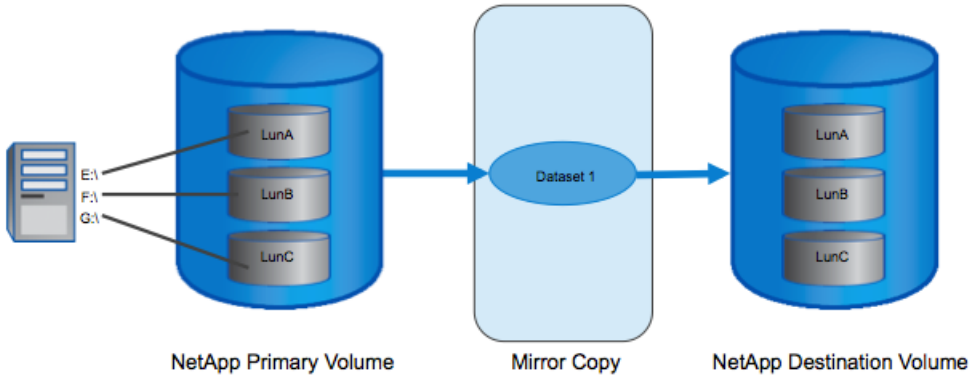
In Figure 10, the LUNs are on separate volumes.

Figure 10) Multiple clients with LUN data on separate volumes.



In Figure 11, the LUNs on the common volume are mapped by a single client and grouped by a single subclient. The result of mirroring this subclient would be a single dataset and a single baseline copy for the common volume.

Figure 11) Single client with LUN data on a common volume.



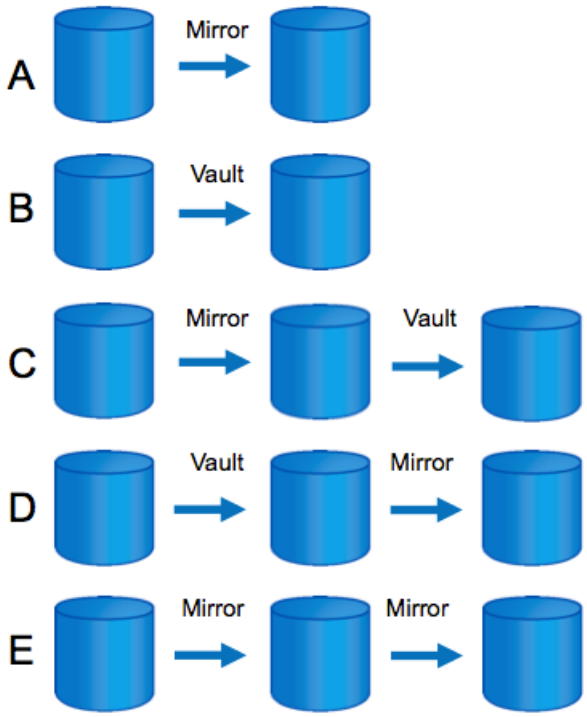
2.4 REPLICATION OPTIONS

There are various ways to architect the mirroring and vaulting strategies for replicating data. Mirroring and vaulting can be used separately or they can be used together and are configured by using storage policies. A storage policy has two copies by default, a Primary(Classic) copy and a Primary(Snap) copy. The Primary(Classic) copy is used for tape copies. The Primary(Snap) copy relates to the NetApp Snapshot copy on the primary system. To vault or mirror that primary data, additional copies must be created in the storage policy.

To create a mirror of the primary data, a mirror copy is created that points to Primary(Snap) as its source. Creating a vault of the primary data is similar, also pointing to Primary(Snap) as its source.

To vault the mirror copy, the source for the vault is set to the mirror copy and so forth. Figure 12 shows some of the replication combinations that can be configured.

Figure 12) Replication combinations.



Note: When vaulting from the mirror copy (example C in Figure 12), the `snapvault.snapshot_for_dr_backup` option must be set to `named_snapshot_only` on the mirror destination system. For more information refer to the SnapVault Best Practices Guide (http://www.netapp.com/us/library/technical-reports/tr-3487.html).

The policy-based replication combinations shown in Figure 12 translate to the following protection policies in DataFabric Manager.

A - SnapProtect Mirror

B - SnapProtect Back up

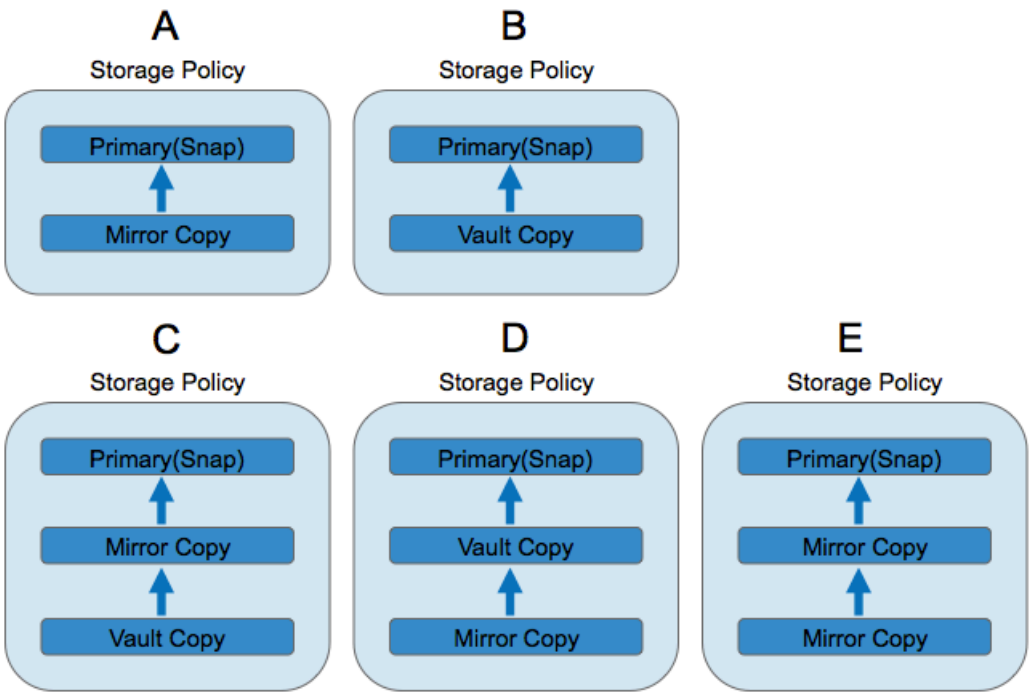
C - SnapProtect Mirror, then back up

D - SnapProtect Back up, then mirror

E - SnapProtect Chain of two mirrors

To create these scenarios, copies within the storage policies should be configured so that the appropriate source dependencies are established. Figure 13 illustrates these dependencies.

Figure 13) Source dependencies within storage policies.



As shown in Figure 14, additional combinations can be added as building blocks to the options listed in Figure 12 to establish fan-out scenarios.

Figure 14) Example fan-out scenarios.

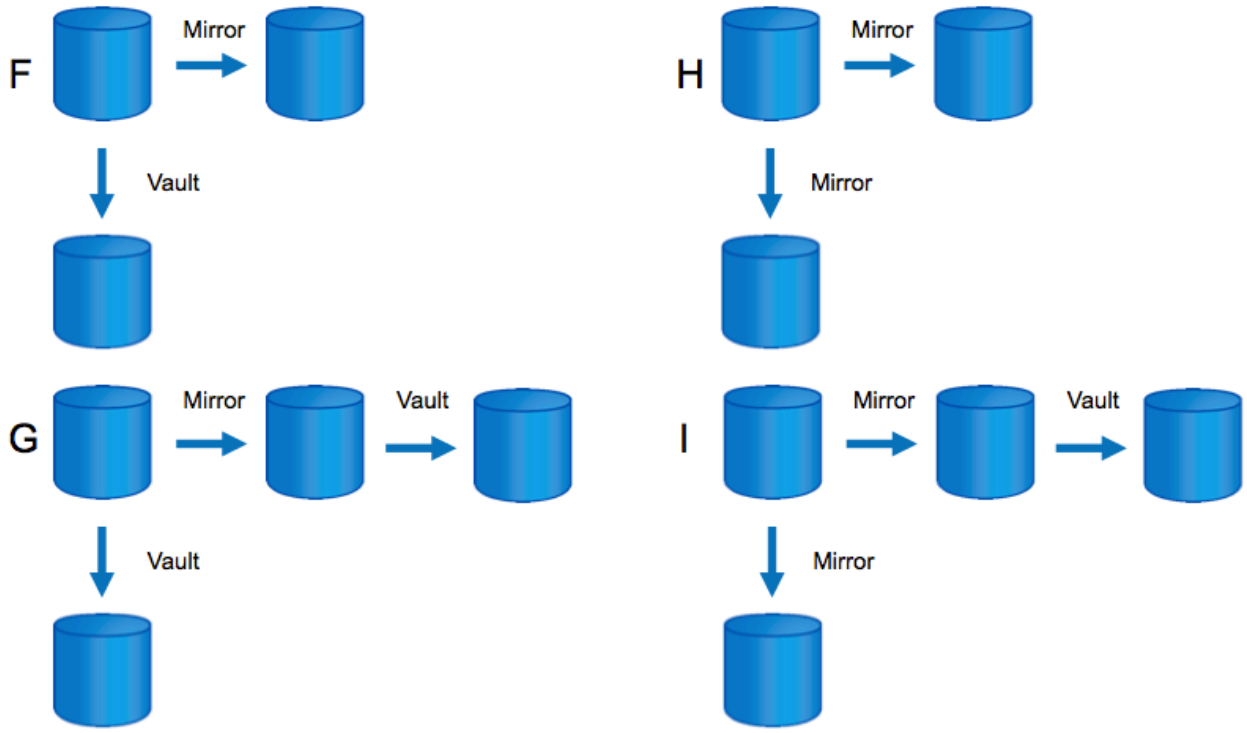
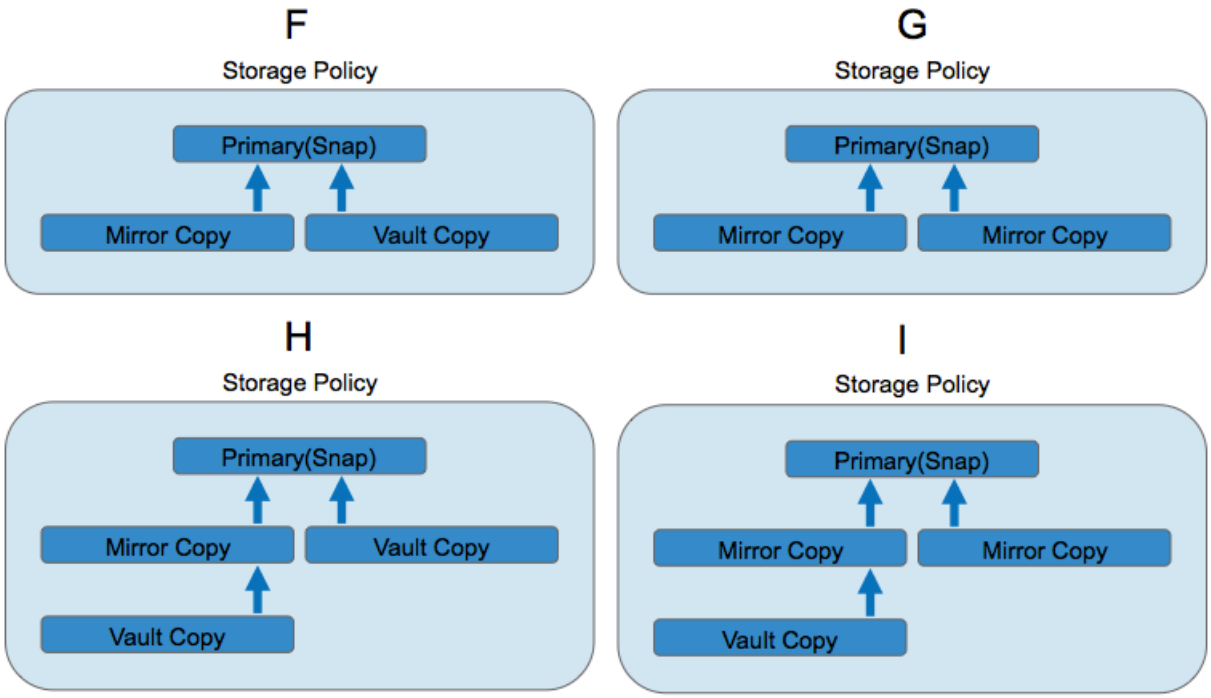


Figure 15 shows copy source dependencies for these fan-out examples.

Figure 15) Example fan-out dependencies within storage policies.



2.5 SCHEDULING AND RETENTION

Scheduling can be done by creating individual schedules or by creating schedule policies. A schedule policy groups various schedules together, each with its own properties. For example, a schedule policy might contain individual schedules for daily, weekly, and monthly backups. Traditional backup scheduling typically calls for weekly full backups and daily incremental backups. With NetApp Snapshot copy technology, however, the SnapProtect model consists almost entirely of full backups. An exception to this is when protecting NAS data. The indexing performance of NAS data backups increases significantly for incremental backups, while maintaining seamless searching for single file recoveries across Snapshot copies.

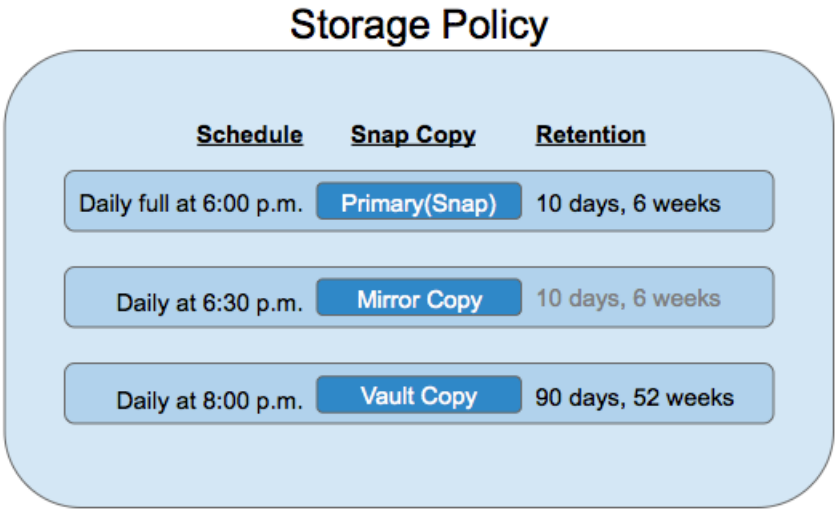
Auxiliary copies can be scheduled based on specific times, or they can be configured to run automatically.

Figure 16 shows a storage policy and its associated snap copies. In this example, a subclient is scheduled to perform a full backup (a local Snapshot copy) each day at 6:00 p.m. Retention for these local Snapshot copies is configured in the Primary(Snap) copy. A retention model of 10 daily backups and 6 weekly backups is established.

Mirroring takes place each day at 6:30 p.m. Retention for the mirror copy matches that of the primary copy.

Vaulting from the mirror copy is done each day at 8:00 p.m. Retention for the vault destination is configured in the vault copy. A retention model of 90 daily backups and 52 weekly backups is established.

Figure 16) Example schedules and retention at specific times.

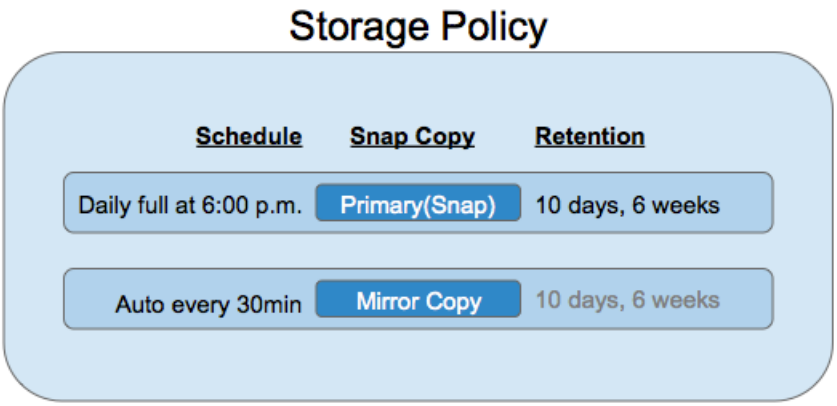


When multiple replication options are used together to protect data, they must be scheduled so that they do not run at the same time. The first replication operation needs to complete before the second operation starts.

Note: An auxiliary copy does not invoke a mirror or vault update if the job it depends on has not been completed.

Figure 17 shows a mirror copy configured to run automatically every 30 minutes. The mirror copy has nothing to do unless a new primary Snapshot copy has been created. So, although these jobs are initiated every 30 minutes throughout the day, in most cases they have nothing to do and do not invoke mirror transfers. This approach to scheduling offers a simplistic solution that can be very useful. This is especially true when many subclients are associated with the storage policy, all of which run at different times. However, with cascading replication copies, NetApp does not recommend the automatic approach. The copies must be scheduled in such a way that they do not run at the same time. The first replication operation needs to complete before the second operation starts.

Figure 17) Example schedules and retention, automatic auxiliary copies.



There are two types of retention rules in the SnapProtect software: basic retention rules and extended retention rules. Basic retention rules apply to daily or hourly backups. Extended retention rules apply to longer-term retention such as weekly full, monthly full, and yearly full backups. These rules are configured in the storage policy and can be set for the primary snap copy, vault copies, and tape copies. Mirror copies do not allow specific retention settings, because they inherit the same retention as the primary copy.

A cycle represents a full backup and the incremental backups that depend on that full backup. In many cases, full backups are used for every backup. However, for NAS data with millions of objects, a strategy that includes incremental backups reduces I/O and improves indexing performance. In addition, if backup jobs require incremental copies to tape, then the SnapProtect backups on the primary storage must include incremental jobs. Considering a full-backup-only paradigm, each backup can be considered a cycle. When incremental backups are included, all of the Snapshot copies in the cycle are retained until the last incremental in the cycle has expired. Performing more frequent full backups reduces the number of Snapshot copies associated with a cycle.

Basic retention rules allow retention entries for days and cycles. The default setting is 7 days and 2 cycles. When using full backups only, NetApp recommends setting days to 0 and defining only cycles.

Extended rules can be applied for longer retention. These rules include options to keep all full backups, weekly full backups, monthly full backups, quarterly full backups, half yearly full backups, and yearly full backups. Extended rules are not tied to a particular backup schedule. Rather, they are tied to full backups that start on a particular day of the week or day of the month. These days can be chosen as required.

To perform both hourly backups and daily backups, separate backup sets and storage policies must be created. One backup set includes a subclient with the hourly schedule and associated with one storage policy. The other backup set includes a subclient with the daily schedule and associated with the other storage policy. When running hourly backups, it is necessary to change the data aging schedule to run hourly instead of the default setting of once per day. The data aging operation is what expires backups and deletes Snapshot copies.

Note: Many of the application iDataAgents do not allow separate backup sets.

The following examples describe how to keep 6 hourly backups, 30 daily backups, weekly backups for 3 months, and monthly backups for 1 year on the NetApp primary system. These examples assume that only full backup jobs are being scheduled.

Hourly backups with 6 hour retention create a daily schedule for the subclient that repeats every hour, then set a basic retention rule in the primary snap copy of the associated storage policy to retain 0 days

and 6 cycles. By default, the data aging schedule runs once per day; therefore, in order to expire backups based on hourly retention, the data aging schedule needs to run hourly.

Daily backups with 30 day retention create a daily schedule for the subclient that repeats every day, then set a basic retention rule in the primary snap copy of the associated storage policy to retain 0 days and 30 cycles.

Retain weekly backups for 3 months retains one daily backup every week and keeps it for 90 days. This requires an extended retention rule. Set an extended retention rule in the primary snap copy of the associated storage policy to retain weekly full backups for 90 days and set the rule to start on the appropriate day of the week. Every daily full backup created on this day of the week is retained for 90 days.

Retain monthly backups for 1 year retains monthly backups by retaining one daily backup every month and keeping it for 365 days. This requires an extended retention rule. Set an extended retention rule in the primary snap copy of the associated storage policy to retain monthly full backups for 365 days and set the rule to start on the appropriate day of the month. Every daily full backup created on that day of the month is retained for 365 days.

For replication (mirroring and vaulting), similar methods can be used to schedule replication and vault retention. However, scheduling is set on the mirror or vault copy in the storage policy rather than the subclient; for vaulting, retention is set in the vault copy in the storage policy.

Table 2 is an expanded example in which virtual machines (VMs) in VMware® datastores are protected with various requirements. Four different storage policies are required because there are mixed retention requirements across the datastores.

Table 2) Scheduling and retention examples.

Backup set	Data-store	Sub-client	Storage policy	Backup schedule	Local retention	Mirror schedule	Mirror retention	Vault schedule	Vault retention
A	DS1	SC1	SP1	Daily full at 6 p.m.[1]	10 days (cycles), 6 weeks, set in primary snap copy	Daily at 6:30 p.m., set in mirror copy schedule	10 days, 6 weeks	Daily at 8 p.m., set in value copy schedule	90 days (cycles), 52 weeks, set in vault copy
A	DS2	SC2	SP2	Daily full at 6 p.m.[1]	30 days (cycles), 8 weeks, set in primary snap copy	Daily at 6:30 p.m., set in mirror copy schedule	30 days, 8 weeks	Daily at 8 p.m., set in vault copy schedule	180 days (cycles), 52 weeks, set in vault copy
A	DS3[2]	SC3	SP3	Daily full at 6 p.m.[1]	10 days (cycles), set in primary snap copy	Daily at 6:30 p.m., set in mirror copy schedule	10 days	Daily at 8 p.m., set in value copy schedule	90 days (cycles), 52 weeks, set in vault copy
B	DS3[2]	SC4	SP4	Hourly, except 6 p.m.	23 hours (cycles), set in primary snap copy	-	-	-	-

[1] The local Snapshot schedules can be set at the subclient level or at the backup set level. In the example in Table 2, all local Snapshot copies run at 6 p.m.; therefore a single schedule at the backup set level could be used. If subclients in a backup set require different local Snapshot schedules, then the schedules need to be set at the subclient level.

[2] Datastore DS3 is defined in two backup sets (A and B), because of the need to perform both daily backups and hourly backups for this datastore. Therefore it is necessary to have the same datastore in two subclients in order to make basic retention rules work. Using two backups sets in this case enables retention for both hourly and daily backups.

2.6 UNDERSTANDING THE RESTORE WORKFLOW

With SnapProtect management software, recovery is simple; data can be restored from virtually any backup copy in a single operation. Restores from recent backups might come from local Snapshot copies, while historical data might come from vault or tape copies.

Data to be restored can be located either by browsing or by using the search feature. When the data is located, a restore job can be initiated. Restores can be done from any of the backup copies by browsing data from a particular copy. The copy order precedence is defined on the Copy Precedence tab of the storage policy properties.

For volumes and LUNs on NetApp primary storage, it is also possible to revert from a Snapshot copy. This feature uses NetApp SnapRestore® data recovery software to revert a volume or LUN back to a particular point in time. This feature should be used with caution, because a revert affects all data in the volume or LUN. To initiate a revert, right-click the subclient and select List Snaps. From the list, right-click a Snapshot copy and select “Use hardware revert capability if available.”

Because SnapProtect uses NetApp Snapshot technology, it is possible to copy data directly from a Snapshot copy via CIFS or NFS.

2.7 DATA CLONING

SnapProtect software enables administrators to create data clones, which allow read/write access to the backup data. Data clones can be used for a variety of purposes. NAS data can be cloned when using the NetApp NAS NDMP iDA. This functionality creates a NetApp FlexClone® volume and makes the entire contents of the Snapshot copy accessible.

By using a File System iDA, LUN data can be cloned as well. This functionality, when performed on a primary Snapshot copy, uses LUN clone technology. When cloning LUN data from a Snapshot copy on a secondary or tertiary NetApp system, a FlexClone volume is created.

To create data clones, right click the subclient and select List Snaps. From the list, right-click a Snapshot copy and select Mount.

3 APPLICATION DATA

SnapProtect can be used to protect applications running on physical servers and hosted on NetApp primary storage. For each supported database application there is an associated iDataAgent. This iDA must be installed on the client system that is running the application. The iDAs prepare the database applications for backup consistency. In addition, they handle things like log truncation during backup, database storage mapping, and log manipulation during restore.

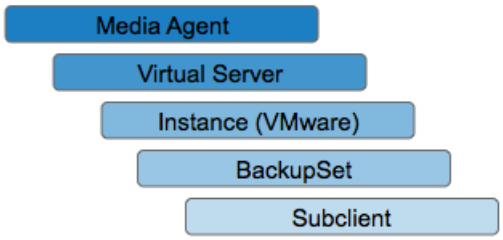
4 VIRTUALIZATION DATA

A key feature of SnapProtect management software is the ability to protect many virtual machines very quickly. In addition, it can index the contents of each VM, and it allows different levels of recoverability, including single file recovery.

SnapProtect software is flexible and allows discovery rules to be established so that new virtual machines can be automatically added to a subclient and protected. For example, using a discovery rule of Datastore Affinity automatically protects new virtual machines on specific datastores.

SnapProtect software uses the Virtual Server Agent (VSA) to perform the data protection operations for virtual environments. The VSA is installed on a system configured as a media agent. Within the VSA, instances are created that define the type of virtualization solution being used. In a VMware environment, a VMware instance would be created under the VSA. Within the instance, a backup set contains the subclients. Figure 18 shows the VSA layout.

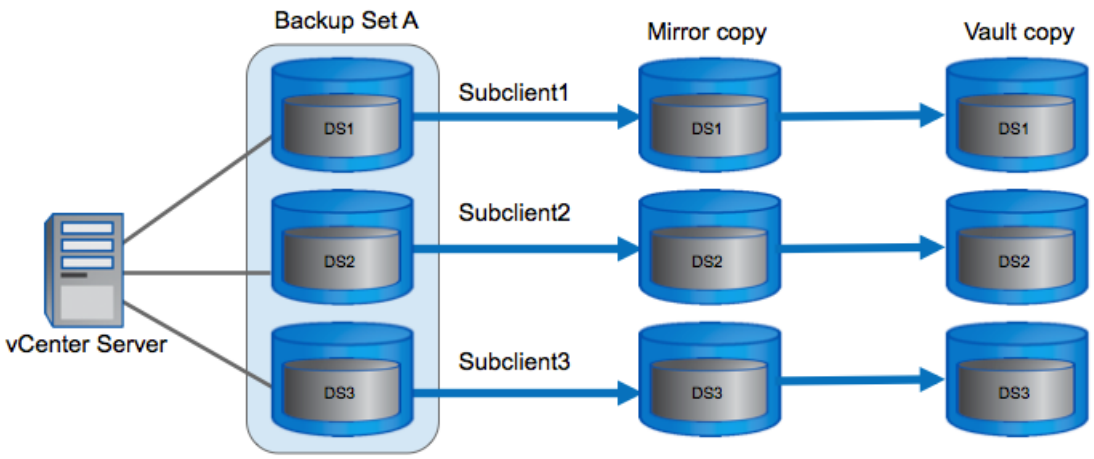
Figure 18) Virtual Server Agent.



Note: Because of the advantages of VMware HotAdd transport mode during restores, NetApp recommends installing the VSA on a virtualized media agent. This virtualized media agent should run on an ESX host that has access to the production datastores, such as an ESX proxy host.

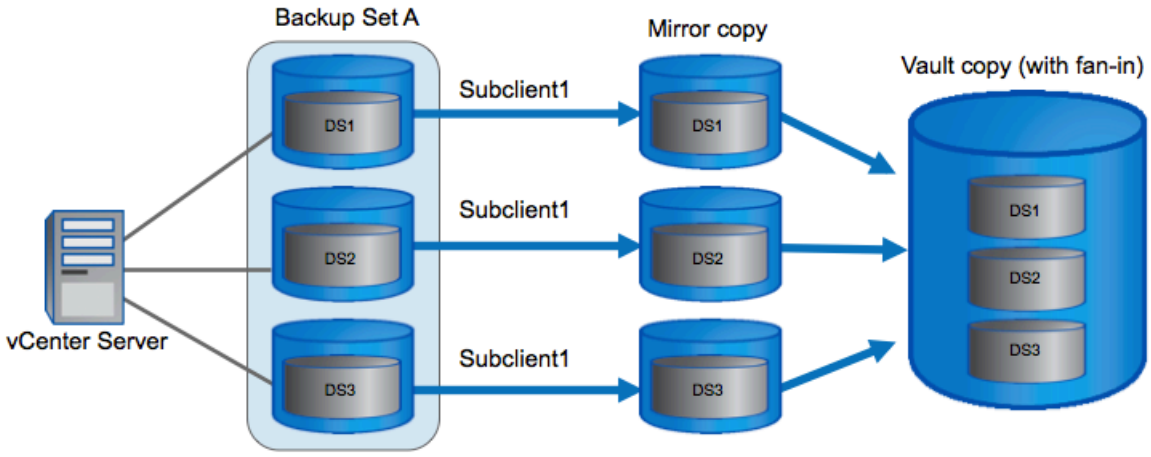
In the example shown in Figure 18, multiple datastores are grouped into a single backup set. However, because the datastores have different scheduling and retention requirements, they are separated into their own subclients (each subclient is associated with a different storage policy). The datastores are mirrored and then vaulted.

Figure 19) Datastores in separate subclients.



In Figure 19, the datastores are grouped into a single backup set and a single subclient. In this example, the datastores have the same scheduling and retention requirements. The datastores are mirrored and then vaulted. Because the datastores are grouped into the same subclient, it is possible to do a fan-in on the vault copy.

Figure 20) Datastores in the same subclient.



If you plan to use VMware vSphere™ Site Recovery Manager, the first replication copy from the primary must be a mirror copy. This allows failover capability as well as the advanced deduplication and WAN efficiencies of SnapMirror for site-to-site replication. With SnapVault, the data can be combined by using a fan-in approach that complements deduplication on the tertiary copy.

Note: VMs cannot have virtual disks that reside in multiple datastores.

Backup settings allow different granularity for restore operations. During restore operations, data for the VMs can be browsed and recovered based on the recovery type selected. A container restore can be performed to recover an entire VM or individual VMDK files. Individual files and folders can be restored to a staging location.

VMWARE AND APPLICATIONS

When run inside a virtual machine, Microsoft Exchange and SQL Server have integration with VSS that allows database consistency during the backup of the virtual machine. The File System iDA and the VSS Provider must be installed on the guest OS to get this functionality. To enable these application-consistent backups, make sure that the “Application aware backup for granular recovery” box is checked under the SnapProtect Operations tab for the subclient. Exchange backups offer the additional option to perform log truncation as part of the backup operation; select Truncate ExDB Logs.

Consistent out-of-place restores of SQL Server and Exchange databases can be performed by restoring the flat database files. The Exchange Offline Mining tool is a standalone utility that allows individual message restores from a backup copy of the Exchange database.

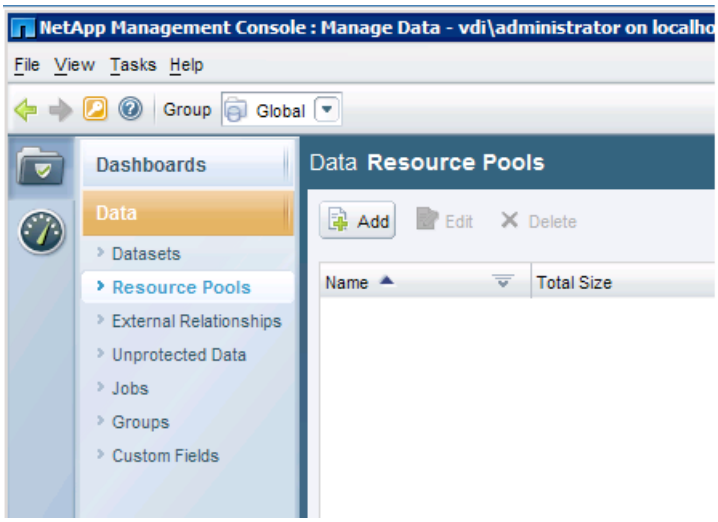
5 BASIC SETUP

The following list outlines the workflow to set up a basic SnapProtect environment. The following subsections contain additional details for many of these steps.

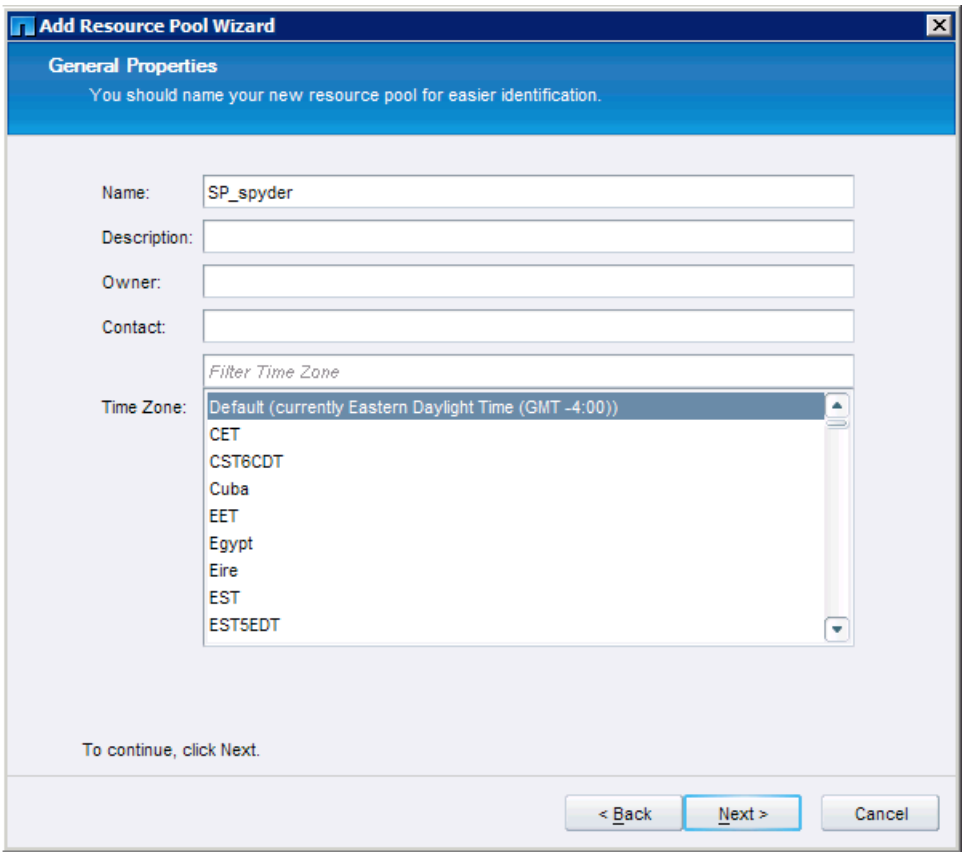
1. Add any required licenses to the NetApp systems in the environment (FlexClone, SnapVault primary, SnapVault secondary, SnapMirror, and so on).
2. Make sure that SnapVault and SnapMirror are enabled on the NetApp systems and that the `options snapvault.access` and `options snapmirror.access` settings are set to allow connectivity.
3. Provision storage for the disk library on the NetApp system for SnapProtect metadata to be stored. Connect this storage to the SnapProtect server. This could be a CIFS share, a LUN, or a virtual disk if the SnapProtect server is a virtual machine.
4. Install the DataFabric Manager server software (on its own system or virtual machine).
    a. Include required DataFabric Manager licenses (Core, Protection Manager, Provisioning Manager).
    b. Make sure that DataFabric Manager discovers the NetApp systems that are using the NMC.
    c. Update login and NDMP credentials for the NetApp systems in the NMC.
5. Install the NetApp Management Console on its own system or virtual machine.
6. Create required resource pools in the NMC for replication to secondary / tertiary storage.
7. Optionally, create any custom provisioning policies by using the NMC.
8. Install the SnapProtect software.
9. Add the DataFabric Manager server, the primary NetApp system, and any secondary / tertiary NetApp systems in SnapProtect array management.
10. Create storage policies in the SnapProtect software as needed.
11. Add clients and iDataAgents as needed: Right-click CommServe → All Tasks → Add/Remove Software → Install Software. Make sure that hostnames are entered using their fully qualified name.

5.1 CREATE A RESOURCE POOL BY USING THE NMC

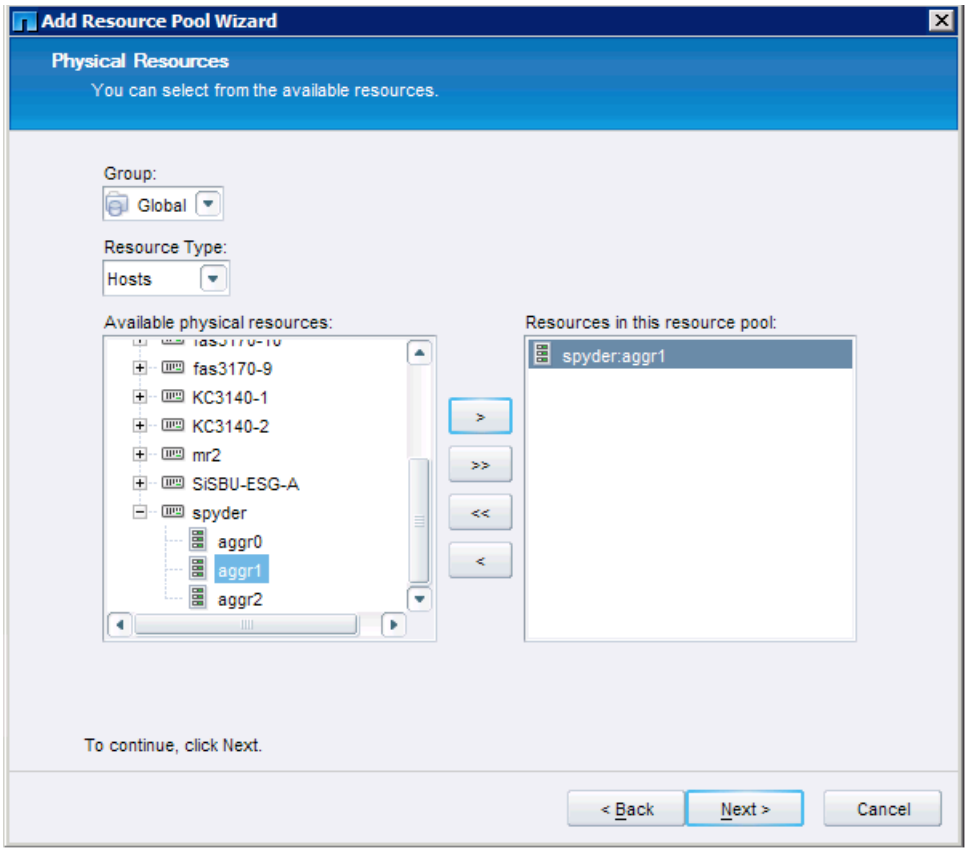
Open the NMC and connect to the DataFabric Manager server. To create a resource pool by using the NetApp Management Console, click Data → Resource Pools → Add to launch the wizard.



Give the resource pool a name.



Choose an aggregate from the destination system to be assigned to the resource pool.

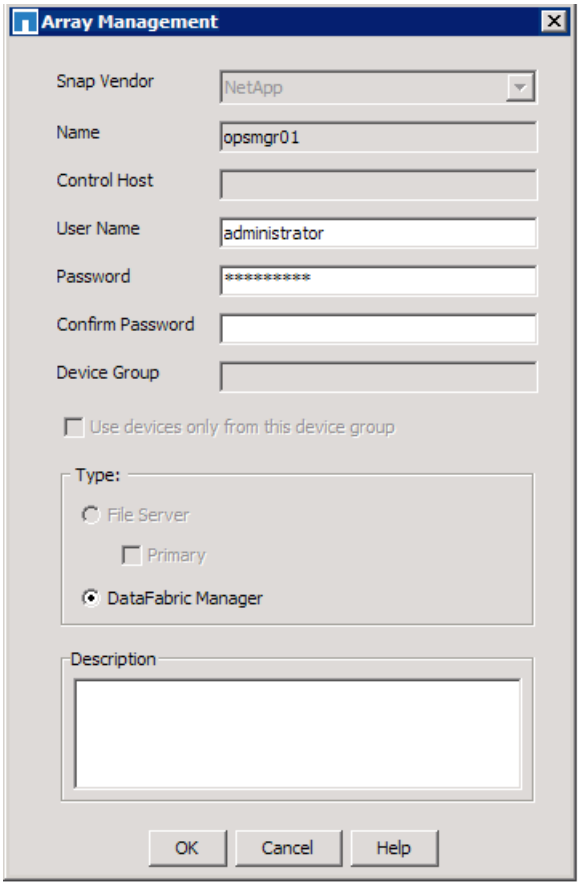


Finish the wizard.

5.2 ADD NETAPP SYSTEMS AND THE DATAFABRIC MANAGER SERVER TO THE SNAPPROTECT SOFTWARE

Before any SnapProtect operations can be performed, information about the NetApp systems and the DataFabric Manager server must be configured. There are a couple ways to add the NetApp systems and the server to the SnapProtect environment. One way is to do this is to use the control panel before any other configuration. To open the control panel, right-click CommServe → Control Panel. Under the control panel, open the Array Management utility.

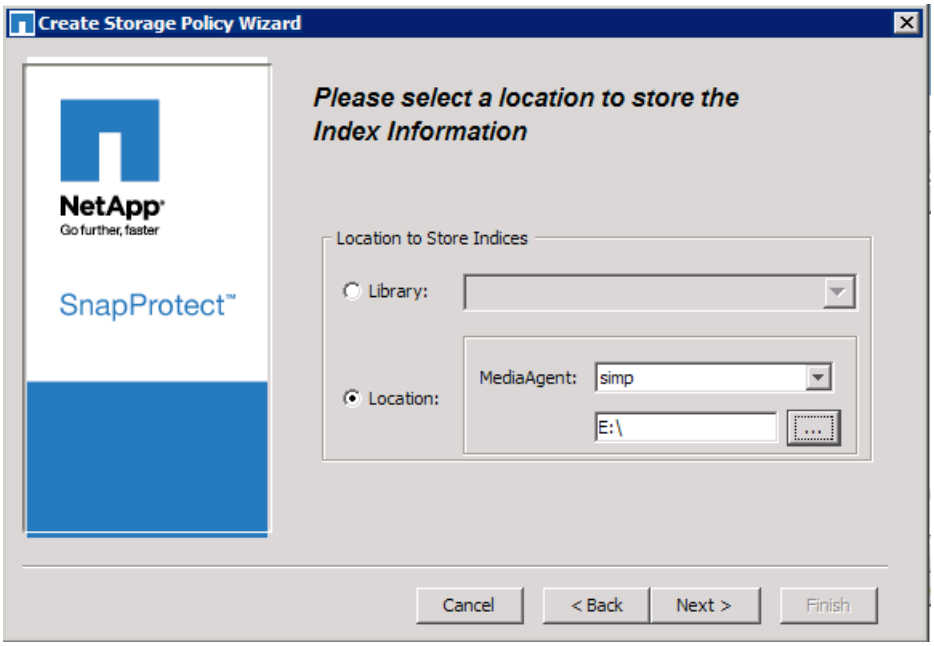
To add the DataFabric Manager server, click Add and populate the Array Management properties. Set the type to DataFabric Manager.



To add a NetApp primary, set the type to File Server → Primary. To add a NetApp secondary or tertiary, set the type to File Server and leave the Primary check box unchecked.

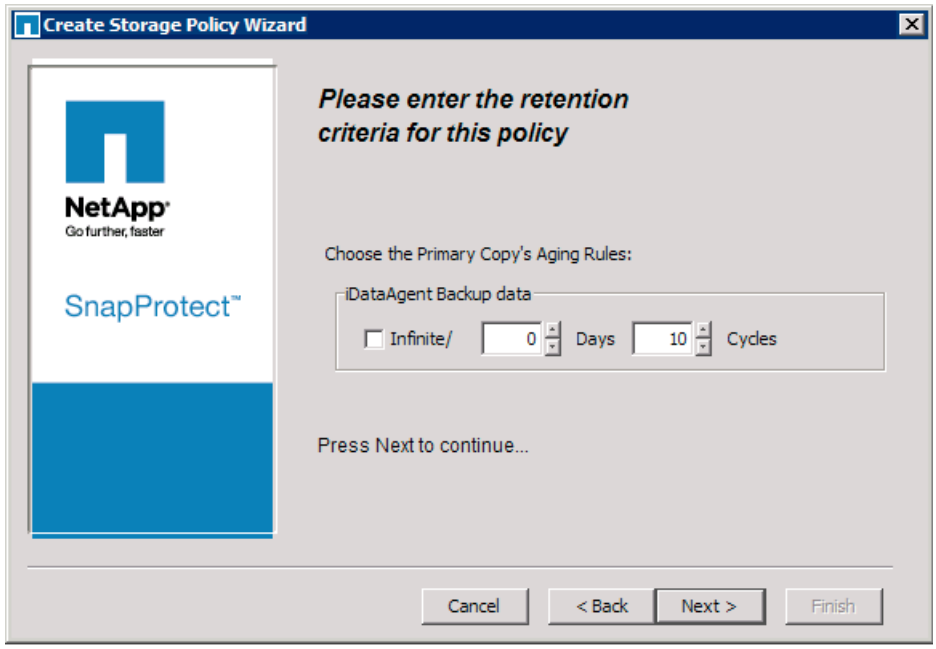
5.3 CREATE A STORAGE POLICY

To create a storage policy in the SnapProtect software, right click Storage Policies → New Storage Policy. Give the policy a name. Enter a location for the index information to live. In this example, the E:\ drive maps to a LUN on the NetApp primary system.

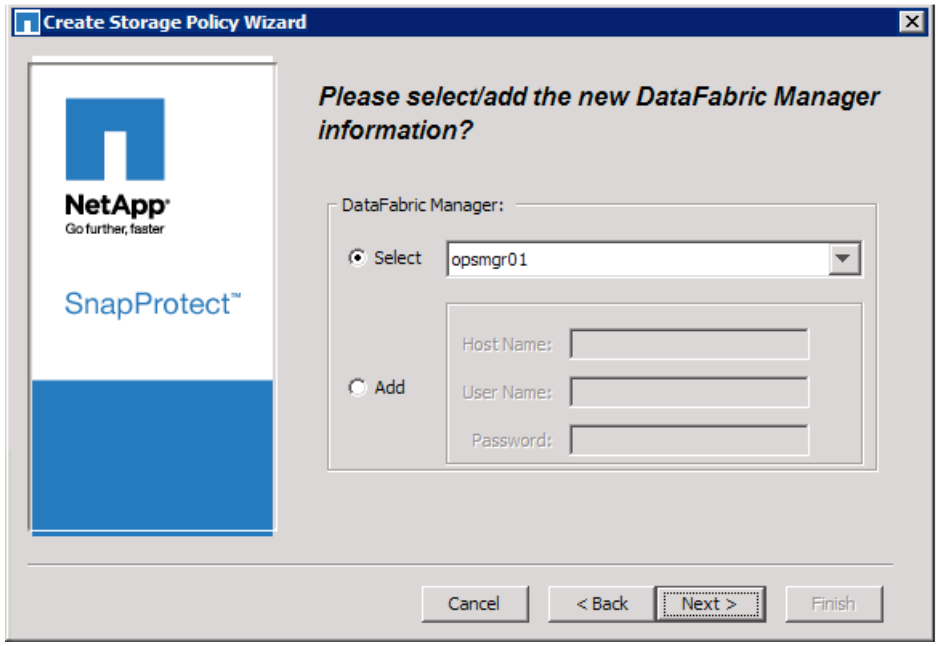


Note: if CIFS is preferred to store the indexes, a disk library must be created manually before creating the storage policy. This disk library can then be selected from the Library drop-down list when the storage policy is created.

On the retention screen, change the retention as needed (for example, 0 days and 10 cycles). This setting can be changed later under the Retention tab for the Primary(Snap) copy in the storage policy.



Select the DataFabric Manager server that was added in section 5.2.



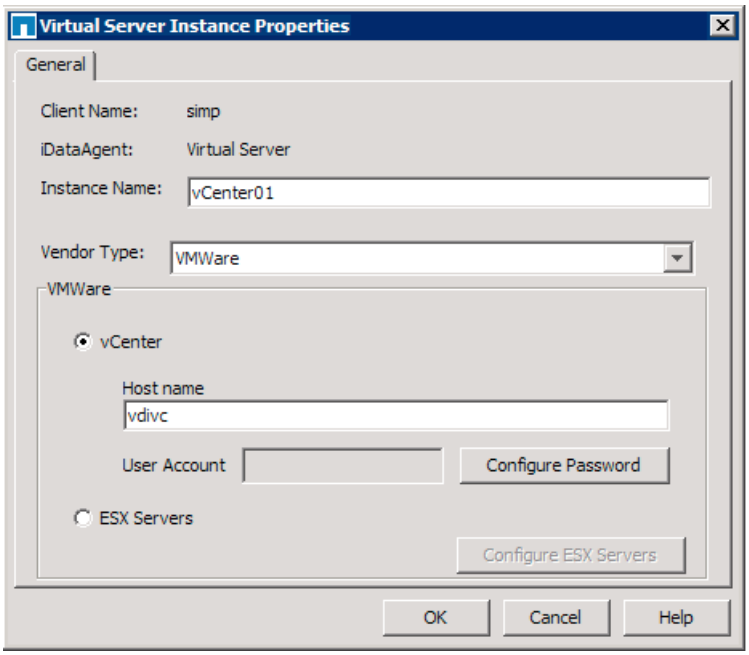
5.4 VMWARE BACKUPS

Before configuring VMware backups, make sure that the Virtual Server iDA is installed on a virtualized media agent. To push iDAs from the CommServe server, right-click CommServe → All Tasks → Add/Remove Software → Install Software and follow the prompts. The Virtual Server Agent module is listed under Client Modules → Backup & Recovery → File System.

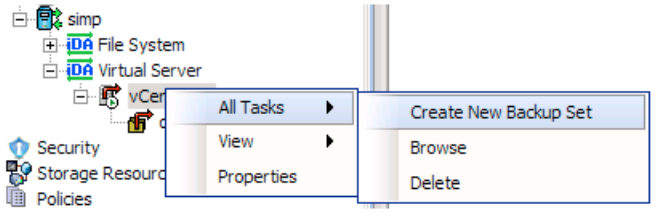
Note: Because of the advantages of VMware HotAdd transport mode during restores, NetApp recommends installing the VSA on a virtualized media agent. This virtualized media agent should run on an ESX host that has access to the production datastores, such as an ESX proxy host.

To configure backups, a VMware instance must be created. Right-click Virtual Server iDA → Create New Instance.

Give the instance a name and select VMware for the vendor type. For vCenter™, provide the hostname and login credentials.

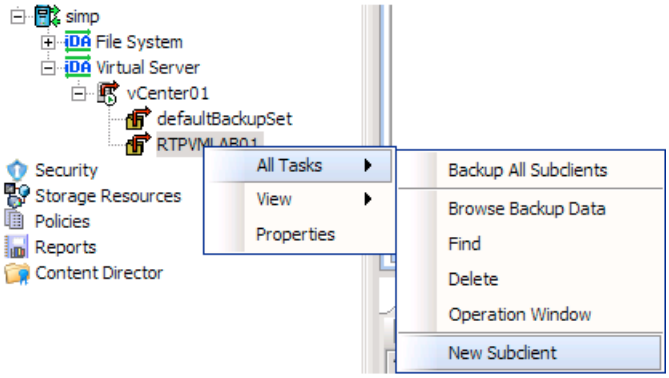


After creating the VMware instance, create a new backup set. Right-click the instance → All Tasks → Create New Backup Set.

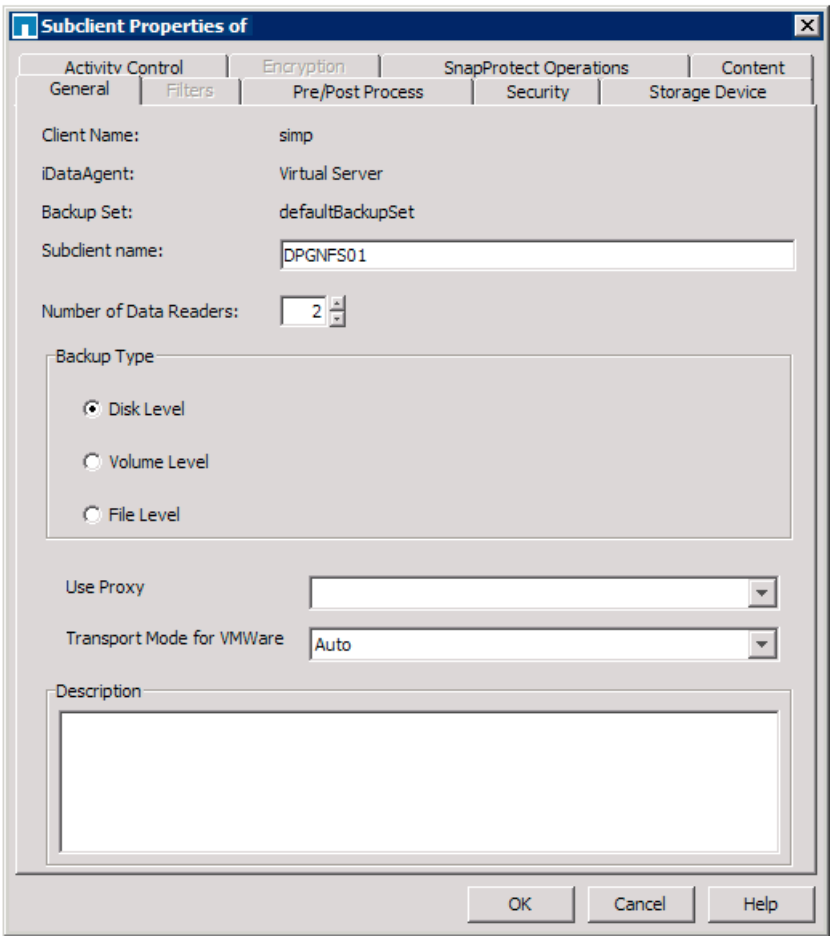


Give the backup set a name and select the storage policy. Do not schedule. Scheduling is configured later.

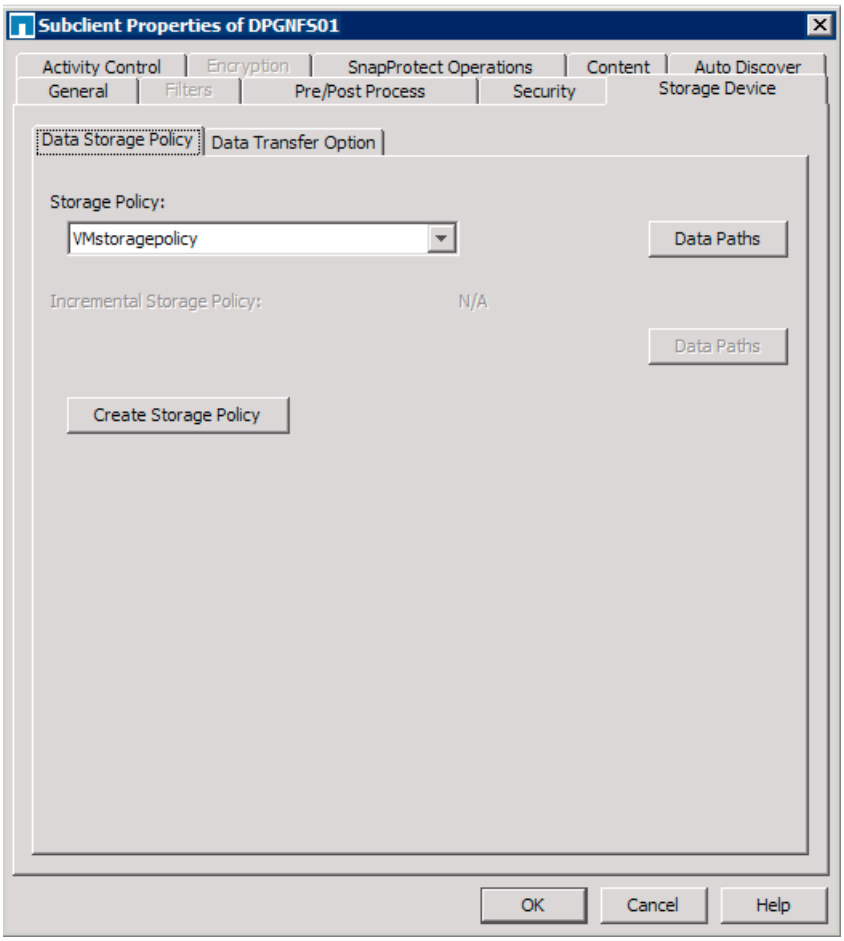
Create a new subclient in the backup set: right-click the backup set → All Tasks → New Subclient.



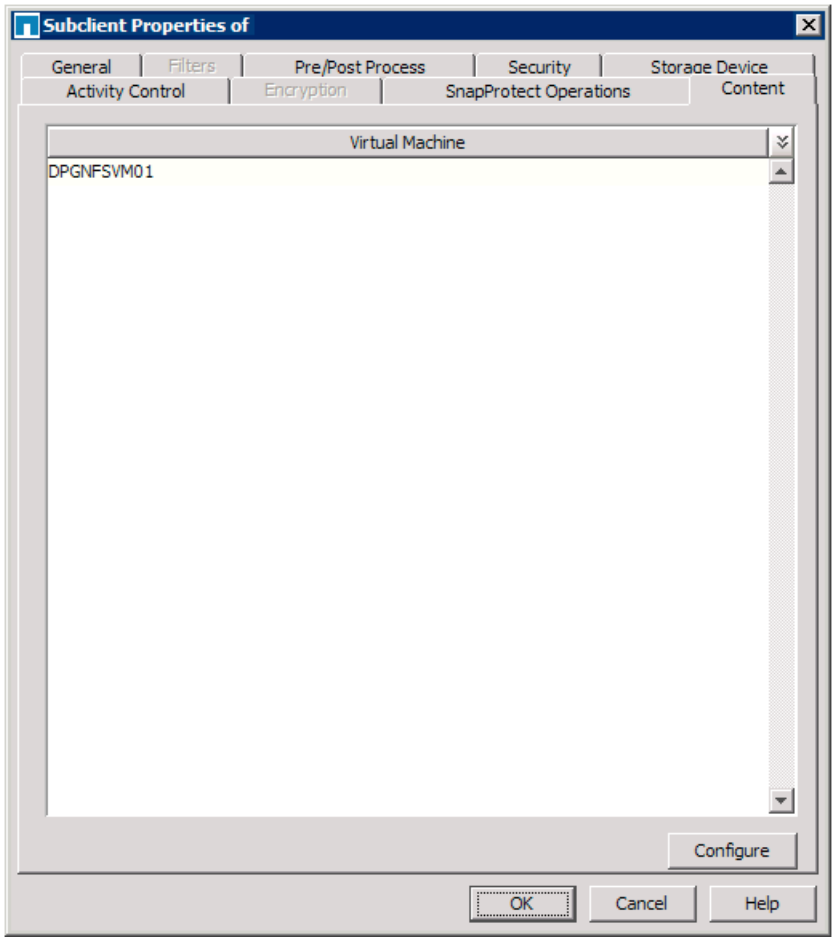
Give the subclient a name.



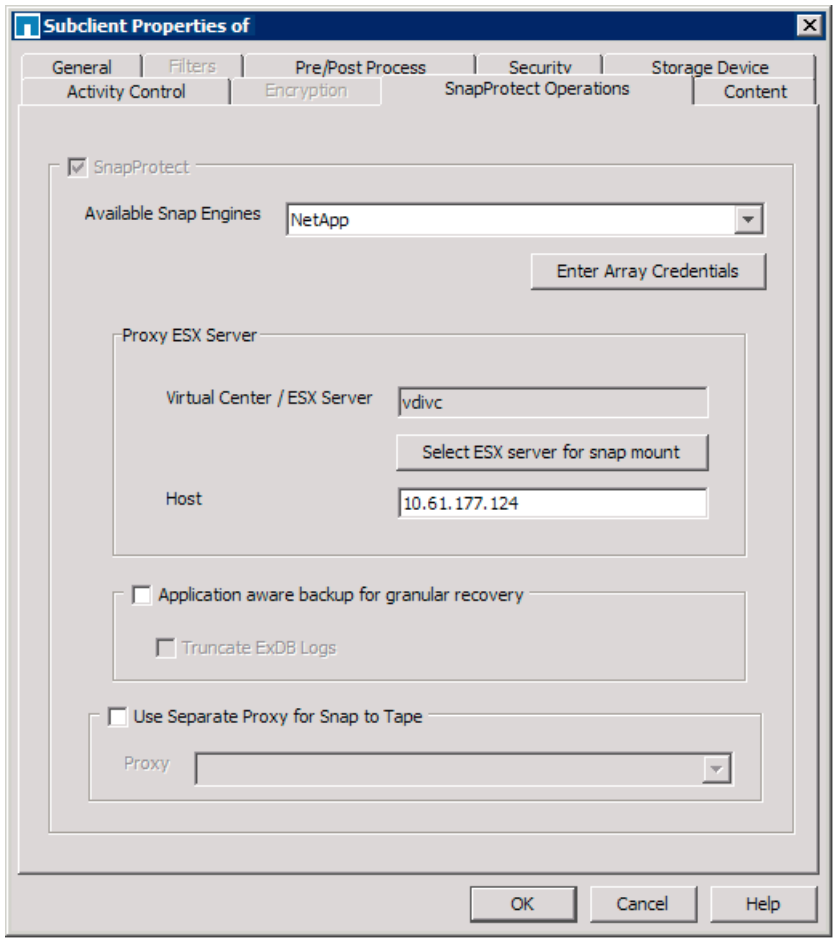
Click the Storage Device tab and select the storage policy.



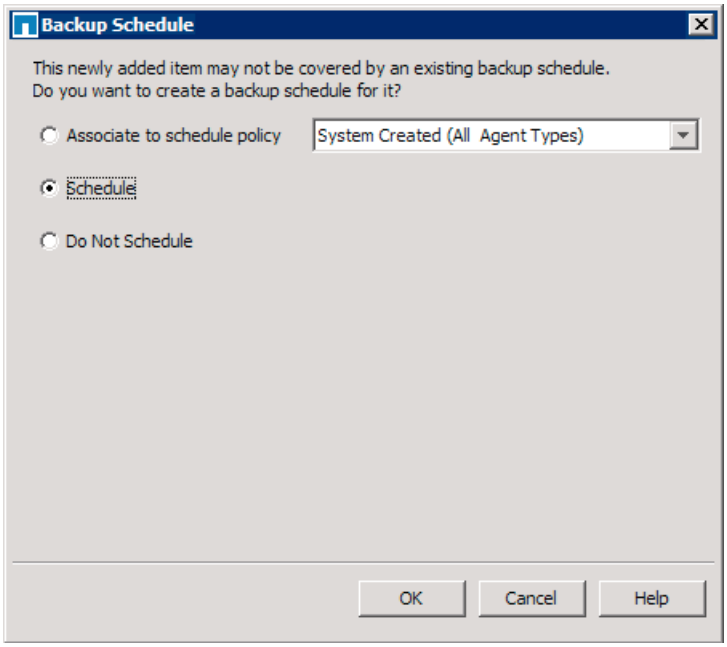
Under the Content tab, assign the VMs in the datastore to the subclient by using the Configure and Discover buttons.



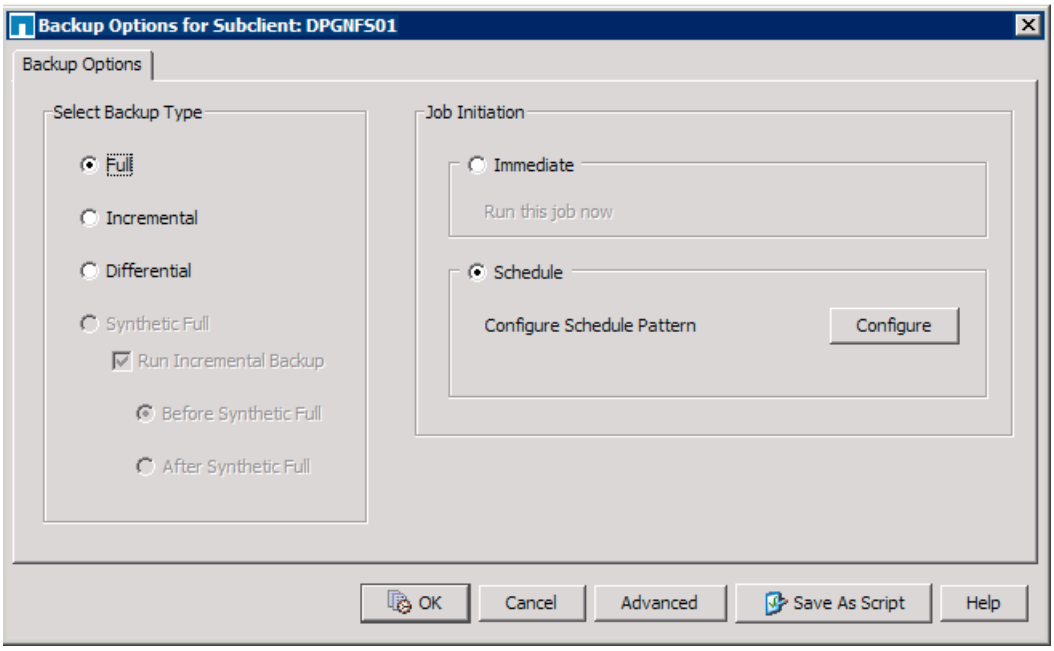
Under the SnapProtect Operations tab, select an ESX server to be responsible for mounting Snapshot copies as needed. Click OK.



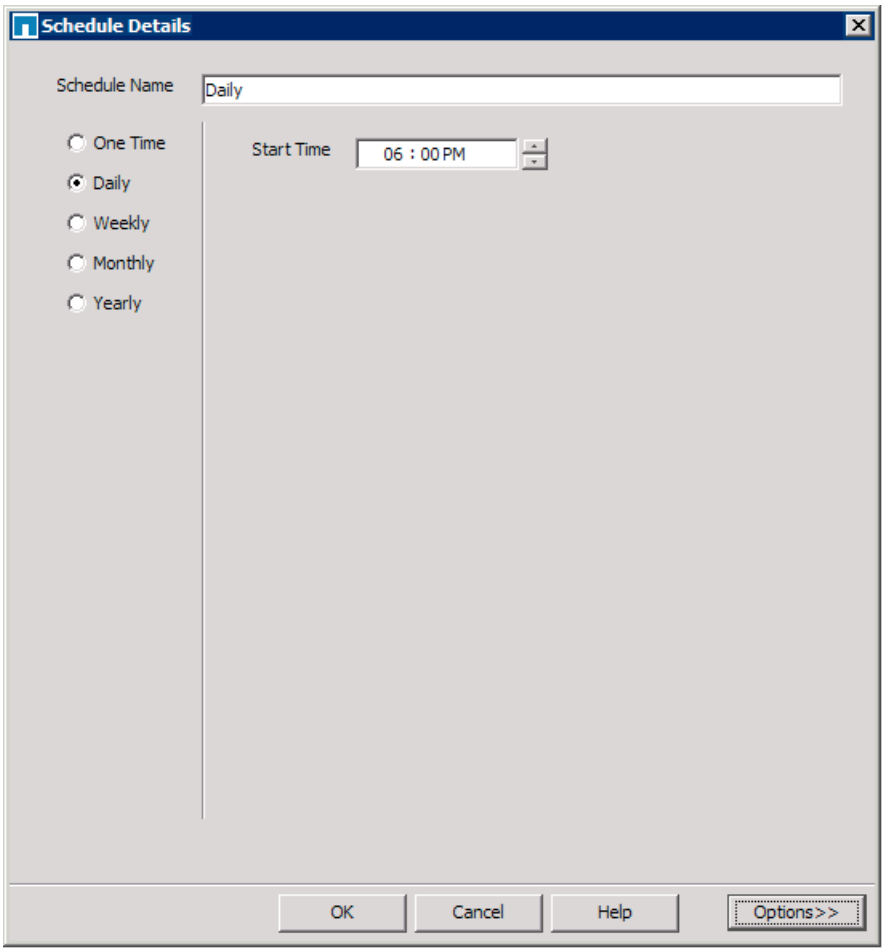
Select Schedule to create a schedule for the backups. Click OK.



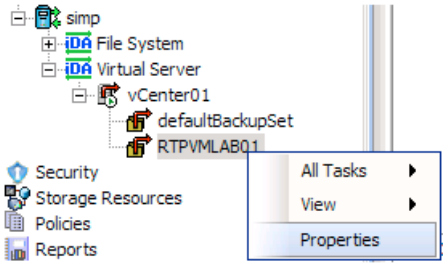
Select Full for the backup type. Under Job Initiation, select Schedule. Click Advanced.



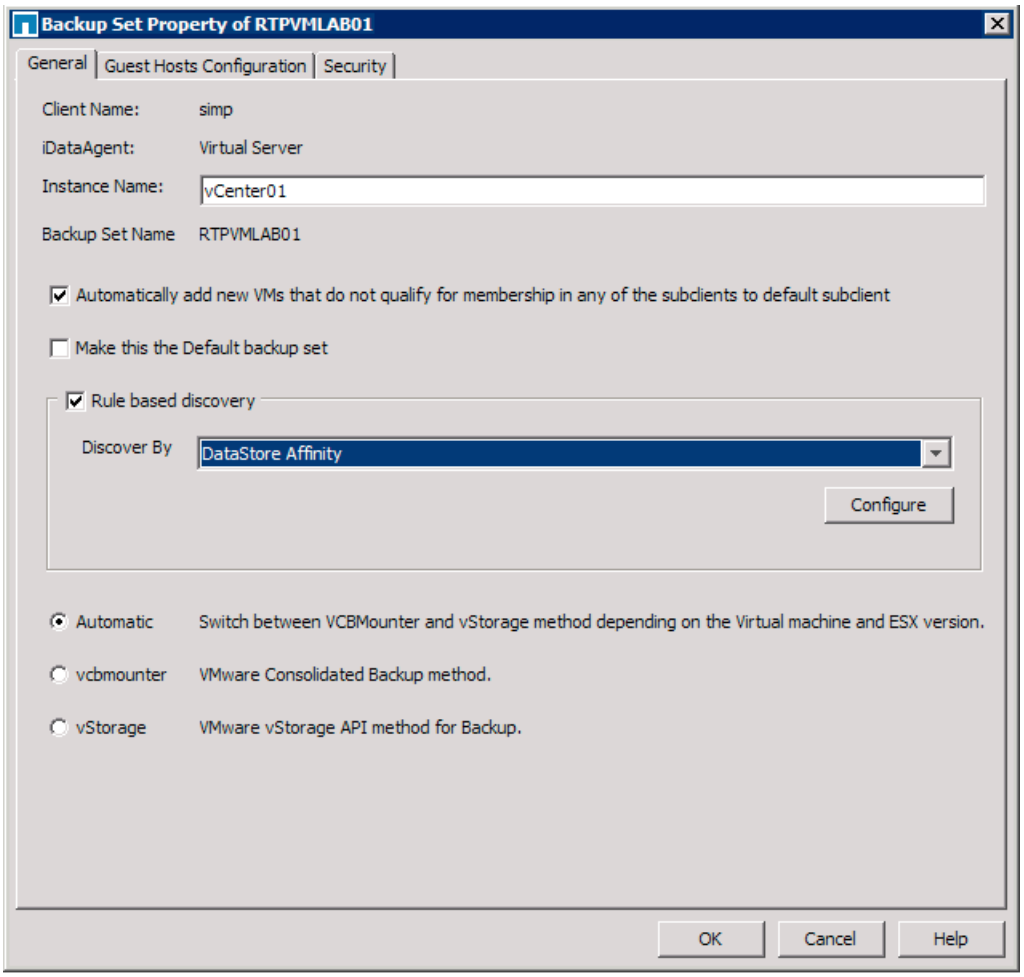
In this example, daily backups will be done at 6:00 p.m.



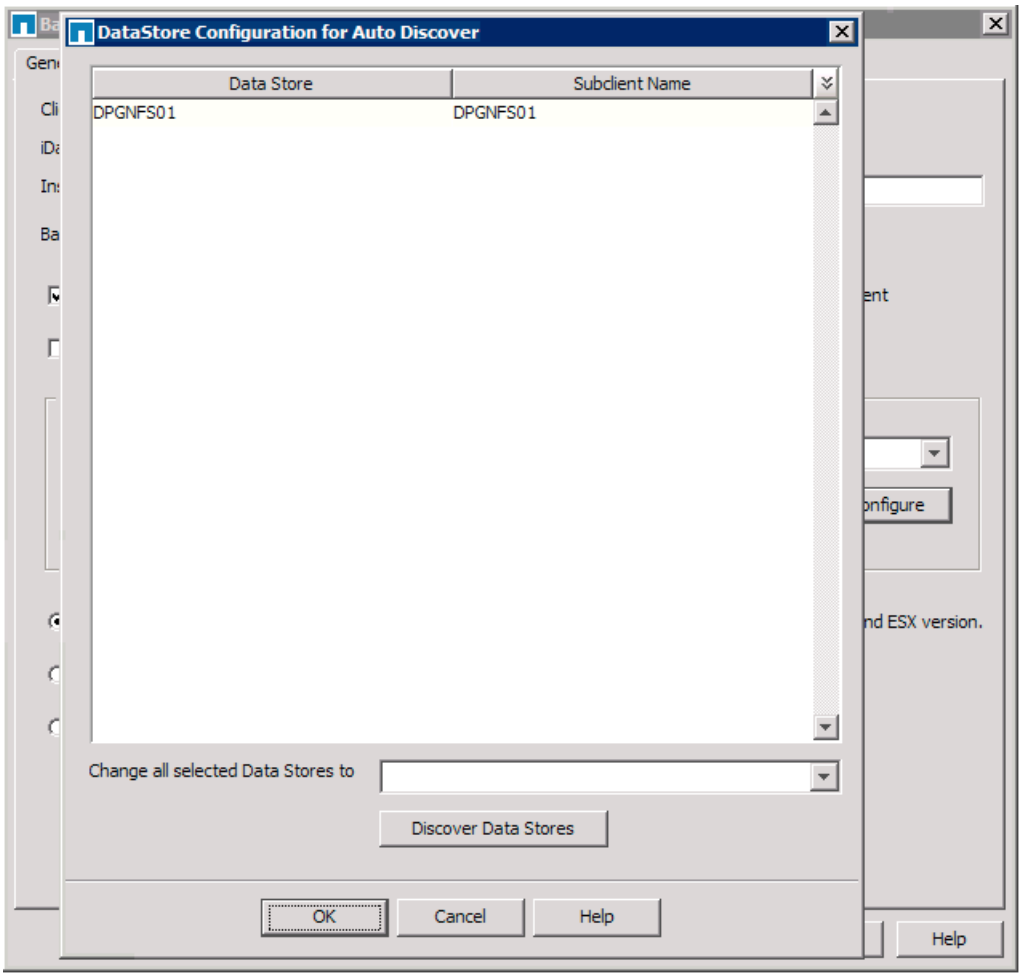
To set up discovery for the datastore so that new VMs are automatically added to the subclient, configure datastore affinity in the backup set. Right-click the backup set → Properties.



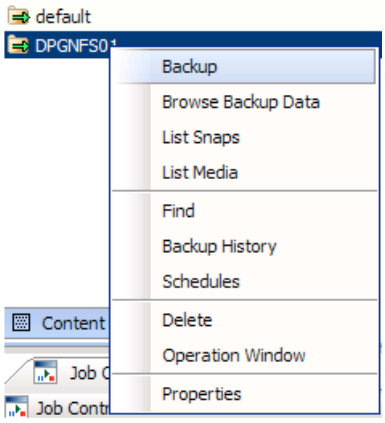
Enable rule-based discovery and select DataStore Affinity. Click Configure.



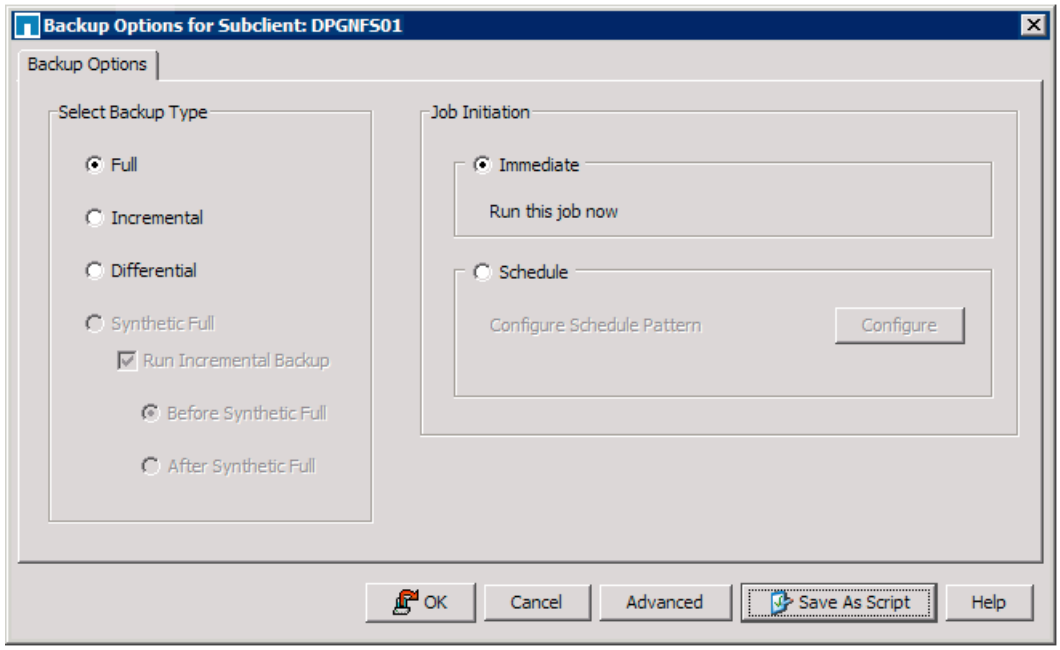
Use the Discover Data Stores button to assign the datastore to the subclient.



To run a manual backup, right-click the subclient → Backup.



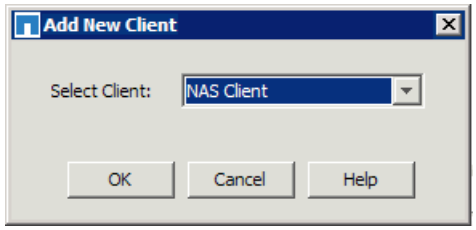
Select Full backup. Select Immediate and click OK to start the backup job.

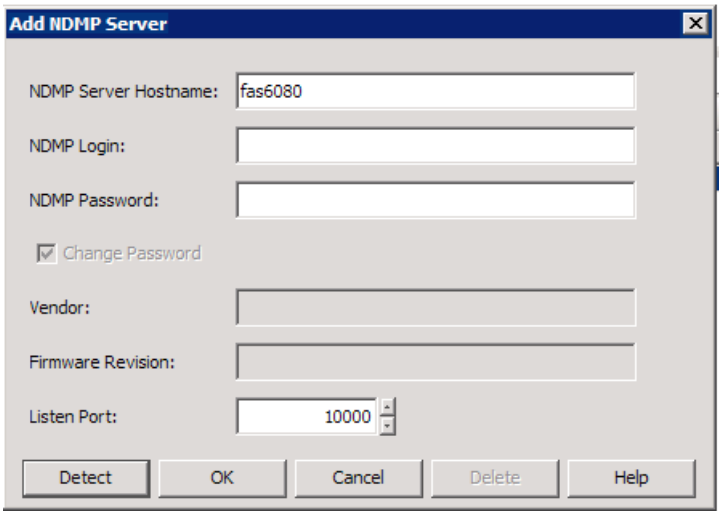


5.5 NAS DATA

Section 5.4 covered many of the fundamental steps to configure backups in general. This section covers steps that are specific to configuring NAS data backups.

To configure backups for a NAS volume, add the NetApp primary system as a client in the SnapProtect software. To add the primary system, right-click Client Computers → New Client. Select NAS Client and click OK. In the Add NDMP Server window, enter the NDMP login credentials and click OK.





The primary system is added as a client with an associated NetApp NAS NDMP iDataAgent. Expand the NetApp NAS NDMP iDataAgent. Create a new backup set and a new subclient.

Populate the subclient properties tabs as follows:

General. Give the subclient a name.

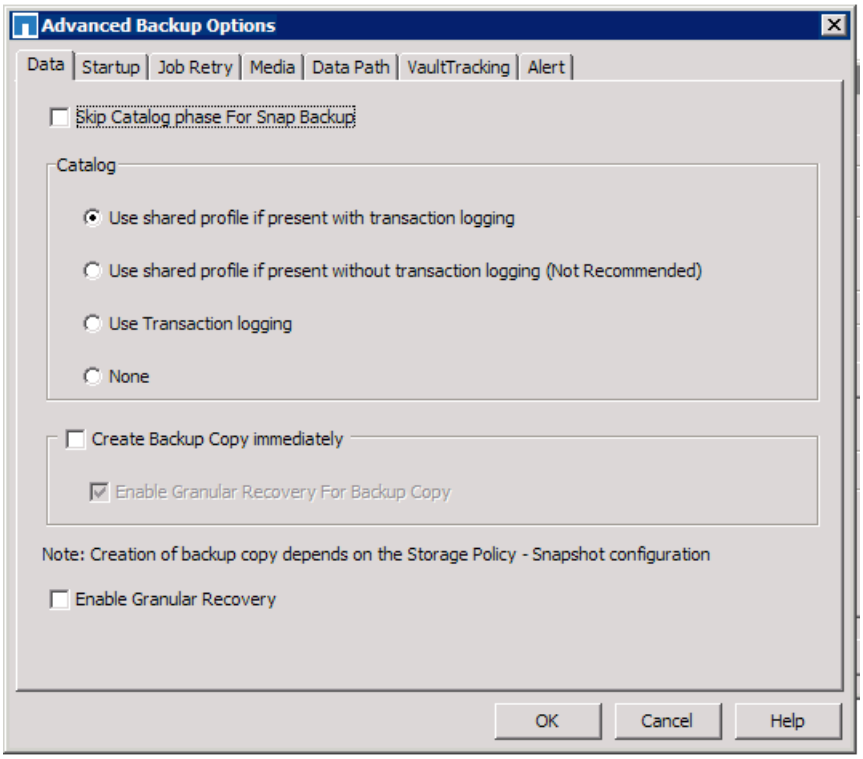
Content. Browse the primary system and select the volumes to be protected by this subclient.

SnapProtect Operations. Optionally select an alternate media agent to act as a proxy for indexing.

Storage Device. Select the storage policy to be used by this subclient.

When the subclient properties are populated, click OK. Create backup schedules as needed.

Indexing NAS data backups (primary Snapshot copies) is optional and is disabled by default. To perform indexing during NAS data backups, uncheck Skip Catalog Phase For Snap Backup. This option is listed under the advanced backup options when initiating a backup job and can be made permanent for schedules.

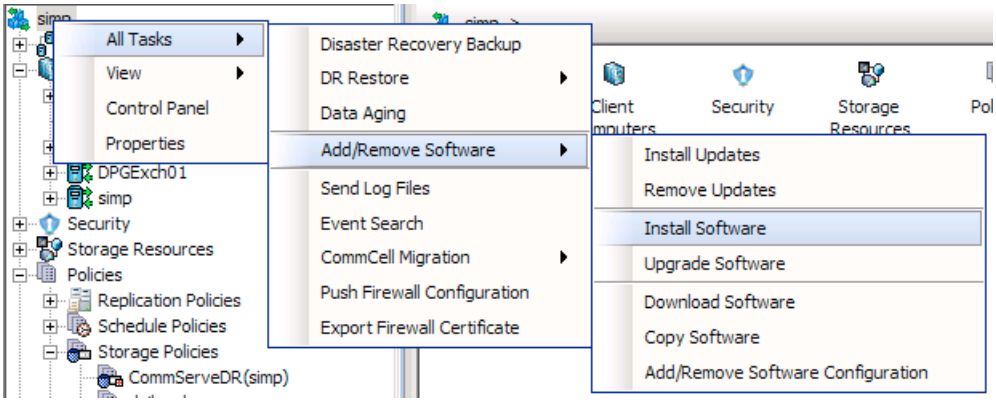


5.6 LUN DATA

This section covers steps that are specific to configuring LUN data backups.

To protect LUN data, a File System iDataAgent on a client is used. For example, a LUN mapped to the F:\ drive on a Windows host can be protected after adding the Windows host as a client in the SnapProtect software.

The client must have the File System iDA, media agent binaries, and the VSS provider installed. To add the client and install the software, right-click CommServe → All Tasks → Add/Remove Software → Install Software. Follow the steps in the wizard to install the required components for the host.



After adding the client, expand the File System iDA. Create a new backup set and subclient.

Populate the subclient properties tabs as follows:

General. Give the subclient a name.

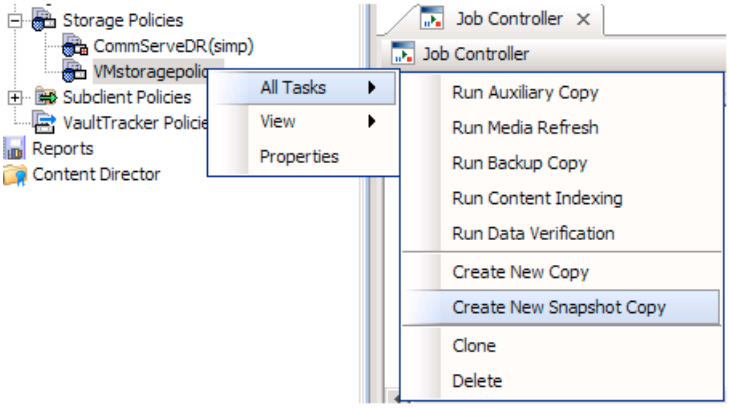
Content. Browse the primary system and select the drive for backup (F:\ drive in this example).

Storage Device. Select the storage policy to be used by this subclient. If a storage policy has not already been created, click Create Storage Policy to create a new storage policy,

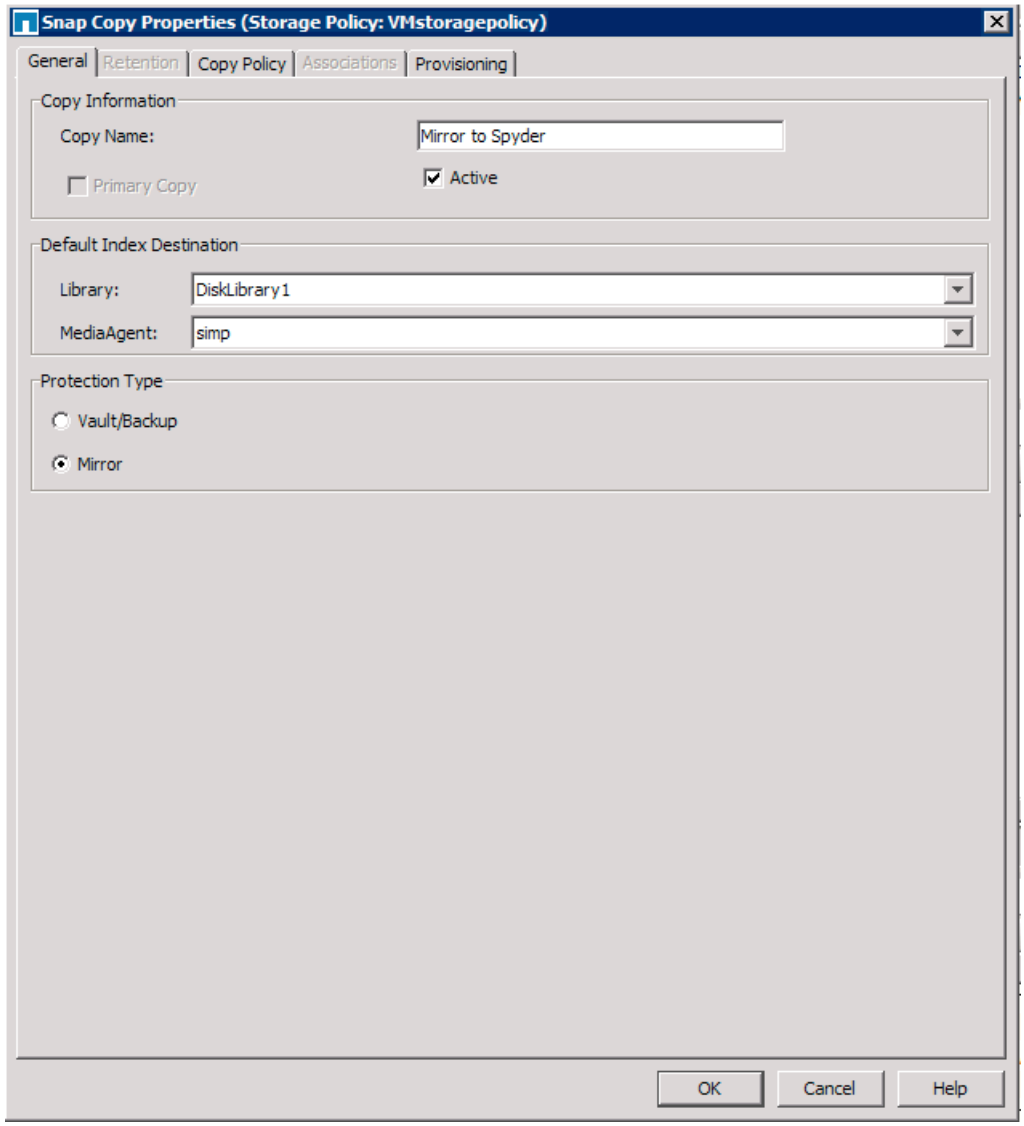
When the subclient properties are populated, click OK. Create backup schedules as needed.

5.7 REPLICATION

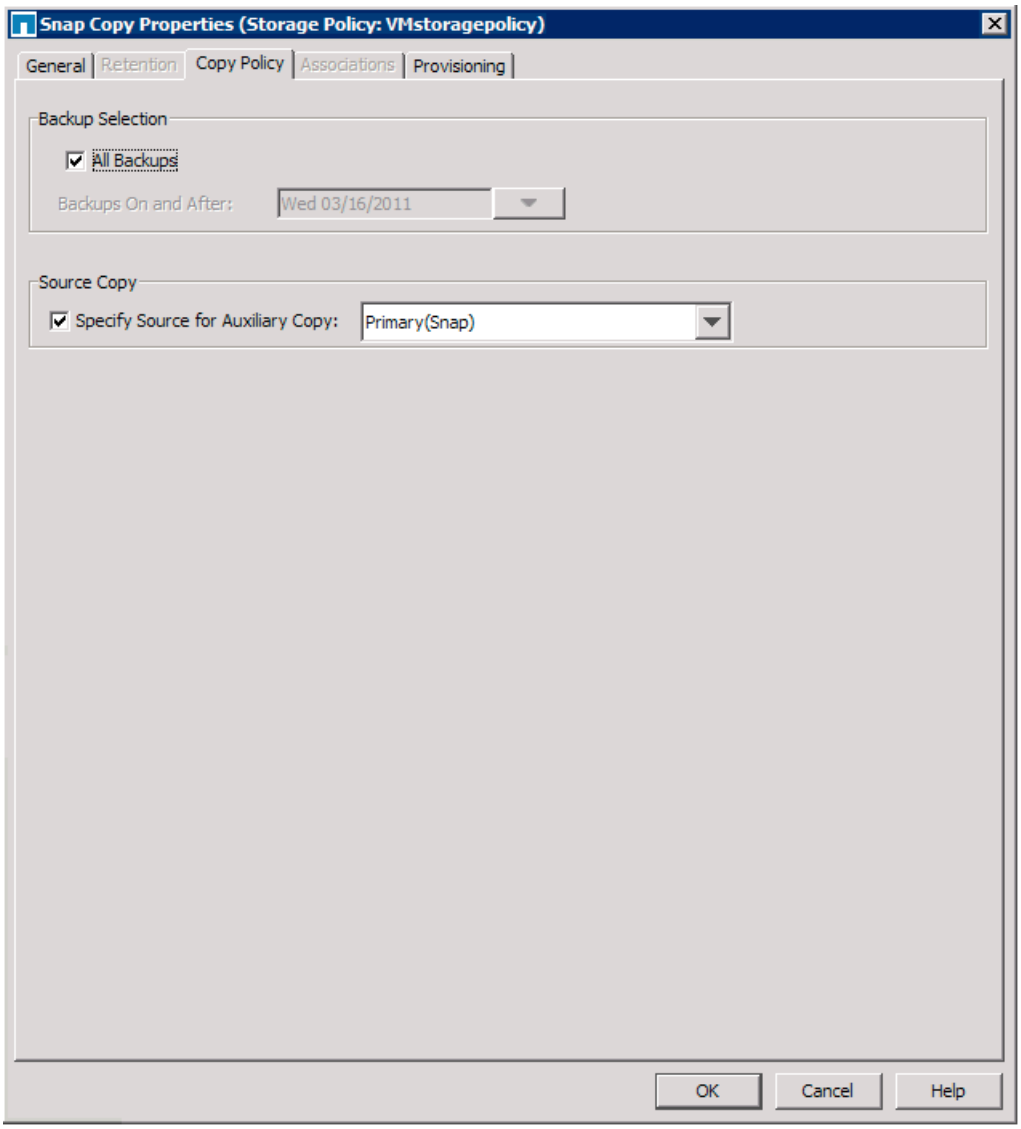
To configure replication, right-click the storage policy → All Tasks → Create New Snapshot Copy.



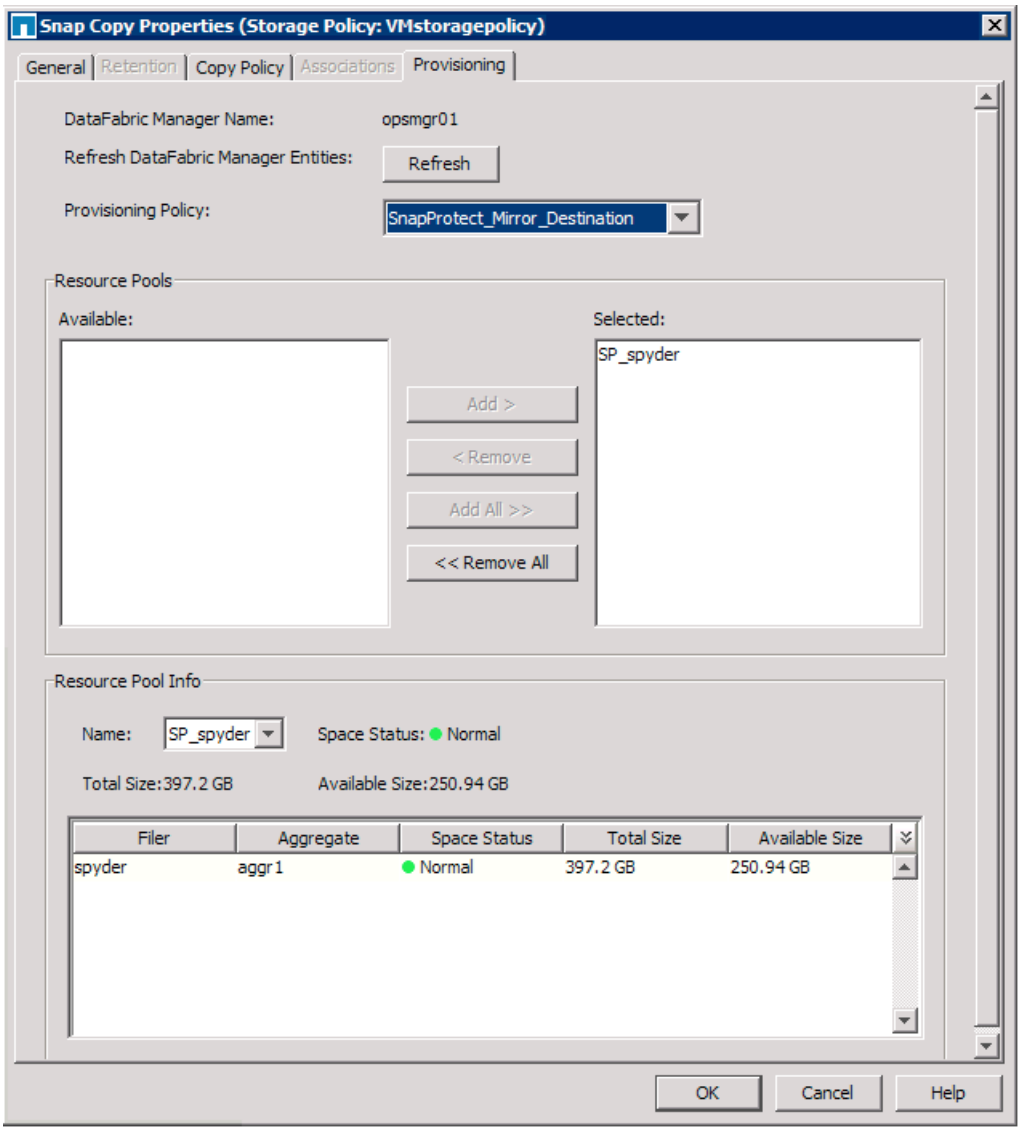
Give the copy a name. Select the library in which to store the indexes, the media agent, and the protection type. In this example, mirroring is used.



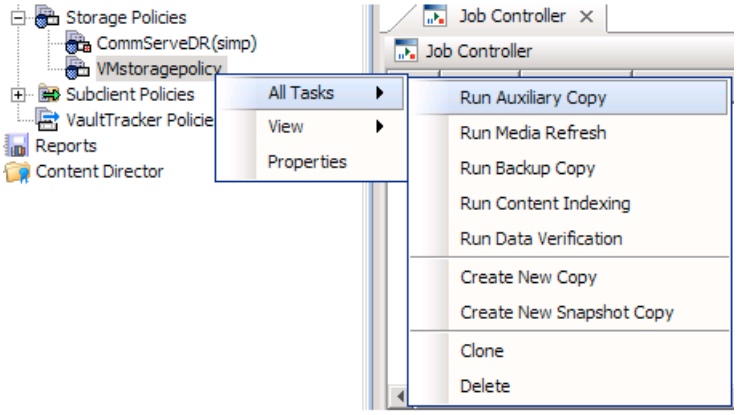
Under the Copy Policy tab, select the source copy for the mirror.



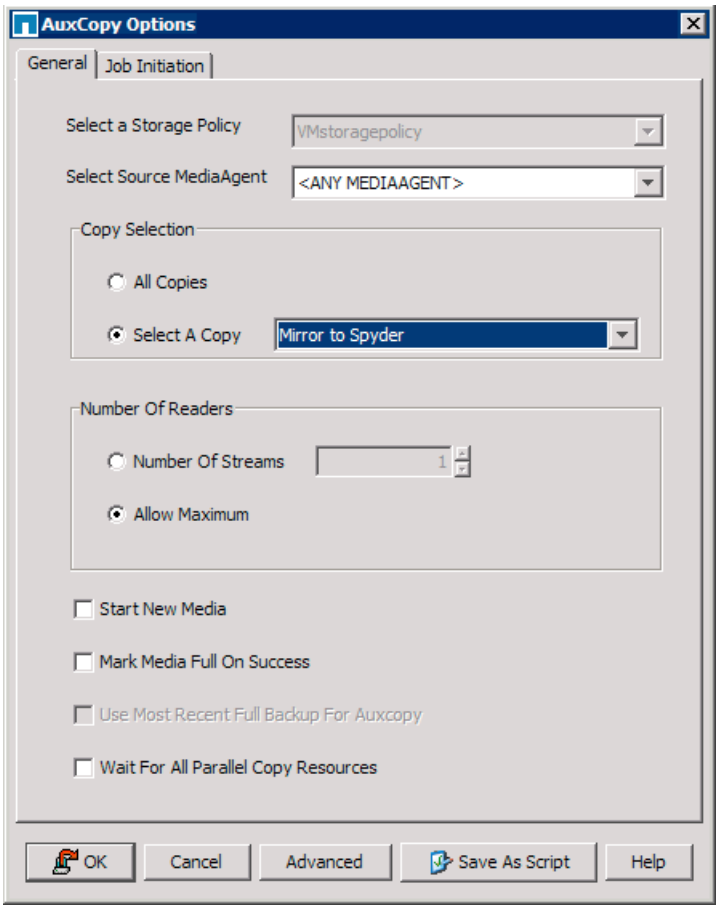
Under the Provisioning tab, select an appropriate provisioning policy. Select the resource pool that was created in section 5.1. Click OK. The new Snapshot copy is added to the storage policy.



To schedule the replication, right-click the storage policy → All Tasks → Run Auxiliary Copy.

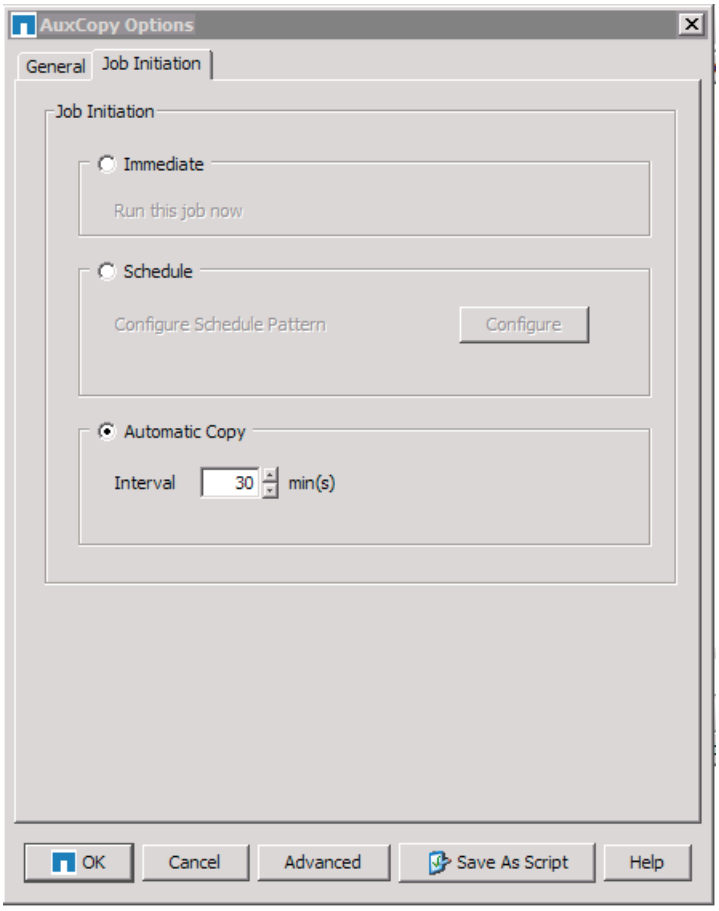


Select the Snapshot copy to schedule.

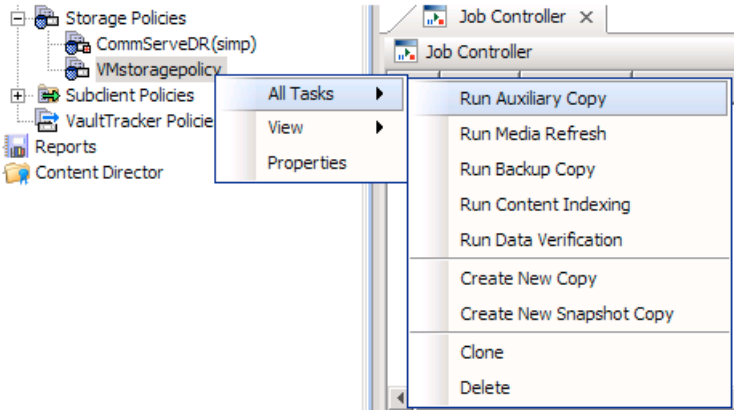


Under the Job Initiation tab, select Automatic Copy and set the interval.

Note: NetApp does not recommend using Automatic Copy with cascading replication copies.



To run a manual mirror job, right click the storage policy → All Tasks → Run Auxiliary Copy.



Select the Snapshot copy and then click OK to start the mirror job immediately.

6 FURTHER READING

Refer to the following documentation for more information.

TR-3487 SnapVault Best Practices Guide
http://www.netapp.com/us/library/technical-reports/tr-3487.html

TR-3446 SnapMirror Async Overview and Best Practices Guide
http://www.netapp.com/us/library/technical-reports/tr-3446.html

TR-3710 Operations Manager, Provisioning Manager, and Protection Manager Best Practices Guide
http://www.netapp.com/us/library/technical-reports/tr-3710.html

TR-3747 Best Practices for File System Alignment in Virtual Environments
http://www.netapp.com/us/library/technical-reports/tr-3747.html

www.netapp.com