Technical Report

Operations Manager, Provisioning Manager, and Protection Manager Best Practices Guide

Adaikkappan Arumugam, Shiva Raja, NetApp

July 2010 | TR-3710

EXECUTIVE SUMMARY

This paper presents best practices to follow when deploying NetApp® Provisioning Manager and Protection Manager software. It also contains recommendations to assist you when configuring resource pools, data sets, provisioning policies, and protection policies.

TABLE OF CONTENTS

1 INTRODUCTION

1.1 INTENDED AUDIENCE

This technical report is designed for storage administrators and architects who are familiar with NetApp storage systems, Data ONTAP® storage provisioning fundamentals, Data ONTAP replication technologies, fundamentals of Operations Manager, Performance Advisor, Provisioning Manager, and Protection Manager.

2 PORTS USAGE

In order for Operations Manager to discover, monitor, alert, and report various aspects of the NetApp storage infrastructure, one should determine that the Operations Manager server is able to communicate with the desired NetApp storage systems. There can be communication issues when a **firewall** sits between NetApp storage systems, the Operations Manager server, and the client (Web-UI and NetApp Management Console). In order to make sure that there aren't any communication issues, the following ports should be enabled and available:

Table 1) Ports used by Operations Manager and targets (part 1)

Port	Description	Importance
The following ports are used by the Operations Manager server to communicate with managed NetApp storage systems, NetApp host agents, and other devices.		
161/udp(SNMP)	Monitoring of appliances	Mandatory
80/tcp(http)	Storage system management (including Performance Advisor, configuration management) and HTTP ping method	Essential
443/tcp(https)	Secure admin-based storage system management	Essential
4092/tcp(http)	Connectivity to NetApp Host Agent (configurable)	Essential
4093/tcp(https)	Secure connectivity to NetApp Host Agent (configurable)	Essential
23/tcp(telnet)	Interactive telnet applet ("Run Telnet" in Operations Manager) to managed appliances	Optional
10000/tcp(ndmp)	Used by Backup Manager to communicate with Open Systems SnapVault® agents (configurable)	Essential
514/tcp (rsh)	Cluster console, storage system takeover/giveback, remote command execution on managed storage systems, vFiler® unit monitoring and management	Essential
162/udp (SNMP)	The DataFabric® Manager server sends SNMP traps to a trap host (configurable)	Optional
25/tcp (SMTP)	SMTP server port (configurable) used when the DataFabric Manager server sends e-mail for alarms	Mandatory
The following ports are used by managed appliances, Operations Manager UI, and NetApp Management Console to communicate with the Operations Manager server.		
8080/tcp(http)	Operations Manager access, software download, and configuration download (configurable)	Mandatory

8443/tcp (https)	Secure Operations Manager access and configuration download (configurable)	Mandatory
8088/tcp (http)	NetApp Management Console access (configurable)	Essential
8488/tcp (https)	Secure NetApp Management Console access (configurable)	Essential
162/udp (SNMP)	Managed appliances send SNMP traps to the DataFabric Manager server to speed up monitoring of important events (configurable)	Optional

Table 2) Ports used by Operations Manager and targets (part 2)

Port	Description	Importance
The following ports are used by the DataFabric Manager server to communicate with managed NetApp storage systems and NetApp host agents.		
514/tcp (rsh)	Cluster console, storage system takeover/giveback, remote command execution on managed storage systems, vFiler unit monitoring and management	Mandatory
22/tcp (ssh)	Secure cluster console, secure storage system takeover/giveback, secure remote command execution on managed storage systems, vFiler unit monitoring and management	Essential

WHY SSH IS BETTER

rsh typically has problems traversing firewalls because it involves connections in both directions. The rsh client will connect to port 514 on the rsh server. The rsh server will then open two connections back to the rsh client to return stdout and stderr; these connections are typically made to dynamic ports on the client, making it difficult for firewall admins. For rsh to work properly, the firewall needs to allow *all* TCP traffic from the managed storage systems destined for the DataFabric Manager server (in addition to allowing rsh traffic from the DataFabric Manager server to reach the managed storage system). Using ssh avoids this problem.

IMPORTANCE

- **Mandatory:** Requires this port to provide basic functionality
- **Essential:** Required only if you use additional functionality
- **Optional:** Less important

3 IMPORTANCE OF AUTOSUPPORT

When Operations Manager is installed, the "autosupportEnabled" global option is initially set to "Unknown." After 24 hours, the option will automatically be set to "Yes" unless you manually set it to "No." However, one should consider the advantages of AutoSupport™ before disabling it.

- Helps NetApp technical support to identify potential problems on your DataFabric Manager server before anyone realizes the problem
- NetApp will help you to solve problems that AutoSupport detects
- If anyone feels that NetApp might get sensitive information about his/her organization, they can configure AutoSupport to not transfer sensitive information
  - Change the "AutoSupport Content" option ("autosupportContent" in CLI) to "minimal" to have the AutoSupport feature blank out all host names, user names, and IP addresses with the "?" character and to not include logs in the AutoSupport transmission
- Enabling AutoSupport is one of the best practices of NetApp Operations Manager

The following table gives you the port information necessary for AutoSupport.

Table 3) Ports used by Operations Manager for autosupport

Port	Description	Importance
The following ports are used by Operations Manager server to enable AutoSupport features.		
8488/tcp (https)	For AutoSupport Protocol	Mandatory
25/tcp (SMTP)	For AutoSupport notification when the "autosupportProtocol" option is set to SMTP	Optional

CHANCES OF AN ANTIVIRUS APPLICATION BLOCKING AUTOSUPPORT TRANSMISSIONS

Many virus-scanning programs block sending TCP packets to another host on port 25. This is to disable mass-mailing worm virus programs. If you have AutoSupport configured to send transmissions using SMTP ("autosupportProtocol" set to "SMTP"), the virus scanner might block the transmission. To work around this, use another protocol such as HTTP or HTTPS for AutoSupport, or configure any virus scanner on the DataFabric Manager server not to block outgoing connections over port 25.

IMPORTANCE

- **Mandatory:** Requires this port to provide basic functionality
- **Essential:** Required only if you use additional functionality
- **Optional:** Less important

4 DATA COLLECTION

This section refers to the frequency of data collection by Operations Manager—“monitoring intervals” from several NetApp storage systems and agents. It is very important to set the appropriate data collection frequency to optimize the processing load on the Operations Manager server.

NetApp recommends leaving all monitoring intervals at their factory defaults.

- Shortening these intervals adversely affects the server performance.
- Lengthening these intervals improves server performance.
  - Lengthening causes detection of status changes and event generation to be delayed, and reports to contain data that is not up to date.

REASON

The performance impact caused by changing the value of “monitoring intervals” is also affected by the number of objects monitored by Operations Manager (such as aggregates, volumes, qtrees, quotas, LUNS, disks, and vFiler units) and the checkpoint frequency of the Operations Manager database.

HOW CHECKPOINT FREQUENCY AFFECTS PERFORMANCE

The database of the Operations Manager server has an internal checkpoint process so that it is accurate, consistent, and keeps up with monitoring intervals. This is necessary for data integrity purposes.

- This checkpoint frequency increases with the increase in changes made in the DataFabric Manager database.
  - The higher the database writes, the greater the checkpoint frequency.
- An increase in checkpoint frequency can further increase the overall burden on your Operations Manager server.

You can refer to “dfmwatchdog.log” under the “log” subdirectory of the Operations Manager installation to see how much memory and CPU each Operations Manager service consumes.

Please refer to the sizing guide TR-3440 to understand the sizing capacity of Operations Manager.

4.1 DATA COLLECTION IN PERFORMANCE ADVISOR

The primary objective here is to get the right performance monitoring without burdening the management server and storage system excessively. In general, the Performance Advisor (PA) tool is designed for minimal overhead. However, you might notice the increase in CPU utilization of storage systems under the following conditions:

- You have selected an option other than the default transport protocol.
  - If SecureAdmin is enabled on your storage system, Performance Advisor requires the HTTPS protocol to monitor your system's performance data.

- You are sampling data at very frequent intervals.
  - Setting the sample rate in performance views of less than 60 seconds
  - Enabling real-time views
    - Use real-time views when really required

- You have selected several different objects and counters—custom views.
  - Data collection in custom views can be redundant because counter data is not shared across views, that is, any data collected previously for a counter in a current performance view is not displayed in a new view.
  - Try to avoid using the same counters across multiple views.
  - If you have Performance Advisor 4.0, create a custom view template instead of a performance view for a particular set of objects.

Increasing the frequency intervals of data collection also helps you to reduce the rate of database file size growth and related disk space usage.

HINT

Always use performance monitoring on systems where required. For example: You can disable the performance monitoring on nonproduction systems.

CONTROL OVER DATA COLLECTION IN PERFORMANCE ADVISOR

With Performance Advisor 4.0 you can control the data collection, sampling rate, and retention time of information collected from the Data ONTAP performance counters, which can help you to minimize the load on targets and efficiently utilize the capacity on the Performance Advisor server.

- For instance, if your storage controllers are only serving NFS protocol, then you can disable the data collection of CIFS-, iSCSI-, and FCP-based counters.
- In order to perform a proper baseline on performance data, allow the PA application to monitor your storage systems for a couple of weeks. This can help you to set the right threshold values.

5 GROUPING

In Operations Manager a *resource group* is a collection of objects that are related in some way. You can group objects based on:

- Characteristics such as storage system (Data ONTAP version)
- Geographical location
- Project or departments or cost centers in your organization
- Business units

Apart from the points mentioned above, grouping in Operations Manager can provide some improved granular control.

- If an admin is responsible for a particular data center, add all the storage systems in the data center to a resource group. Since he/she is interested only in the events of this group, create an alarm with that group.
  - First, set a capacity-related threshold on that independent aggregate/volume/qtree.
  - Now create a subgroup in which you can add that particular object (the corresponding aggregate/volume/qtree).
  - Create an alarm for that particular subgroup. This will enable you to have granularity of raising alarms on single/multiple objects of your choice!
- Groups and subgroups can help you to perform hierarchical configuration management.
  - For example, there is a department "Global Sales" that has 20 NetApp storage systems. They have a subgroup called "APAC Sales" that has 4 storage systems out of 20. Let's assume that all 20 storage systems in Global Sales need to have common values under "options.dns"; however, the httpd option needs to be disabled for all the storage systems under the APAC subgroup: Create the respective group and subgroup and create a configuration template for each group.
  - The subgroups end up having two sets of changes: The Parents configuration (options.dns values) The Groups configuration (option.httpd disabled)
- Make sure that groups created for storage system configuration management contain only storage systems and that MultiStore® configuration management contains only vFiler units.
  - For example: Configuration management for groups containing data sets
- Temporary groups can be created to move NetApp storage systems during maintenance mode in order to group all the events that occur during their maintenance.
  - This can be helpful to isolate the events raised during maintenance mode, which can be of less importance.
- Temporary groups can also be used to run a common operation on all NetApp storage systems.
  - A common example is where "snmp traphost" needs changing on all (or many) appliances. The trick here is to use groups.

Always remember that resource grouping is the key component of Operations Manager. It has an impact on several aspects of Operations Manager such as alerting, reporting, and providing role-based access control. Efficient grouping is always the best practice in Operations Manager.

NOTE:

As a best practice, try to avoid:

- Creating unnecessary groups
- Frequent object addition/deletion/modification from resource groups
- Creating deep hierarchies unless absolutely necessary
- Adding storage systems to all groups

5.1 GROUPING IN PERFORMANCE ADVISOR

Grouping performed in Operations Manager is also carried over to Performance Advisor. You can select a resource group as a source and apply a performance threshold or a performance template. However, you don't need the help of resource groups to achieve maximum granularity. You can drill down a storage system based on its logical construct or physical construct and set the threshold for an individual object, irrespective of its resource group. For example, groups of two volumes of different storage systems (each storage system belongs to a different resource group) can have the same threshold template.

6 REPORTING

The Reporting feature in Operations Manager provides a wide variety of options for report controls that you can use to display NetApp storage infrastructure data in your reports. Operations Manager reports can be rendered into a wide range of formats such as HTML, XML, CSV, Perl, and Microsoft® Office Excel files.

Although reporting in Operations Manager makes it easy and flexible for users to create and manage reports that meet their requirements, users need to create custom reports to meet specific requirements given the complexity of business needs. You may want to consider the following when choosing the best way to design and create a custom report.

- What are the best practices for designing and generating a custom report?
  - Retrieve the minimum amount of data needed in your report. (Avoid unnecessary data in your report by avoiding unnecessary fields.)

  - Try generating reports during off-production hours. (Use Report Scheduling (on a daily, weekly, and monthly basis) and schedule FSRM path walks appropriately.)
- How do I avoid common mistakes when choosing a report layout and picking an output format?
  - Choose the appropriate tab where the custom report is meant to reside. If it is a LUN performance and capacity report, choose the tab LUN.
  - Add the report to the Favorite Reports section. This helps to locate the desired report easily.
  - Choose the appropriate output formats. Operations Manager supports various formats such as Perl, csv, HTML, and Excel.
- How do I take advantage of existing features to generate the customized reports I want when the custom reports in Operations Manager don't help?
  - Custom reports help you to customize reports only within related catalogs. For example, the base catalog "SRM path" can only relate to the base catalog "Agent" (NetApp Host Agent). What if I need to combine the mapped LUN/volume/qtree capacity? I need to combine the base catalog "LUN"/"Volume"/"Qtree," which Operations Manager wouldn't allow. You can also export the Operations Manager database to text files. This will help you to create your own customized reports and you can load these text files to user-specific databases.

NOTE:

Use custom scripts to combine existing canned reports (using Perl, preferably) from different base catalogs to create your own custom reports.

Charge-back reports can be very critical in certain organizations. Operations Manager helps you to introduce a more granular method of calculation:

dfm option set chargebackIncrement=daily

For example, the monthly rate calculated will be independent of the number of days in the month. For example, if the Chargeback Increment is set to Monthly, the rate for both February and March will be same, whereas if it is set to Daily, the two rates will be slightly different.

6.1 REPORTING IN PERFORMANCE ADVISOR

You can generate point-in-time performance reports when a threshold breaches on a set of storage resources.

- For instance, let's assume that you are trying to monitor the latency of your LUNS serving a critical application.

  - Group your storage resources (volumes/LUNS) using Operations Manager resource groups (refer to section 5.1 for grouping and threshold settings).
  - Set the threshold on the latency counter (refer to section 4.1, the last point, for effectively setting performance thresholds) monitoring the group created in the previous step.
  - Associate an alarm with this threshold. For Example: You can associate a script with this alarm that would execute a `dfm` command to generate a performance report on LUNS for the group created in the very first step.

7 ALARM MANAGEMENT

There are several events in Operations Manager for which alarms can be configured. Hence, it is important to decide on a strategy for using alarms in a NetApp storage infrastructure.

- Decide on the appropriate user or user group to create and edit alarms.
- Avoid unnecessary or redundant alarms.
  - Avoid creating alarms for all events.
- Maintain logical groups to organize all events and alarms.

- Set appropriate e-mail recipients in order to send alarm notifications only to the people who need to see or acknowledge them.

Use the following tips to set alerts and notifications:

- Try setting a single alarm for multiple events.
  - Try specifying event classes or use wildcards to specify a set of events. You type in the command `dfm eventtype list` to list the event class.
  - Try specifying alarms by severity level. You can also reduce the severity of events that you feel are unnecessary. For example, not interested in cpu_busy events? Then use: **dfm eventtype modify [-v <event-severity>] <event-name>**
  - This helps you to reduce the number of alarms.
- Monitor similar resources together using the same alert:
  - As mentioned in the grouping section, group critical objects together and configure the desired alarm.
- For resources managed by different teams or individuals, create separate alerts.
  - You can divide objects based on business needs by having different groups and assigning alarms separately to each group.
- For critical resources, select more events to create alarms.
- Associate custom scripts with alarms that can be invoked when the alarm triggers.
  - You can use the Performance Advisor GUI to associate a script with any event.
  - Efficient scripts should be written and shouldn't take much time.

Maintain an appropriate practice of managing alarms when you receive one, such as,

- Make sure that the corresponding user acknowledges the event by taking immediate corrective actions.
  - Monitor unacknowledged but important events using events reports.
- Once the corrective action is taken, make sure that the user deletes the event.
  - The deleted events can be viewed from the history page (deleted events).
  - You can always visit this page to see what problems were resolved on your NetApp storage infrastructure.

7.1 ALARM MANAGEMENT IN PERFORMANCE ADVISOR

In addition to the above-mentioned points, consider the following:

- You can configure a custom event with a combination of thresholds to judge performance issues more accurately.
  - For example, if you set thresholds on the avg_latency counter for all LUNs in a storage system, you might receive alarms from LUNs that are hardly used. Hence, you can create a rule to generate alerts only when the total_ops counter and the avg_latency counter breach the threshold values.
- You can configure alarms for these events in Performance Advisor or Operations Manager.
- Since these are custom events, you can use a naming convention containing notifications or mini comments.
  - If these comments are further documented, it helps users to respond to the alarm and troubleshoot.

Alarm creation can be refined over a period of time involving a considerable amount of iteration. During this time an administrator can watch out for the following opportunities to refine alarm creation:

- Determine whether trigger values are set appropriately.
  - For example, determine that the trigger values are not set too low, causing the alert to trigger frequently. This can only be done after you baseline the performance metrics of your storage infrastructure.
- Determine whether any event is spiking (frequently exceeding a trigger value and then returning to normal levels).
  - If so, adjust the trigger values or the values that prevent the event from triggering until it exceeds a trigger value for several consecutive intervals.

- Observe how quickly users respond to an alarm. Determine whether the alert schedule is appropriate considering the user response.
  - For example, if users typically respond to an alert after several days, then a schedule that evaluates the alert every hour may not be necessary.

8 USER SECURITY

Operations Manager provides user security to both Operations Manager users and Data ONTAP users. Hence, one must keep the following points in mind while designing proper user security for Operations Manager users and Data ONTAP users.

8.1 OPERATIONS MANAGER AND PERFORMANCE ADVISOR USERS

Access control enforcement on storage administrators who configure the storage environment is a critical part of providing proper user security. To provide proper user security, the following points should be enforced:

- Each Operations Manager user must have an individual account; there shouldn't be any shared or common user accounts.
  - Only people who need to use Operations Manager should be made Operations Manager users.
- Strict password policies should be enforced:
  - Passwords should be complex.
  - Passwords should be changed regularly.
  - Passwords should be closely held secrets.
- Each Operations Manager user should be authorized to perform only the management actions required to perform their job.
  - Create user groups corresponding to the specified roles in an organization so that users may be placed in those groups that correspond to their job/duties. **There are no user groups in OM. You can have Active Directory® groups added as users to OM.** Limit the number of users in the administrative group.
  - Be frugal in assigning "performance monitoring" access rights. Be careful in assigning these rights because this action might affect storage system CPU utilization, especially **"PerfView RealTimeRead."** Make sure that only the appropriate users possess this ability, because it enables viewing real-time data.
- Ensure that users are given authority only over objects they must use to do their duties.
  - Use resource groups in which you can add the required objects (storage systems, hosts).
  - Use subgroups under these logical groups to introduce more granularities.
- Create only as many users and user groups as required.
  - Delete Operations Manager users / user groups that are not being used.
- Be cautious when changing the membership of the users in user groups. Care should be taken to not move users to user groups that do not relate to their job description.
- Operations Manager users' actions should be regularly audited.
  - Check the audit logs of NetApp storage systems to see a history of operations performed.

Note: An upgrade from an older release of Operations Manager would make the "everyone user" have read access to everything by default.

8.2 SECURING STORAGE INFRASTRUCTURE ELEMENTS

Access control enforcement on storage administrators who work with NetApp storage systems is as critical as providing proper user security to Operations Manager users. To provide proper user security over the NetApp storage infrastructure, the following points should be enforced:

- Maintain strict access to the configuration management capabilities of NetApp storage systems and MultiStore.
  - Users should only have access to configuration management if they are authorized to change configurations.
  - Users should only have access to storage systems and MultiStore if they need them to do their jobs.

- The passwords of root user IDs for NetApp storage systems and MultiStore should be held secretly because they are required by Operations Manager for monitoring purposes.
- Ensure the usage of SNMPv3 protocol for monitoring and reporting on the NetApp storage systems with Data ONTAP 7.3 and above.
  - SNMP v1 is not a secured protocol.
- Always remember that the Operations Manager administrator controls the capabilities of Data ONTAP users.
  - Create a user ID and provide the required access for a Data ONTAP user in Operations Manager if needed.
  - Care should be taken to provide the user access to perform only the management actions required to perform his/her job.
- Remember to secure your NetApp storage system according to its capabilities and best practices.
  - Refer to the Data ONTAP administrative guide for more information.

9 REPOSITORY MAINTANANCE

The Operations Manager repository is a relational database that holds the current and historical data of both the NetApp storage environment and Operations Manager itself. This data includes configuration details about storage systems, hosts (NetApp host agents), statistical data for capacity planning, alerts, and status information about any given device.

- Make sure that the database is backed up regularly.
  - Try creating a strategy for the database backup based on the retention count that can be set under options. (7 daily database backups for the current week, 4 weekly database backups for the current month and 12 monthly database backups for the current year)
- The Operations Manager database can operate from a NetApp LUN if the local disk doesn't have enough space or for better performance and high-availability reasons.
  - To manage a huge database optimally, the database itself can be moved to a NetApp FCP LUN (whose volume is striped with four to five high-performance disks; refer to Data ONTAP best practices for more details).
  - To reduce the I/O load, more memory can be assigned to the database cache: dfm database set dbCacheSize=1024**.** This is only an optional setting
  - **With the help of SnapDrive®, you can take application-consistent snapshots of your Operations Manager repository.**
  - Moving the Operations Manager database to a LUN can also help you to perform migration from one server to another. You can also use SnapMirror® and SnapDrive to migrate it to a different storage system's LUN.
- Monitor and maintain the size of the database.
  - Use the Purge Events Interval option to purge old and unnecessary events. Use dfm options set EventsPurgeInterval = <Time_duration>
  - If FSRM is being used, monitor the database size carefully because it can be responsible for increasing the size of the database exponentially.
- You can configure the Operations Manager database and services with high availability and disaster recovery solutions.
  - Please refer to https://now.netapp.com/Knowledgebase/solutionarea.asp?id=kb41917 for Operations Manager configured with an HA solution.
  - Please refer to TR-3655 for Operations Manager configured with a DR solution: http://media.netapp.com/documents/tr-3655.pdf
- Make sure that the Operations Manager database and DataFabric Manager options are backed up prior to an upgrade operation.

10 PROVISIONING MANAGER AND PROTECTION MANAGER INTEGRATION

Though Provisioning Manager and Protection Manager are independent applications (Protection Manager 4.0 automatically enables Provisioning Manager, but vice versa isn't true), the integration between these products brings out a complete picture of provisioning and protecting NetApp storage that involves creating automated workflows for storage provisioning, data protection, and data migration.

In further sections, you learn about a significant number of components and concepts that are common to both Provisioning Manager and Protection Manager. The objective of this document is not only to highlight the best practices of Provisioning Manager and Protection Manager, but to highlight integrating Provisioning Manager and Protection Manager to achieve operational efficiency for storage provisioning and data protection of NetApp storage resources.

10.1 PROVISION AND PROTECT AT THE SAME TIME

When a data set is associated with a provisioning policy and a protection policy, you don't have to go through separate workflows for provisioning and protecting storage. Through a single click you have the capability to provision and protect your storage at the same time.

Figure 1 shows the example of a data set associated with primary and secondary resource pools through provisioning and a protection policy.

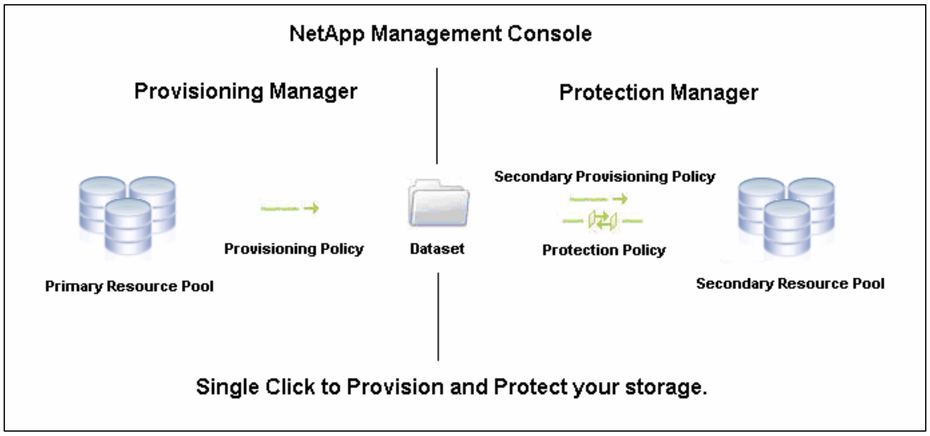


Figure 1) Example of data set associated with primary and secondary resource pools through a provisioning and protection policy.

10.2 INTEGRATION THROUGH PROVISIONING POLICIES

Provisioning Manager's provisioning polices can also be applied to secondary storage in two ways:

- Provisioning Manager can create a dedicated "Secondary Provisioning Policy." However, unlike other provisioning policies (NAS/SAN) this policy does not have the capability to specify thin provisioning settings.
  - Dedupe settings specified in this policy can only be applied to data sets with backup- and DR–based protection policies.
- Provisioning Manager's NAS- and SAN-based provisioning policies can be applied to the secondary storage of a data set if it is associated with a DR-based protection policy.

10.3 INTEGRATION IN DATA MIGRATION

Data migration in Provisioning Manager can be performed in two ways:

- Data set migration (follows the underlying principle of vFiler unit migration)
- vFiler unit migration (useful when you have multiple data sets associated with the same vFiler unit)

It is interesting to observe that data set/vFiler unit migration in Provisioning Manager preserves the replication relationships managed through Protection Manager. Hence, you don't have the trouble of rebaselining when data with a replication relationship is migrated.

10.4 STORAGE SERVICE CATALOG

Instead of associating resource pools and provisioning and protection polices separately, Provisioning Manager 4.0 introduces the “Storage Service Catalog,” which is a set of policies, resource pools, and vFiler unit templates that define a storage service level, usually corresponding to a service-level objective. A data set may be provisioned by specifying a desired storage service. The resource pool used to provision the storage and the policies (provisioning and protection) governing the data set are determined by the storage service specified.

A Storage Service Catalog contains the following:

- Provisioning policy
- Protection policy
- Primary and secondary (tertiary) resource pools
- vFiler unit templates (used to specify common DNS/NIS settings for a vFiler unit associated with a data set)

The advantage of the Storage Service Catalog is that it frees storage administrators to create data sets and associate the appropriate policies and resource pools. All a storage administrator needs to do is to create a set of service catalogs. Once these service catalogs are created, end users (server or database) administrators can create their own data sets and associate the appropriate Storage Service Catalog to provision and protect their storage within a single workflow.

ADDING A STORAGE SERVICE CATALOG

The most important point that needs to be considered for this step is to ensure that you have the appropriate policies, resource pools, and vFiler unit templates created well in advance.

10.5 DATA SETS

A data set is a collection of data (logical containers such as volumes, qtrees, LUNS) to be managed as a unit in provisioning, mounting, protecting, and recovering. The physical resources of a data set should share the same storage provisioning and data protection requirements. For example, a data set might consist of all the home directories that need to be thinly provisioned, deduped, and backed up three times a day. Another might be all of your tier-1 application data that needs to be replicated to secondary storage for D2D backup and then mirrored for DR.

BEST PRACTICES TO CREATE A DATA SET

- NetApp does not recommend having only one volume per data set since it would break the very purpose of creating a data set. Group the primary data that has identical provisioning and / or protection requirements.

  - Fan-in capability in Protection Manager only applies to a data set level.
- A data set name should be meaningful. For example, it could be named with the application for which the storage is being provisioned.
- A volume can be imported in a data set if the administrator wants the volume to conform with a particular provisioning policy.
  - NetApp does not recommend importing the same volumes to multiple data sets with different provisioning policies.
- When you think about how you want to group your data, think about the most logical way to segregate it and how you want to protect it. Also think about how you want to assign resources to contain backups or mirrors.
- When creating data sets in a large environment it is a good idea to map your data environment on paper or in a spreadsheet and assign each component to a data set prior to creating it in the Provisioning/Protection Manager application.

NOTE:

- A data set can have multiple resource pools.
- For Provisioning Manager 3.8 or earlier, when a user who is trying to create a data set decides to provision the storage with the creation wizard and for some reason the provisioning fails, then a data set

is not created. NetApp recommends choosing the “Provision later” option and provisioning the storage for the data set under the Provision tab.

- For Protection Manager 3.8 or earlier, a single storage system that is licensed with both SnapVault primary and SnapVault secondary licenses cannot be included in a data set.
- Once data sets are configured and storage is provisioned, NetApp does not advise changing the policy type. For example, if a data set is created with NAS storage, then later changing the policy to SAN is not a good practice.
- If you want to have a data set (associated to a vFiler unit) with CIFS shares to be migrated, create a vFiler unit manually (instead of automated vFiler unit creation through the data set creation wizard) and associate it with a data set (in the data set creation wizard). Only offline migration is allowed.

GUIDELINES FOR USING TIME ZONES WITH DATA SETS

When creating or editing a data set, you can select a time zone to associate with the data set. Review these guidelines to help you determine whether to use the default time zone or select a new one for your data set.

- You can select a time zone other than the default for any data set, but you do not have to.
- If you do not select a time zone, the default value of the DataFabric Manager server or the default value set with the `dfm option set timezone=<timezone` is used.
- If you associate a time zone with a data set, the protection application uses the time zone you selected, rather than the server's system default, to interpret schedules affecting data in that data set.
- When you set a time zone for a data set, the time zone applies only to the primary data in the data set. It does not impact any destination nodes or resource pools.
- The start and end time of a local or a remote replication is determined by the time zone selected for the primary data in the data set. The time zone selected for any resource pool associated with the data set **does not** impact the time of the replication event starting on the primary.

There are times when you might want to change the default time zone:

- If most of the primary data is in a different time zone than the DataFabric Manager server, you might want to set the DataFabric Manager global command, `dfm option set timezone=<timezone>`, to specify which time zone Protection Manager should use for interpreting all schedules.
- If most of your data sets are distributed across different time zones and you want the schedules interpreted in the "local time" for each data set, specify each data set's time zone in the Provisioning Manager or Protection Manager user interface.

10.6 RESOURCE POOLS

Resource pools describe physical resources such as aggregates or storage systems that have the same attributes, such as size, cost, performance, and availability. These attributes are used by the storage admin when matching physical resources with provisioning or when balancing data set operations.

BEST PRACTICES TO CREATE A RESOURCE POOL

- Prior to creating resource pools, consider how you will use the available storage. This will help you to determine how you want to combine those resources into resource pools to meet your setup's replication needs.
- NetApp does not recommend having primary and secondary storage systems within the same resource pool. Create separate resource pools for primary and secondary storage.
- Provide a meaningful name that briefly describes the storage or the intended use of the storage in the resource pool: for example, server1_homedirs, tier1_mktg, or china_eng.

Following are examples of ways that you might combine your physical resources to make the most efficient use of your resource pools:

- A set of aggregates composed of inexpensive, slow ATA drives
- A set of high-performance aggregates composed of 15K Fibre Channel disk drives in a RAID-DP® configuration
- A set of resources categorized based on cost or performance
- A set of storage systems that are suitable for provisioning for certain departments within an organization

- A set of homogenous resources grouped together as a way of restricting access to high-performance storage

NOTE:

- When a resource pool contains a storage system, all aggregates and disks on the storage system are available for provisioning.
- Resource pools may not nest and may not overlap (that is, storage in one resource pool may not be in another resource pool).
- A resource pool can be shared by multiple data sets.
- Resource pools appear in DataFabric Manager resource groups and therefore cannot contain resource groups.
- If more than one OM server is deployed, make sure to not include any particular aggregate or a storage controller in a resource pool on one server and an aggregate from that storage controller in a resource pool on the other DFM server.
- The default values for overcommitment thresholds available in Provisioning Manager are a best practice. These will be used for thin-provisioning scenarios. If the user does not want a high degree of thin provisioning, then these thresholds can be set to a lower value.
- If you intend to associate a resource pool with a data set in a mirror relationship, the volumes on the primary node and those on the secondary nodes must be of the same type. You cannot combine traditional volumes and FlexVol® volumes on nodes that are part of a mirror relationship.

GUIDELINES FOR USING TIME ZONES WITH DATA SETS

When creating or editing a resource pool, you can select a time zone to associate with the resource pool. Review these guidelines to help you determine whether to use the default time zone or select a new one for your resource pool.

- You can select a time zone other than the default for any resource pool, but you do not have to.
- If you do not select a time zone, the default value of the DataFabric Manager server or the default value set with the `dfm option set timezone=<timezone` is used.
- When you assign a time zone to a resource pool, any protection policy node associated with that resource pool will use the resource pool's time zone.
- When you set a time zone for a resource pool in the protection application interface, the time zone applies only to the data on the secondary destination node. It does not impact the primary data in a data set. Different time zones can be selected for replication of primary data to a secondary resource and for replication of data from a secondary to a tertiary resource.
- For two-node policies, the protection application uses the time zone of the primary data to interpret schedules for data transfers. This is true even if a resource pool is associated with the data set and that resource pool has a time zone assigned to it that is different from the data set's time zone.

There are times when you might want to change the default time zone:

- You might want to change the default time zone for a resource pool when working with a three-node policy. If you want data transfers from the second node to the third node to be scheduled in the "local time" of the second node, specify a time zone for the resource pool that contains the second node.
- If you change the default time zone by using the `dfm option set timezone=<timezone>` command, then all resource pools that use that default setting begin using the new default value.
- Be sure any changes to the time zone are thoroughly evaluated for potential impact to schedules, because changes could disrupt the schedules for future jobs.
- Data sets and resource pools for which the time zone is set by using the protection application user interface are not affected by changes made using the time zone option. This is because the Protection Manager setting overrides any default setting.

RESOURCE LABELS

You can assign a resource label as a way to narrow the available resources to be considered for a provisioning or protection request. This feature is only available when Provisioning Manager is licensed.

When a provisioning request is processed, do you want to restrict the resources available for provisioning to only those with a specific label assigned to them?

- A label set on an individual member of a resource pool takes priority over a label applied to the entire resource pool.
- Labels can be edited in line in the table. For both resource pool and members, an existing label can be selected from the drop-down list or a new label can be typed in.

The resource label can be assigned when you create a resource pool. This is an optional custom comment that you can assign to a resource pool or to the individual members of the resource pool, such as storage systems or aggregates. You might assign a resource label based on factors such as cost, reliability, or specific configurations. The resource label essentially functions as a filter. It allows you to identify specific resources to be used to fulfill a provisioning request, so only those resources that have the label assigned to them are considered. This allows granular control when matching provisioning requests with available resources.

When you create a provisioning policy you can specify a resource label to be associated with the policy. If a label is specified for a policy, only the resource pools and resource pool members that match the label are used to fulfill the provisioning request; otherwise it errors out. Storage that has an assigned resource label can still be used to fulfill provisioning requests that do not specify a label.

For example, assume an administrator assigns a resource label of Tier 1 to a resource pool containing the highest-cost, most reliable storage. The administrator also creates a provisioning policy named prov-pol-1, with the resource label Tier 1 specified. When a provisioning request is made with the prov-pol-1 policy, Provisioning Manager searches for storage with the Tier 1 label. If no resources with that resource label are available, the provisioning request fails. For this reason, you should use resource labels with care in the provisioning policy.

10.7 PROVISIONING POLICIES

Provisioning policies define the parameters of how primary and secondary storage should be provisioned. If you have Protection Manager licensed (provisioning is autoenabled starting with 4.0, so you need to check), you can also apply a secondary provisioning policy to the secondary storage through Provisioning Manager. The secondary provisioning policy is only dedicated to secondary storage. You aren’t allowed to associate other provisioning polices (created for primary storage) to secondary storage. However, there's an exception. If the disaster recovery license (refer to the protection policy section of this report) is enabled, you can also apply primary provisioning policies to secondary storage for DR policy-based data sets.

BEST PRACTICES TO CREATE A PROVISIONING POLICY

- It is a best practice to have RAID-DP selected for provisioning your storage containers (provided that your resource pools have RAID-DP aggregates).
- Please refer http://media.netapp.com/documents/tr-3505.pdf for best practices on Deduplication.
- While provisioning NFS export, the option “Guarantee initial size, allocate maximum size on demand, and allow automatic deletion of Snapshot copies” under “container settings” is optimal for space utilization. This option is not available for CIFS in Data ONTAP 7.3.1.
- While provisioning LUNS, the option “Guarantee space for LUN/volume, grow space for Snapshot copies on demand, and allow automatic deletion of Snapshot copies when necessary” under “container settings” is considered to be the best option.
  - The Snapshot™ copies created by Protection Manager will not be affected in this option (it might override if the retention time is too long). These Snapshot copies will be deleted automatically through the protection policy’s retention expiration or they can be manually deleted through the secondary space management feature in Protection Manager 4.0.
- The default values for capacity thresholds available in Provisioning Manager are a best practice. If the user wants to increase the restriction with respect to capacity, then these thresholds can be set to a lower value.
- If there are several data sets with the same provisioning requirements it is ideal to have a single provisioning policy.

NOTE:

- The deduplication schedules, once specified, are not maintained by Provisioning Manager. The schedules are actually pushed to the appropriate storage controller. If Provisioning Manager fails, the dedupe process is still carried out by the appropriate storage system.
- If a data set has a NAS provisioning policy, the exports are done only at qtrees. By default the max number of qtrees per volume is 25. If required, you can change this option: `dfpm dataset set <dataset_name_or_id> maxqtreespervolume=x.`
- If a data set has a SAN provisioning policy, the default number of LUNs per volume is 25. If required, you can change this option: `dfpm dataset set <dataset_name_or_id> maxlunspervolume=x`**.**

10.8 PROTECTION POLICIES

Protection policies define how the members of a data set should be protected. If you have the Disaster Recovery license, protection policies can also define how to fail over to DR secondary storage on the disaster recovery node when disaster occurs.

When a protection policy is applied to a data set, it defines how data stored in a data set's members should be backed up or mirrored. You can configure a protection policy that specifies a single protection method (local backup, remote backup, or mirroring) or a combination of those methods. For example, a protection policy might specify that the primary data is backed up to a secondary location and that the secondary copies are mirrored to a tertiary location.

If the Disaster Recovery license is installed, protection policies that use qtree SnapMirror to back up data can also invoke your site's disaster recovery script. After the problem is resolved, you can move data set member access manually from the secondary storage back to the primary storage.

On a broader perspective, there are two things involved in creating a protection policy:

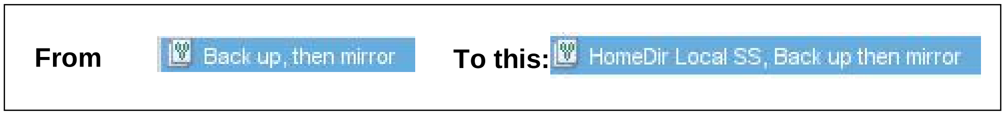
- The Backup/Data transfer schedule on:
  - Primary Node (local backups)
  - Primary Node and Secondary Node (local Snapshot copies with SnapVault/SnapMirror)
  - Primary Node, Secondary Node, and Tertiary Node (local Snapshot copies with cascaded SnapVault/SnapMirror)
  - Secondary Node (remote backups such as OSSV backup)
- Retention time of Snapshot copies
  - Primary Node (local backups)
  - Secondary/Tertiary Node (SnapVault target)

BEST PRACTICES TO ADD A PROTECTION POLICY

- To specify a solid backup/data transfer schedule, you can specify timings of hourly, daily, weekly, and monthly backups within a single schedule. A common practice is to:
  - Apply a schedule of **frequent hourly, one daily, one weekly, and one monthly** local backup and remote backup operation on the primary node and on the backup connection between the primary node and the secondary/tertiary backup node.
- Please refer:
  - http://media.netapp.com/documents/tr-3487.pdf for best practices on SnapVault.
  - http://media.netapp.com/documents/tr-3446.pdf for best practices on SnapMirror.
  - http://media.netapp.com/documents/tr-3466.pdf for best practices on OSSV.
- When specifying a data transfer schedule, a user needs to have a good idea about the backup/data transfer window. This will allow him/her to group the primary data into a data set more ideally.
  - For example: Let's assume that a data set has 10 volumes and it is associated with a protection policy named "Mirror, then backup." Now let's assume that the 10th volume takes a longer time to complete its SnapVault transfer from secondary to tertiary compared to other volumes; then the SnapMirror schedule between the primary and secondary (for all the 10 volumes) will not trigger until the 10th volume has completed its backup to the tertiary node. Hence, care should be taken in

grouping the primary volumes to a data set and specifying the data transfer schedule to a protection policy.

- Before selecting a schedule for a replication job, you need to know the time zones associated with the primary data and destination storage for that job. That information helps you determine when you want your local backups to occur and when you want the remote replication to take place to achieve your data protection goals.
  - When Protection Manager interprets a scheduled replication from primary data in a data set to secondary destination storage, Protection Manager uses the time zone of the primary data. When interpreting a scheduled replication from secondary storage to tertiary storage, Protection Manager uses the time zone of the secondary node.
  - If you misapply time zones, unexpected and unwanted results could occur. For example, a weekly mirror to tertiary storage might occur before completion of the daily backup that you want to capture, or replication jobs might occur at a time of day when network bandwidth is already servicing a heavy load, and so forth.
  - Some time zones observe daylight saving time (DST) and some do not. Protection Manager automatically observes the local time zone and adjusts for daylight saving when appropriate. As a result, backups scheduled during the DST transition, between 1:00 a.m. and 3:00 a.m. local time, may not perform as intended.
- In a backup policy (SnapVault), you can specify a shorter “retention time” on the primary and longer “retention time” on the secondary/tertiary node for effectively utilizing the capacity on the primary node.
  - Common retention durations for hourly backups are for short durations of a day or two if you are also maintaining daily, weekly, or monthly backups of this data for longer durations.
  - Common retention durations for daily backups are from five days to several weeks.
  - Common retention durations for weekly backups are from one month to several months.
  - Common retention durations for monthly backups are from five months to several months.
- The lag in a protection policy refers to the RPO. Always determine that the lag time (error/warning) is greater than the Snapshot/SnapMirror/SnapVault monitoring interval.
  - For example: If the monitoring interval is set to 1 hour and Lag Warning and Lag Error are set to 15 and 30 minutes, you might never receive that lag warning error message.
  - As a best practice, make a copy of the protection policy you want to use and rename it something meaningful to you or to the backup administrator(s).
- It is a best practice to rename the node(s) of the policy you want to use for a particular job with a name that is meaningful to you or your organization instead of using the default of primary, backup, and mirror, as shown below:



From Back up, then mirror **To this:** HomeDir Local SS, Back up then mirror

Figure 2) Policy naming

- If there are several data sets with the same protection requirements it is ideal to have a single protection policy.

NOTE:

- In data sets configured for disaster recovery backup protection, SnapMirror licenses on the primary and disaster recovery node systems are required to support the backup operation. The protection application will configure underlying qtree SnapMirror relationships that support backup and failover processes between the primary and disaster recovery nodes.
- If you are configuring disaster recovery protection, you have the option to assign an export protocol to the disaster recovery node so that, in case of failover, users can access data in the disaster recovery node using the same protocols they used to access data in the original primary node.

- To enable preservation, Protection Manager always retains copies of at least the last two successful backups even if those copies are older than their specified minimum retention duration of the most recent successful backups.

11 VFILER UNIT MANAGEMENT AND MIGRATION

This option is only available in Provisioning Manager. This section deals with providing guidelines for vFiler unit management and migration.

- You can create and set up (even specify a VLAN ID, available only in Provisioning Manager 3.8 or later) through a resource pool only when you have the entire storage controller assigned to it.
- It is a best practice to have a vFiler unit template with common settings and to use that template to set up multiple vFiler units with the same setting.
- You also have the capability to migrate vFiler units and data sets (associated with a vFiler unit) in Provisioning Manager 3.8 or later.
  - Provisioning Manager 3.8 only allows offline migration of vFiler units/data sets (with Data ONTAP 7.3.1 or later).
  - Provisioning Manager 4.0 allows nondisruptive migration of vFiler units/data sets (with Data ONTAP 7.3.3).
- Data sets/vFiler units can be migrated within or across the members of a resource pool, provided these members contain the individual storage systems and not just aggregates of the system.
- With nondisruptive migration of vFiler units/data sets (NetApp Data Motion™, Data ONTAP 7.3.3) in Provisioning Manager 4.0, the following points should be taken into account:
  - NetApp Data Motion is not possible between a CFO pair because synchronous SnapMirror is not supported within a CFO pair.
  - NetApp Data Motion is not allowed if the source or destination is taken over by its partner in a CFO pair.
  - NetApp Data Motion cannot happen from a high-end platform to a low-end platform.
  - For example, it is not possible to migrate from FAS6070 to FAS3020.
  - During cutover, if the source or destination storage system crashes or if the DFM server crashes, you need to use the command `dfpm migrate fix` once the storage systems/DFM server are up.
- Although NetApp Data Motion is nondisruptive, NetApp advises performing migration during off-production hours. This is because NetApp Data Motion uses synchronous SnapMirror to perform a cutover, and the applications being served from this vFiler unit might experience performance degradation. As a matter of fact, Provisioning Manager performs some checks so that the source and destination storage systems are not overloaded:
  - Provisioning Manager calculates the current CPU/disk and anticipates additional load because of the pending cutover process.
  - The CPU threshold (on the source and destination storage systems) is 60% and the disk threshold (on the destination aggregates) is 50%. The cutover will error out if either one of the two samples collected by Provisioning Manager breaches the threshold limit.
- You can also use Provisioning Manager's CLIs or APIs to migrate a vFiler unit to a desired aggregate on the target controller.
- You have the capability to roll back after the cutover process in Provisioning Manager 4.0. This is a quick process; however, the rollback process will not be rapid if the vFiler unit has FlexClone® volumes.
- It is not possible to migrate (offline/online) a vFiler unit if it serves destination volumes for SnapMirror/SnapVault.

12 SNAPMIRROR, SNAPVAULT, OSSV, AND DEDUPLICATION

This section deals with Protection Manager's inner workings with SnapMirror, qtree SnapMirror, SnapVault, OSSV, and deduplication in Data ONTAP.

WHEN DOES PROTECTION MANAGER USE QTREE SNAPMIRROR?

Protection Manager can use SnapVault or qtree SnapMirror to perform backups. The protection application determines which technology to use based on the licenses enabled on the source and destination storage system and the schedule applied to the backup connection of the protection policy.

If you use only SnapVault licenses in your environment, the protection application uses SnapVault for backups. However, if you use both SnapVault and SnapMirror licenses in your environment, the protection application uses the following to determine whether to use SnapVault or qtree SnapMirror for backups:

- If both the source and destination storage system have the SnapMirror license enabled and not licensed for SnapVault, the protection application uses qtree SnapMirror for backups.
- If both the source and destination host have the SnapVault license enabled but not the SnapMirror license, the protection application uses SnapVault for backups.
- If the data to be backed up is located on a host running the Open Systems SnapVault agent, the protection application uses SnapVault for backups.
- If the schedule applied to the backup connection specifies that the data needs to be backed up more frequently than once an hour, the protection application uses qtree SnapMirror for backups even if all SnapVault licenses are enabled on the source and destination storage system.
- If the secondary provisioning policy is used for the backup node and it has scheduled dedupe enabled, , the protection application uses qtree SnapMirror for backups even if all SnapVault licenses are enabled on the source and destination storage system.
- Customers may have SnapMirror and SnapVault licenses but still want NetApp Protection Manager to create SnapVault relationships instead of qtree SnapMirror licenses. For this to occur, change the option `pmQSMBackupPreferred` to “no.”

OTHER NOTES

- Set the options SnapMirror.access and SnapVault.access to “all” on destination storage systems if they do not have access to the source storage systems.
- By default, Protection Manager estimates each Open Systems SnapVault secondary volume to be 10 gigabytes. You may change the initial projected size using the following CLI command:
  ```
  dfm option set pmOSSVDirSecondaryVolSizeMB=51200
  ```
  - The projected size is expressed in megabytes. In this case, we set it to 51,200 megabytes (50 gigabytes).
- Protection Manager supports the capability of backing up (SnapVault/qtree SnapMirror) multiple-source volumes to a single destination volume. The default number of primary volumes is “one”. To configure Protection Manager so that it is capable of backing up (SnapVault/qtree SnapMirror) multiple-source volumes to a single destination volume, use the following option: `dpMaxFanInRatio`. This is applicable only starting with DFM 3.8. The tested value for this option is “four”.
- Protection Manager can recognize the Snapshot copy on deduped SnapVault secondary volumes and automatically initiate the dedupe job on the secondary volume after the completion of the SnapVault transfer.
  - Requires Data ONTAP 7.3.1 or later
- Although Protection Manager can support dedupe on secondary volumes, you need Provisioning Manager (with the secondary provisioning policy with dedupe enabled) to enable Protection Manager to create deduped volumes for the SnapVault secondary. If Provisioning Manager isn’t licensed, this can still be achieved by manually enabling dedupe on the secondary volume from the storage system.
  - With Protection Manager 4.0 you would have Provisioning Manager licensed.
- When importing SnapVault relationships coming from an Open Systems SnapVault system, choose the “Remote backups only” policy. This policy will only import SnapVault relationships that match the backup connection of a policy.
- Protection Manager 3.8 or later has the capability to dynamically resize the backup volumes (SnapVault /qtree SnapMirror secondary volumes). Set the option `dpDynamicSecondarySizing` to `enable` to automatically resize the secondary volume before each backup update job.
- Prior to DFM version 3.8, Protection Manager deleted the relationships that had at one time been part of a data set after the reaper time of two hours, but not any more. You can now use the option

`dpReaperCleanupMode` (available only in Protection Manager 3.8 or later) to control, at a granular level, which relationships PM can delete.

  - Use the value `Orphans` if you don't want to delete the imported relationships. The relationships created by Protection Manager will still be deleted.
  - Use the value `Never` if none of them should be deleted.
  - Use the value `Automatic` (not recommended) to go with the legacy operation.

- NetApp recommends the following options when using Protection Manager 3.7 and later:
  - Set the `maxActiveDataTransfers` for storage systems (use dfm host set) based on the stream count limit for Data ONTAP. The maximum value can be set at up to 140.
  - `dpScheduledJobExpiration=12h` Suppose that 150 jobs are started in Protection Manager and that 100 of those jobs finish in 12 hours. The remaining 50 jobs are dropped and an error is logged in the dfmscheduler.log file. There is no other record of those dropped jobs and a retry will NOT be done after 12 hours. Depending on the number of jobs you have you might want to set this value high for retries.
  - If the option `dpDynamicSecondarySizing` is disabled, then you can set the option `pmAutomaticSecondaryVolMaxSizeMb=250000` (or a higher value as appropriate). Protection Manager uses the entire aggregate size for creating destination target volumes; this causes a performance issue when the WAFL® scanner is activated. To avoid impact, you can limit the target size.
- Refer to "Data Transfer Reports" in Operations Manager to understand the data transfer rate, amount, duration, and so on for SnapMirror, SnapVault, and OSSV relationships.

12.1 IMPORTING EXISTING SNAPVAULT AND SNAPMIRROR RELATIONSHIPS

Please verify that you are assigned an administrator role, such as the GlobalBackup and GlobalDatabase roles or the GlobalFullControl role, that enables you to import relationships.

Please make sure that you have determined the data set to which you will import the external relationships and the connections in that data set that you will associate with each external relationship.

DECISIONS TO MAKE BEFORE IMPORTING EXTERNAL RELATIONSHIPS

Before you use the Import Relationships wizard to import external protection relationships into a data set, you need to decide which data set(s), if any, meet(s) the requirements of the relationships, whether you need to create a new data set, and which connection to associate with each external relationship.

SELECTING THE DATA SET

Review the existing data sets listed on the Data Sets window Overview tab to see if any are suitable for the external relationships. Consider the policy applied to the data set and the protection schedule used by the policy. Are the protection requirements of the other data set members the same as for the external relationships you want to import?

You cannot import an Open Systems SnapVault host or directory into a data set when a local backup schedule is already defined on the primary node; you can import them only into data sets that have policies specifying either no protection or remote backups, or those with no local backup schedule.

When selecting a relationship to associate with a policy connection, take into consideration that the licensed protection application adds the source and destination storage objects to the data set. For example, when importing a SnapVault relationship, you select the source qtree or directory to be added to the source node and the licensed protection application adds the destination volume to the destination node. If the user adds a volume SnapMirror relationship, the source and destination volumes are added to the corresponding data sets.

If there are no existing data sets that meet the requirements of the external relationships you want to import, create a new data set.

DETERMINING WHICH CONNECTION TO ASSOCIATE WITH A RELATIONSHIP

The data set you select might have more than one connection to which you can import an external relationship. For example, a data set protected by the chain-of-two-mirrors policy has two mirror connections into which you might import a SnapMirror relationship.

Review the policy settings for each connection and node and the resources assigned to each node to help you decide which connection is the better match for the external relationship.

To review the policy settings for each connection and node in the data set, go to the Protection Policies window, select the policy applied to the data set, and then click Edit → Nodes & Connections.

To review the resources assigned to each destination node in the data set, go to the Data Sets window Overview tab, select the data set, and then, in the Graph area, click the node you want to check.

- When importing relationships, NetApp Protection Manager deletes SnapMirror schedules (only when using pre-3.8 versions; starting in 3.8 the snapmirror.conf entries are not deleted) but leaves Scheduled Snapshot and SnapVault "snap sched" schedules in place on the storage system. If you want all schedules centrally managed, they need to connect to the storage system and you need to disable all storage system resident schedules.
- As a best practice, NetApp recommends staggering backup schedules if performance issues occur due to too many backups getting started at once.
- To view the number of backup jobs starting at the same time, look at the file dfmcmd.log on the DFM server.

12.2 MIGRATING THE SECONDARY VOLUMES OF SNAPMIRROR AND SNAPVAULT

Protection Manager 4.0 has the capability to migrate secondary volumes of SnapMirror and SnapVault. The following are the points that need to be considered for this feature:

- In order to migrate the secondary volumes through Protection Manager, the relationships should be managed by Protection Manager. Hence, if you intend to migrate a secondary volume of a relationship created outside Protection Manager, you must import it to a data set.
- This feature will not work on volumes with any of the export protocols enabled (NFS, CIFS, FCP, and iSCSI). Hence, in a DR-capable data set, you cannot migrate a secondary volume if the secondary volumes have exports configured.
  - You need to remove the exports manually (through FilerView®/CLI, perform a refresh on the storage system object in Operations Manager to recognize the change) and migrate the secondary volume.
  - Once the migration is complete, you can recreate these exports.
  - This feature will not work if the status of the DR-enabled data set is "Failing over" or "Failed Over."
- This feature will not work on secondary volumes that have clones.

The following points deal with the impact of data transfer on the secondary volume during its migration process:

- While migrating a SnapMirror secondary volume, both scheduled and on-demand backups are suspended for the entire course of migration.
- While migrating a SnapVault/qtree SnapMirror secondary volume, the backups happen in parallel when the migration is in process; however, the backups are suspended for a brief period of time when there's a cutover from the old secondary volume to the new secondary volume.
- SnapMirror needs to be licensed on both source and destination storage systems for this feature to work properly.
  - If SnapMirror licenses are not available, Protection Manager can still migrate the SnapVault secondary volume, but the data transfer from the primary to the secondary is suspended for the entire course of migration.

13 RESTORE GUIDELINES

Review these guidelines prior to restoring data using Protection Manager.

- Protection Manager restores data components as small as single files and as large as a volume.
- When restoring data to its original location, you can choose the option “`Warn about overwrite and out-of-space conditions on the destination`” and Protection Manager will shoot a warning if the existing destination files are going to be overwritten.
  - Protection Manager can also shoot a warning message if the destination doesn’t have enough capacity.
- In the case of restoring files from a tertiary volume (for example, Primary -> Mirror -> Archive), the restore can be directed to the original location or a custom location.
- In an Open Systems SnapVault relationship, if the secondary storage is in a different domain than its Open Systems SnapVault host, the /etc/hosts files must include FQDN-IP mapping.
- When Protection Manager restores copies of the primary data, Data ONTAP adds a `restore_symboltable` file in the destination directory. After you successfully restore the desired data, you can delete this file from the directory.
  - This is fixed in release 4.0.1 or later.
- For more information about the `restore_symboltable` file, see the Data ONTAP main pages.
- Restore copies’ data from a “backup” to an active file system either at the original location or another location. If the user wants to restore an entire volume, including its Snapshot copies, the user must use VSM restore from outside Protection Manager.
- Use the option `dpRestoreTransfersPerHost` if you want to achieve parallel restore to the same host. The default value is 8; the maximum value is 64.

14 NETWORK INTEGRATION

For those environments in which storage controllers are hosted on multiple networks or there is a segregated “backup” network created for the backup and replication of data, you need to set parameters to indicate to Protection Manager which interface(s) to use. By default the IP address Protection Manager uses is the one that the storage controller was discovered with.

- Protection Manager is directed to use interfaces for SnapVault by setting the NDMP preferred interfaces on the primary storage controllers. This is set on the storage controller with the option `ndpmd.preferred_interface`; the default value is disabled, indicating no preference. This should also be changed if you wish to use a particular interface on the storage system.
- `ndmpd.preferred_interface 10.187.24.87`

Protection Manager is directed to use interfaces for SnapMirror by setting “hostpreferredIP1” and “hostpreferredIP2” for storage systems in DFM at the individual system level. To view and set the preferred interface, use the following process:

- Determine the object ID of the source and destination storage controller with `dfm host list.`
- View the current settings for the object with `dfm host get object_id_from_command_above.`
- Set the preferred IP address for VSM for this object: `dfm host set object_id_from_command_above hostPreferredAddr1=new_IP_address`
- Confirm that the setting was written with `dfm host getobject_id_from_command_above.`

The returned list should indicate the “new” IP address you wish to use.

NetApp provides no representations or warranties regarding the accuracy, reliability or serviceability of any information or recommendations provided in this publication, or with respect to any results that may be obtained by the use of the information or observance of any recommendations provided herein. The information in this document is distributed AS IS, and the use of this information or the implementation of any recommendations or techniques herein is a customer’s responsibility and depends on the customer’s ability to evaluate and integrate them into the customer’s operational environment. This document and the information contained herein may be used solely in connection with the NetApp products discussed in this document

www.netapp.com