NETAPP UNIVERSITY

Data ONTAP 8.0 7-Mode NCDA NSO-154 Certification

Study Guide

Course Number: STRSW-ILT-ANCDA-D87M
Revision, Date: 2.1, 20MAR2011

ATTENTION

The information contained in this guide is intended for training use only. This guide contains information and activities that, while beneficial for the purposes of training in a closed, non-production environment, can result in downtime or other severe consequences and therefore are not intended as a reference guide. This guide is not a technical reference and should not, under any circumstances, be used in production environments. To obtain reference materials, please refer to the NetApp product documentation located at http://now.netapp.com/ for product information.

COPYRIGHT

NetApp reserves the right to change any products described herein at any time and without notice. NetApp assumes no responsibility or liability arising from the use of products or materials described herein, except as expressly agreed to in writing by NetApp. The use or purchase of this product or materials does not convey a license under any patent rights, trademark rights, or any other intellectual property rights of NetApp.

The product described in this manual may be protected by one or more U.S. patents, foreign patents, or pending applications.

RESTRICTED RIGHTS LEGEND

NetApp Documentation is protected by Copyright and is provided to U.S. Government Agencies with LIMITED RIGHTS as defined at FAR 52.227-14(a). Use, duplication, or disclosure by the U.S. Government is subject to the restrictions as set forth therein. In the event of use by a DOD agency, the Government's rights in Documentation are governed by the restrictions in the Technical Data Commercial Items clause at DFARS 252.227-7015 and the Commercial Computer Software and Commercial Computer Software Documentation clause at DFARS 252.227-7202.

TRADEMARK INFORMATION

NetApp, the NetApp logo, Go Further, Faster, Data ONTAP, Appliance Watch, ASUP, AutoSupport, Bolt Design, Center-to-Edge, ComplianceClock, ComplianceJournal, ContentDirector, Cryptainer, Data Motion, DataFabric, DataFort, Decru, Decru DataFort, Evolution of Storage, Exec-Vault, FAServer, FilerView, FlexCache, FlexClone, FlexShare, FlexVol, FPolicy, Get Successful, gFiler, LockVault, Manage ONTAP, MultiStore, NearStore, NetApp Availability Assurance, NetApp IT As A Service, NetApp ProTech Expert, NetCache, NOW, NOW (NetApp on the Web), ONTAPI, Raid-DP, Replicator-X, SANscreen, SecureAdmin, SecureShare, Shadow Tape, Simulate ONTAP, SmartClone, SnapCache, SnapCopy, SnapDrive, SnapLock, SnapManager, SnapMirror, SnapMover, SnapRestore, Snapshot, SnapStore, SnapSuite, SnapValidator, SnapVault, Spinnaker Networks, Spinnaker Networks logo, SpinCluster, SpinFlex, SpinFS, SpinHA, SpinMove, SpinServer, SpinStor, StoreVault, SyncMirror, Tech OnTap, Topio, vFiler, VFM, VFM (Virtual File Manager), WAFL, and Web Filer are either trademarks, registered trademarks, or service marks of NetApp, Inc. in the United States and/or other countries.

Not all common law marks used by NetApp are listed on this page. Failure of a common law mark to appear on this page does not mean that NetApp does not use the mark nor does it mean that the product is not actively marketed or is not significant within its relevant market.

Apple and QuickTime are either trademarks or registered trademarks of Apple Computer, Inc. in the United States and/or other countries.

Microsoft and Windows Media are either trademarks or registered trademarks of Microsoft Corporation in the United States and/or other countries.

RealAudio, RealNetworks, RealPlayer, RealSystem, RealText, RealVideo, RealMedia, RealProxy, and SureStream are either trademarks or registered trademarks of RealNetworks, Inc. in the United States and/or other countries.

All other brands or products are either trademarks or registered trademarks of their respective holders and should be treated as such.

NetApp is a licensee of the CompactFlash and CF Logo trademarks.

TABLE OF CONTENTS

PURPOSE OF THIS GUIDE

This guide helps you prepare for the NCDA (NSO-154) certification exam. The guide reviews generally the material that is provided in the Accelerated NCDA Boot Camp Data ONTAP 8.0 7-Mode course.

CERTIFICATION DESCRIPTION

As a NetApp Certified Data Management Administrator (NCDA), you are an expert in managing and optimizing a NetApp® storage system that runs the Data ONTAP® operating system in NFS and Windows® (CIFS) multiprotocol environments:

You have proven your ability to:

- Work with multiple issues and within multiple environments—CIFS, NFS, FC, SCSI, iSCSI, TCP/IP
- Provide in-depth support
- Perform administrative functions
- Manage performance
- Implement high-availability (active-active) controller configurations and SyncMirror® and MetroCluster® solutions to ensure data availability and recovery
- Use the SnapMirror®, SnapRestore®, and SnapVault® products to manage and protect data

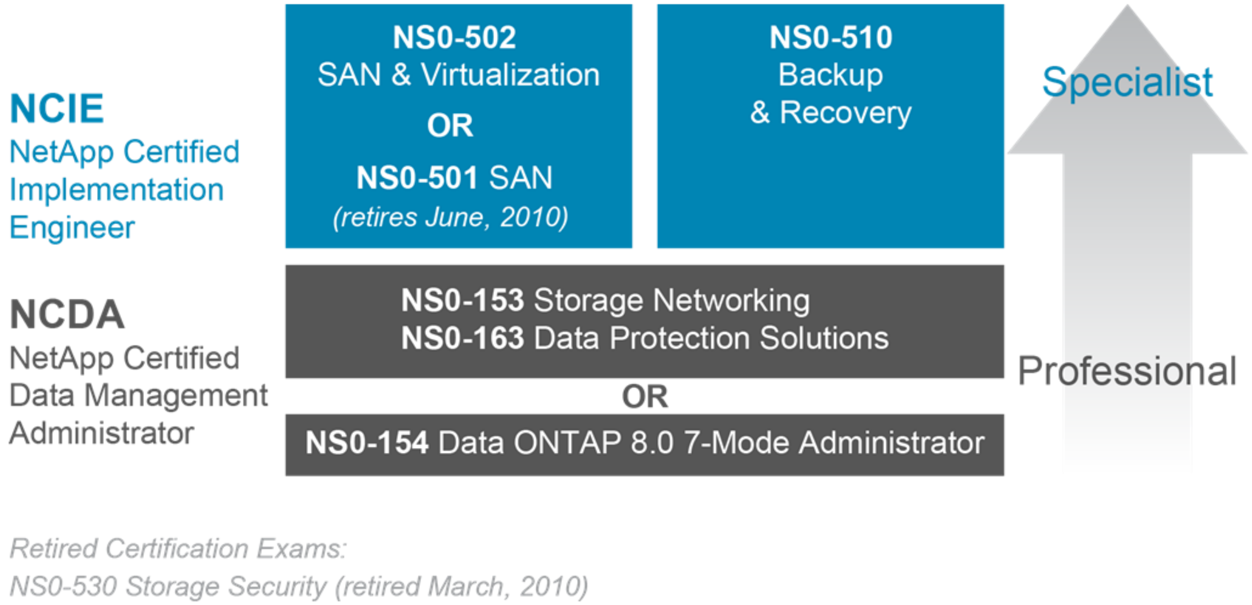

NCIE
NetApp Certified Implementation Engineer
NS0-502
SAN & Virtualization
OR
NS0-501 SAN
(retires June, 2010)
NS0-510
Backup
& Recovery
Specialist
NCDA
NetApp Certified Data Management Administrator
NS0-153 Storage Networking
NS0-163 Data Protection Solutions
OR
NS0-154 Data ONTAP 8.0 7-Mode Administrator
Professional
Retired Certification Exams:
NS0-530 Storage Security (retired March, 2010)

FIGURE 1: NETAPP CERTIFIED DATA MANAGEMENT ADMINISTRATOR (NCDA)

Previously, to achieve NCDA certification, candidates had to pass two exams (NS0-153 and NS0-163). While this option is still available at the time of this writing, NCDA candidates may now sit for one exam (NS0-154) based on Data ONTAP 8.0 7-Mode operating system. This study guide focuses on the newer exam but maintains the structure of the NS0-153 and NS0-163 exams for backwards compatibility.

NOTE: Neither successful completion of the Accelerated NCDA Boot Camp Data ONTAP 8.0 7-Mode course nor the successful learning of the contents of this guide guarantees passing the exam. At least six months of experience administering NetApp storage systems is expected.

EXAM NS0-154: STORAGE NETWORKING CONCEPTS

As a NetApp Certified Data Management Administrator, you will have proven skills in performing in-depth support, administrative functions, and performance management for CIFS, NFS, and FC for SCSI or iSCSI for TCP/IP protocols on a NetApp storage system that runs the Data ONTAP operating system in NFS and Windows (CIFS) multiprotocol environments.

SKILLS TESTED

- Describe the configuration requirements for NFS in a storage system environment
- Configure the storage system to support Kerberos™ security
- Explain and identify the frequently used CIFS components
- Create, configure, and manage CIFS shares, groups, and permissions
- Analyze storage system statistics and identify potential performance issues
- Analyze NFS performance by using the `sysstat` and `nfsstat` commands
- Investigate, identify, troubleshoot, and implement solutions in CIFS and NFS environments
- Identify the supported storage area network (SAN) configurations and the necessary hardware and software components (including NetApp Support Console)
- Identify concepts, commands, and procedures for using the UNIX® file system, especially in relationship to creating and mounting file systems that use UFS on storage system-based logical unit numbers (LUNs)
- Identify the components and tools that are required to create a LUN
- Install or update HBA drivers and firmware to enable Fibre Channel (FC) protocols for SCSI and TCP/IP protocols for iSCSI
- Use the sio_ntap utility, as well as storage system commands, to gather data for performance and problem analysis
- Collect data to assist with troubleshooting hardware, operating systems, and applications
- Configure 64-bit aggregates for a system running Data ONTAP 8.0 7-Mode

RECOMMENDED COURSES

- Web-based: *Data ONTAP 8.0 7-Mode Fundamentals*
- Instructor-led: *Data ONTAP 8.0 7-Mode Administration*
- Instructor-led: *Accelerated NCDA Boot Camp Data ONTAP 8.0 7-Mode*

EXAM PREPARATION

This section describes various NetApp FAS learning points. The points are relevant to the NS0-153 and NS0-154 exams and are related to storage networking concepts. However, in addition to discussing the exam topics, the section provides a summary of a range of NetApp technologies.

Figure 2 highlights the subjects covered in the NS0-154 exam (white text) and the topics covered within each subject (black text).

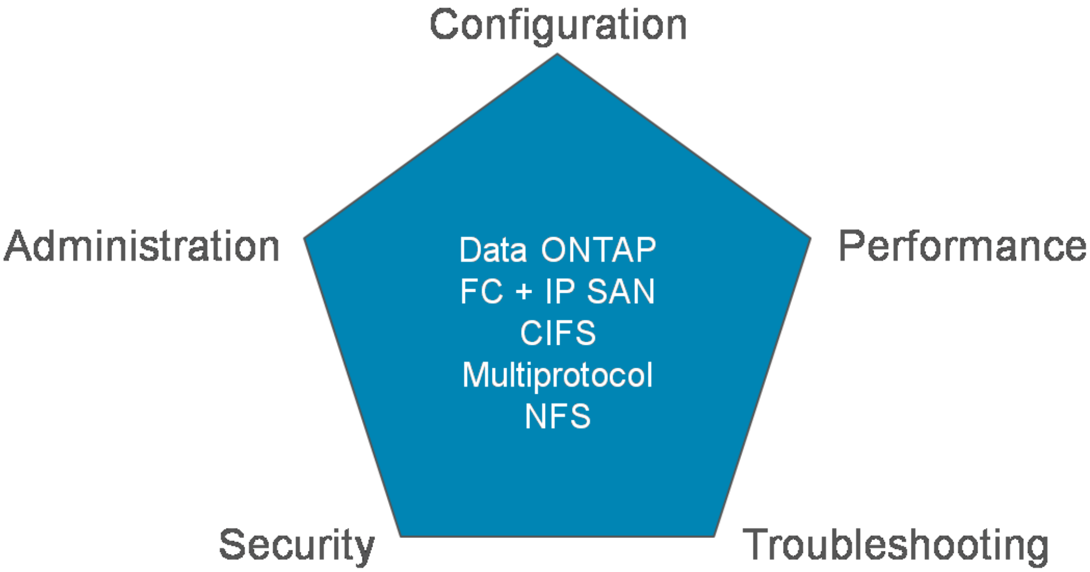


FIGURE 2: TOPICS COVERED IN THE NS0-153 AND NS0-154 EXAM

The 154 exam covers these topics and additional topics that are not identified here.

DATA ONTAP SOFTWARE

The operating system of the NetApp FAS storage controller is named Data ONTAP. Data ONTAP is the foundation of the NetApp Unified Storage architecture, supporting a wide range of storage protocols and storage-management software features.

The NetApp FAS storage controllers are designed around the NetApp Unified Storage architecture and support numerous storage protocols, software features, and storage tiers in an appliance-like design. Refer to Figures 3 for identification of the supported protocols.

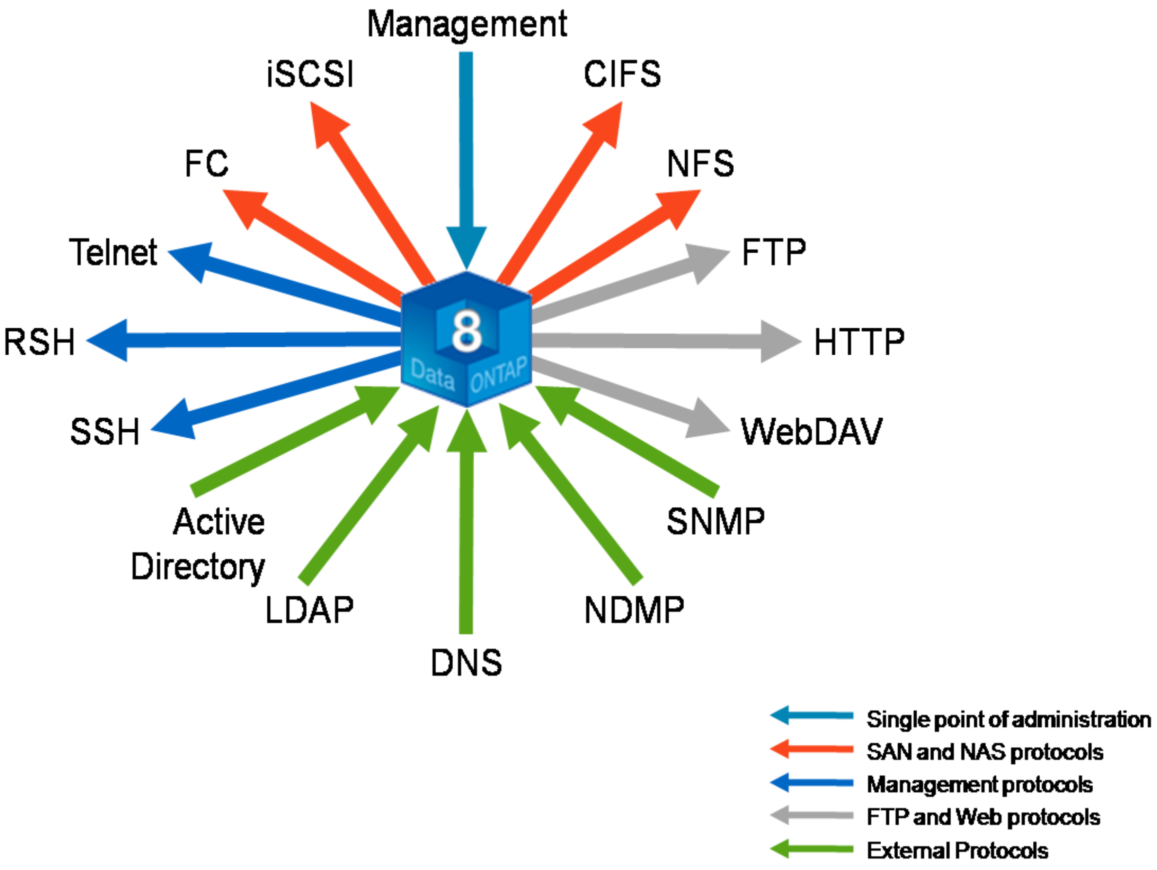


FIGURE 3 - PROTOCOLS SUPPORTED BY THE NETAPP FAS CONTROLLERS

NOTE: This figure identifies only the major storage, management, and authentication protocols. For brevity, some protocols, such as the NetApp API (Manage ONTAP Solution or ZAPI), and Network Information Service (NIS) are not shown.

CONFIGURATION

When the storage controller is powered on for the first time, it runs the `setup` script. This script configures fundamental parameters such as hostname, network IP addresses, and CIFS authentication (if licensed). You can rerun the `setup` script manually at any time to reconfigure the parameters.

To manually reconfigure the FAS controller's Ethernet network interfaces, you use the NetApp System Manager interface or the `ifconfig` command; for example: `ifconfig ns0 192.168.1.1 netmask 255.255.255.0`

NOTE: Most changes made from the command-line interface are transient. To enable persistence after a reboot, you must enter the changes into the `/etc/rc` file.

You may also need to configure the built-in FC ports (located on the motherboard). The built-in FC ports can function in one of two modes:

- Initiator mode (the default)
  - Use the initiator mode to connect to disk expansion drawers.
  - Use the `fcadmin config <adapter_name> -t initiator` command to set an adapter to initiator mode.
- Target mode
  - Use target mode to connect to host systems (that is, Windows or UNIX servers).
  - Use the `fcadmin config <adapter_name> -t target` command to set an adapter to target mode.

NOTE: The `fcadmin` command applies only to the built-in FC ports. Any FC HBA adapter that you purchase is predefined as either an initiator or target adapter.

The `fcp show adapter` command returns only the FC ports that are configured to function in *target* mode. The `fcp show initiator` command returns all host FC *initiator* ports that are visible via the controller's FC *target* ports.

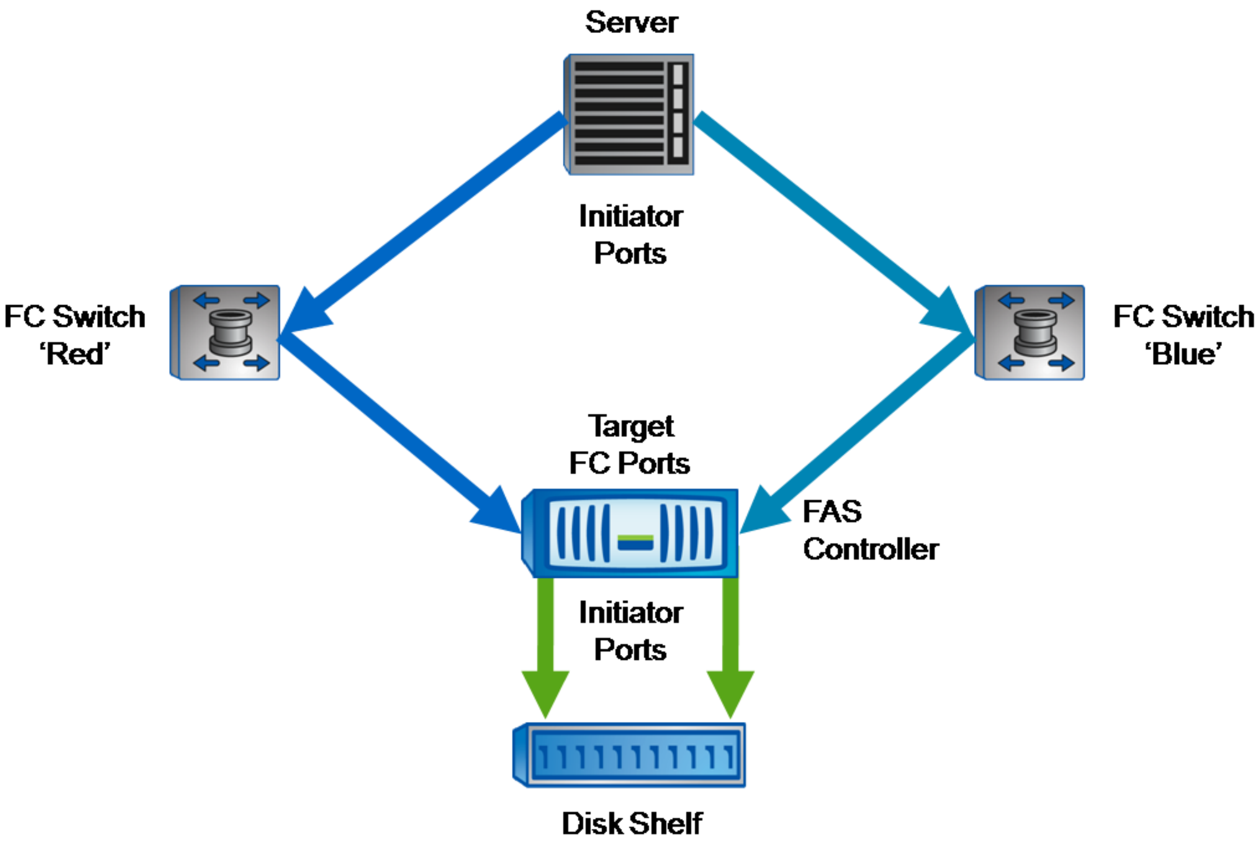


FIGURE 4: FIBRE CHANNEL INITIATOR-TARGET PORT RELATIONSHIPS

NOTE: Similar initiator-target relationships exist in iSCSI connections.

ADMINISTRATION

The administration of the NetApp FAS controller can be performed via various administrative interfaces.

1. **NetApp System Manager**: a GUI designed for single-system configuration
2. **Command-line interface**: accessed via Telnet, Remote Shell (RSH), or Secure Shell (SSH)
   - The `aggr status` command displays the existing aggregates.
   - The `cifs shares -add` … command defines a CIFS share.
   - The `options dns.enable on` command enables Domain Name System (DNS) name resolution (which requires further configuration).

   NOTE: In Data ONTAP 8.0 7-Mode, only secured protocols (such as SSH) are enabled by default. Open protocols (such as RSH and Telnet) are disabled by default. NetApp recommends use of the default.
3. **Operations Manager**: a licensed tool that is installed on a host and that provides sophisticated management tools for one or more storage systems. Operations Manager includes the following products:
   - Performance Advisor
   - Protection Manager
   - Provisioning Manager

The license for Operations Manager enables Performance Advisor, Protection Manager and Provisioning Manager are separately licensed products.

NOTE: Management of the FAS storage infrastructure can be promoted to the host operating system (Windows or UNIX) by the SnapDrive® tool and to the application layer by the SnapManager® tool. Both tools are optional purchases and require licenses.

The various SAN and network-attached (NAS) protocols differ in many technical details: SAN provides block-level access, and NAS provides file-level access. Additionally, some protocols use FC connections, and others use Ethernet. However, both SAN and NAS provide remote systems (hosts) with access to centralized, shared storage.

Perhaps the simplest way to define the difference between SAN and NAS is by describing how the file system is managed and how access is shared. For a SAN device, the controller provides block-level access to the hosts. The hosts then create and manage their own local file system (for example, ext3), and the local file system is typically not shared with other hosts. On the other hand, NAS storage uses the file system on the controller—the WAFL® (Write Anywhere File Layout) system, and the controller provides shared file-level access to multiple remote systems (hosts).

Both FC and iSCSI SANs support multipathing. With multipathing, there are redundant physical paths (two or more) between a host and the controller. Refer to Figure 4 for an example.

Multipathing ensures storage availability, even if a SAN hardware failure occurs. NetApp supports various multipathing options for both FC and iSCSI.

- FC and iSCSI multipathing
  - With the host platform’s native multipath I/O (MDIO) driver
  - With the NetApp device-specific module (DSM) for Windows
  - With the Veritas Dynamic Mulitpath (VxDMP)software
- iSCSI only multipathing, with iSCSI’s inherent multiple connections per session (MCS)

You choose the multipath solution that best meets the needs of your environment.

The NetApp on the Web (NOW)® online support and services site is available to all customers and business partners. You should become familiar with the NOW site, as it is a primary resource for questions regarding the configuration or performance of a FAS storage controller. Some of the resources available on the NOW site include:

- A knowledgebase (searchable) of NetApp product information
- Online manuals, for all NetApp products
- Software downloads, including updates and evaluation versions
- Return Merchandise Authorization (RMA), used to request replacement of failed hardware
- Bugs online, used to review all known software bugs
- Release Comparison, used to identify the version of Data ONTAP in which a particular bug was fixed

64-BIT AGGREGATES

Before Data ONTAP 8.0, aggregates and, consequently, the volumes within the aggregates were based upon a 32-bit architecture. This architecture effectually limits the size of an aggregate to 16 TB.

As storage demands increase, the 16-TB limitation causes three major issues. First, as the size of each disk drive increases, the number of drives per aggregate must decrease. For example, in Data ONTAP 7.3, an aggregate can accommodate only nineteen 1-TB data drives. Second, performance is directly affected. As the number of drives decreases, the read and write performance that can be realized from an aggregate also decreases.

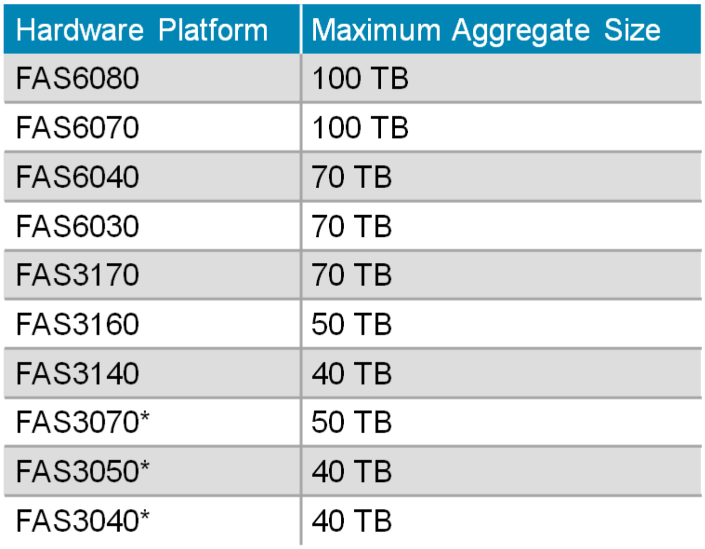
The third issue with the 16-TB limitation is that aggregate management becomes more complex and less efficient. For example, a FAS6070 system with 500 TB of storage requires a minimum of 32 aggregates. Such a requirement greatly increases the complexity of managing large storage arrays.

NetApp is pleased to announce that its Data ONTAP 8.07-Mode operating system supports 64-bit aggregates. This architecture overcomes the disadvantages of the 16-TB limitation.

NOTE: In Data ONTAP 8.0 Cluster-Mode, 64-bit aggregates are not available.

Data ONTAP 8.0 7-Mode supports aggregates of up to 100 TB—running on top-of-line hardware with greater aggregate sizes available in the future.

See Figure 5 for information about platforms and maximum aggregate sizes.



Hardware Platform	Maximum Aggregate Size
FAS6080	100 TB
FAS6070	100 TB
FAS6040	70 TB
FAS6030	70 TB
FAS3170	70 TB
FAS3160	50 TB
FAS3140	40 TB
FAS3070*	50 TB
FAS3050*	40 TB
FAS3040*	40 TB

* Supported only through a Policy Variation Request (PVR)

FIGURE 5: 64-BIT AGGREGATES—MAXIMUM AGGREGATE SIZE PER PLATFORM

In Data ONTAP 8.0, NetApp added the -B switch to the `aggr create` command. This switch accepts 32 or 64 as a valid parameter.

```
aggr create <aggr-name> [-B {32 | 64}]…
```

The 32 parameter designates a 32-bit aggregate architecture, and the 64 parameter designates a 64-bit aggregate architecture. The 32 parameter is the default.

To create a 64-bit aggregate, you must explicitly use the 64 parameter. For example, if a storage administrator enters the following: `aggr create sales aggr -B 64 24`, a 64-bit aggregate is created. The new aggregate is named “sales aggr,” and 24 spare disks, chosen by Data ONTAP, are added the new 64-bit aggregate.

NOTE: The command succeeds only if the specified number of spare disks exists. As in earlier Data ONTAP releases, the storage administrator can manually choose the disks to be added by specifying a `-d` switch and the disk names.

When a storage administrator upgrades a storage system from Data ONTAP 7.3 or earlier to Data ONTAP 8.0, all existing aggregates are designated as 32-bit aggregates. In Data ONTAP 8.0, you cannot directly convert a 32-bit aggregate to a 64-bit aggregate. Similarly, volumes within a 32-bit aggregate cannot be moved directly to a 64-bit aggregate. If you want to take advantage of the 64-bit aggregate feature, you must migrate data from a 32-bit aggregate to a 64-bit aggregate. To migrate data, you may use the `ndmpcopy` command.

PERFORMANCE

You can use several commands on the FAS storage controller to collect system performance data. Performance data can consist of a broad summary of system performance or of details about specific performance parameters.

Some common commands for collecting performance data are:

- `sysstat`
  - The default output includes CIFS, NFS, HTTP; CPU, nonvolatile RAM (NVRAM), network interface card (NIC), and disk performance values. The `–b` parameter adds block-level (that is, FC and iSCSI) performance to the output.
  - Example: The `sysstat –s 5` command runs every five seconds and prints a summary on termination. The default interval is 15 seconds.
  - Example: The `sysstat –u` command displays extended utilization data, including consistency point (CP) time and type. The command can be used to identify when a controller is busy. Outputs such as cp_from_log_full (‘F’) and cp_from_cp (‘B’) indicate that the controller is busy.
- `statit`
  - This command can output performance data on various object types, specific instances of object type, and other performance counters.
  - The command displays many performance items, including per-disk performance data.
- `stats`
  - This advanced mode command can provide low-level performance data.
  - Example: The `stats show cifs:cifs:cifs_latency` command displays the average latency value for the CIFS protocol.
  - Performance counters can also be accessed via the Windows PerfMon GUI (but only if CIFS is enabled).

Some commands for collecting statistics about the performance of the Ethernet network interface and other data are:

- `netstat:` The `netstat -i` command identifies the network interfaces, the number of in and out network packets, errors, and collision values
- `ifstat:` The `ifstat –a` command displays a low level view of interface performance data, including transmitted (Tx) and received (Rx) bytes per second

If you cannot resolve a performance problem by using the `systat, stats, status, netstat,` or `ifstat` commands, then you may need to download the `perfstat` command from the NOW online support and services site.

Characteristics of the `perfstat` command:

- Captures all needed performance information
- Captures information from hosts and filers
- Captures all information simultaneously and thus enables cross correlation
- Operates on all host platforms and all filer platforms
- Records all captured data in one plain-text output file

DATA ONTAP SECURITY

The security of the FAS controller, the security of the Data ONTAP operating system, and the security of the data stored on the controller are different topics. This section deals only with Data ONTAP security.

The Data ONTAP operating system supports role-based access control (RBAC), where defined roles, with specific capabilities, can be bound to groups. Individual users are then assigned to the groups. The assigned users can perform only the tasks that the roles assigned to their groups allow them to perform.

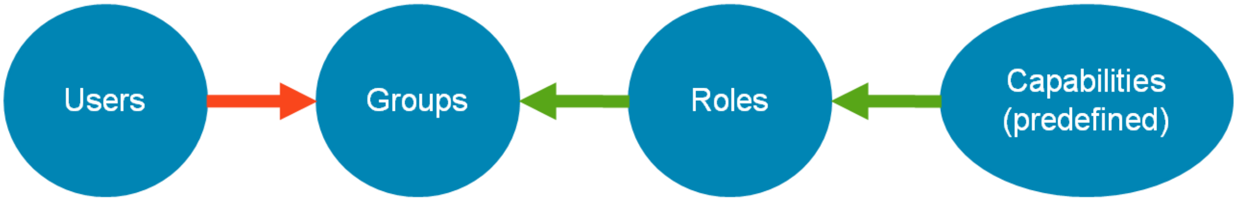


FIGURE 6: ROLE BASED ACCESS CONTROL

For administrative purposes, one or more accounts are usually defined on the FAS storage controller. The `useradmin` command is used to create and modify local admin accounts.

Examples of the `useradmin` command:

- To create an administrative user

```
useradmin user add <username> -g <groupname>
```

- To list the local groups

```
useradmin group list
```

- To delete a role

```
useradmin role delete <rolename>
```

- To assign a capability to a role

```
useradmin role modify <rolename> -a <capability>
```

You can use the `useradmin` commands, with their obvious variations, to add, list, delete, or modify users, groups, or roles. The capabilities are predefined and cannot be changed.

NOTE: There is always at least one administrative account, `root`, which cannot be deleted.

TROUBLESHOOTING

If you need to capture a network packet trace, so you can analyze a protocol level problem, then you could use the `pktt` command on the FAS controller.

Characteristics of the `pktt` command:

- `Is` normally run for a short period, capturing network traffic on a particular interface, potentially filtered to target a particular client.
- Produces output in standard tcpdump format. The output can be analyzed in third-party network monitoring tools, such as the following:
  - `ethereal` (download from wireshark.org)
  - `netmon` (download from 14icrosoft.com)

    NOTE: You must first convert the tcpdump output to netmon format by using the `capconv.exe` command that you can download from the NOW online support and services site.

SAN

The NetApp FAS storage controllers support both the FC and iSCSI SAN standards.

One feature (among many) that distinguishes the NetApp SAN architecture from competitive solutions is that the LUNs are defined below the SAN protocol layer. Therefore, a NetApp LUN is not an FC LUN or an iSCSI LUN, but it can be exposed to a host via either or both protocols, even simultaneously. This flexibility allows for easy migration between the SAN access protocols.

CONFIGURATION

Refer to Figure 4 and the accompanying text for a description of the initiator and target configuration of the controller's built-in FC ports.

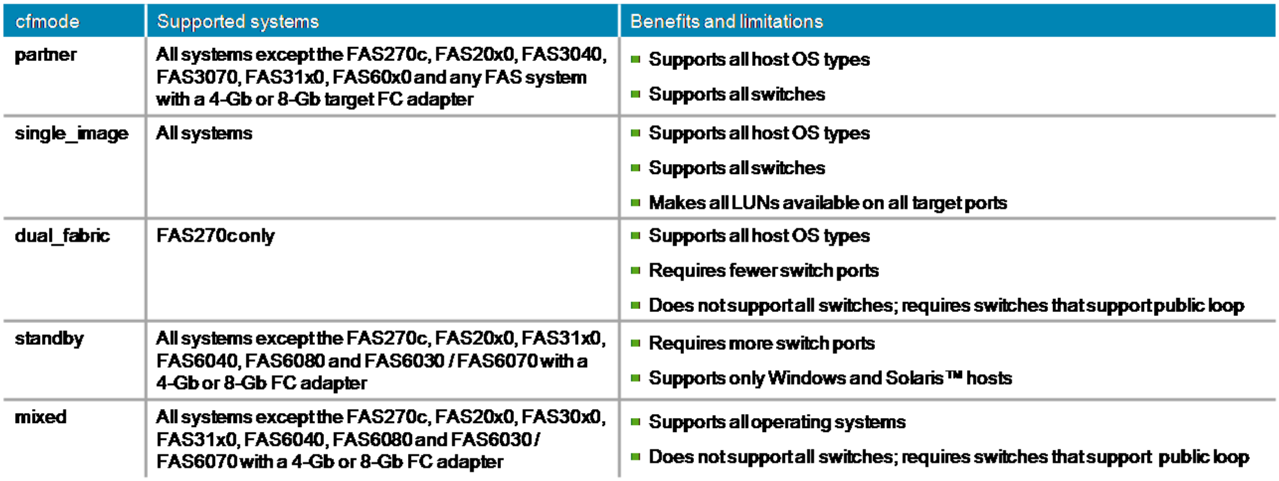
In a high-availability configuration, a LUN that is being accessed via FC is visible from every FC target port on both FAS storage controllers (even though the LUN is actually "owned" by only one controller). This FC access mode, which is called "single image," has been the default failover mode for FC controllers since the release of Data ONTAP 7.2 software. You can change from one to another failover mode (with caution) by using the `fcp cfmode` command. After you change modes, you must reboot the controller.

The failover modes for FC controllers are:

- single_image (the default since Data ONTAP 7.2)
  - LUNs are visible on all FC target ports labeled by their worldwide port names (or WWPNs) on both controllers.
  - A worldwide node name (WWNN) is shared by both controllers.
- partner
- mixed
- dual_fabric
- standby

If you require a description of the earlier FC cluster modes, refer to the product documentation.

Although all of the failover modes for FC controllers are supported for existing systems, only the single_image mode is supported for new storage systems. The following table details some of the features and limitations of the various failover modes:



cfmode	Supported systems	Benefits and limitations
partner	All systems except the FAS270c, FAS20x0, FAS3040, FAS3070, FAS31x0, FAS60x0 and any FAS system with a 4-Gb or 8-Gb target FC adapter	▪ Supports all host OS types ▪ Supports all switches
single_image	All systems	▪ Supports all host OS types ▪ Supports all switches ▪ Makes all LUNs available on all target ports
dual_fabric	FAS270c only	▪ Supports all host OS types ▪ Requires fewer switch ports ▪ Does not support all switches; requires switches that support public loop
standby	All systems except the FAS270c, FAS20x0, FAS31x0, FAS6040, FAS6080 and FAS6030 / FAS6070 with a 4-Gb or 8-Gb FC adapter	▪ Requires more switch ports ▪ Supports only Windows and Solaris™ hosts
mixed	All systems except the FAS270c, FAS20x0, FAS30x0, FAS31x0, FAS6040, FAS6080 and FAS6030/ FAS6070 with a 4-Gb or 8-Gb FC adapter	▪ Supports all operating systems ▪ Does not support all switches; requires switches that support public loop

FIGURE 7: FAILOVER MODES FOR FC CONTROLLERS

NOTE: Failover mode for FC controllers is an FC-specific concept. The iSCSI protocol handles controller failover in a completely different manner.

Conceptually, FC and iSCSI are very similar, although they vary greatly in detail and implementation (for example, FC is a *wire-level* protocol, whereas iSCSI travels on top of a *TCP* connection). However, the end result is the same; they both provide block-level services to a host, so the host can access a LUN.

Both FC and iSCSI are licensed protocols. The only difference between the protocols is that the iSCSI license is provided for no charge with the every FAS controller. You use the `license add <licnum>` command to license a feature (such as FC or iSCSI) on the controller.

After the SAN protocols are licensed, they must be started. No configuration or host access can occur until the protocols are started. Use the following commands to start each of the SAN protocols:

- FC
  - `fcp start`, to start the protocol
  - `fcp status`, to return the status of the protocol
- iSCSI
  - `iscsi start`, to start the protocol
  - `iscsi status`, to return the status of the protocol

Another difference between the FC and iSCSI protocols concerns ports and adapters. The FC protocol can use only a dedicated, target-mode FC port on a specialized FC HBA adapter. An iSCSI protocol can use either a specialized iSCSI HBA adapter or any standard Ethernet port. If a standard Ethernet port is used, then the controller uses the iSCSI software target (ISWT).

If you are using ISWT support in a high-availability controller configuration, then you may notice two ISWT devices:

- ISWTa, for the local controller
- ISWTb, for the partner controller, used for controller failover

The SAN protocols provide block-level access to the LUNs on the storage controller. You can create the LUNs by using various tools, as follows:

- NetApp System Manager, a GUI downloadable from the NOW online support and services site
- Command-line interface, using either of the following two methods
  - To create each item manually, run the `lun create`, `igroup create` and `lun map` commands.
  - To create everything from a wizard, run the `lun setup` script and follow the prompts. When you use the wizard, you do not need to manually create the igroups or map the LUNs.
- SnapDrive, which allows you to manage storage from the host

  NOTE: When you create a LUN, you need to know certain facts, such as the location (path) of the LUN, the OS of the respective host, the capacity required, and the LUN ID.

The LUN ID is the means by which a host identifies a LUN and distinguishes between multiple LUNs. Therefore, LUN that is presented to a host must be unique to the host. However, of course, each host has its own LUNs and LUN IDs, and the LUNs and LUN IDs of each host are independent of the LUNs and LUN IDs of all other hosts.

- Multiple LUNs to the same host must have unique LUN IDs.
- Each host can have its own IDs. IDs can conflict only within a host, not between hosts.

When a LUN is created via the `lun setup` wizard, multiple commands are rolled up into the script. However, if you create a LUN manually, then you must run two additional commands—to make the LUN accessible from a host. To filter which LUNs are visible to which hosts (sometimes called “LUN masking”), you must use the `igroup create` and `lun map` commands.

- `igroup create` defines a relationship between a host (WWPN) and the OS type and identifies whether the current connection is an FC or iSCSI connection.

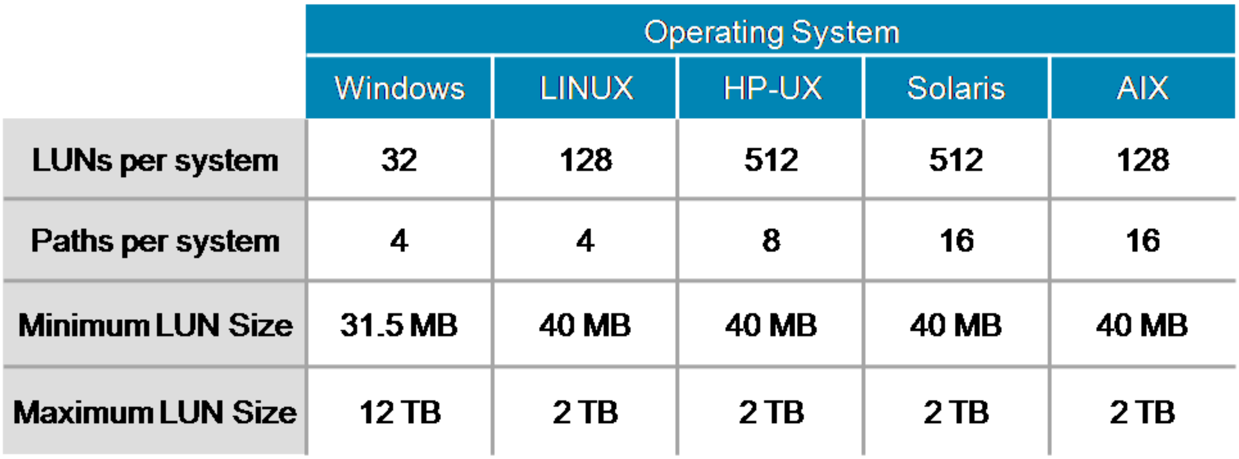
  ```
  igroup create -f -t windows <ig_name> <wwpn>
  ```

- `lun map` defines a relationship between a LUN and the igroup and sets the LUN ID,

  ```
  lun map </path/lun> <ig_name> <lun_id>
  ```

NOTE: Assuming that the SAN zoning is correct, the host should now be able to rescan for and connect to the new LUN.

You need to be aware of the minimum and maximum LUN sizes that are supported by each of the host operating systems. Refer to the Figure 8:



	Operating System				
	Windows	LINUX	HP-UX	Solaris	AIX
LUNs per system	32	128	512	512	128
Paths per system	4	4	8	16	16
Minimum LUN Size	31.5 MB	40 MB	40 MB	40 MB	40 MB
Maximum LUN Size	12 TB	2 TB	2 TB	2 TB	2 TB

FIGURE 8: SUPPORTED LUN SIZES PER HOST PLATFORM

NOTE: Not all supported operating system platforms are shown here. For more detailed information, refer to the NOW online support and services site and host platform documentation.

How a LUN consumes capacity on its parent volume and how LUN capacity is affected by the creation of Snapshot® copies is a complex topic. If a volume that contains a LUN is incorrectly configured, the host may think that the volume has space available for writing data while, in reality, the volume is filled with Snapshot data. When the host attempts to write data, the LUN goes offline, and manual rectification is required. Obviously, the LUN capacity problem creates a very undesirable scenario.

There are several ways to prevent the LUN capacity problem from occurring:

- Space reservation—former best practice
  - Until recently, space reservation was the recommended method to guarantee space for writes to a LUN (regardless of the number of Snapshot copies).
  - Some of the parameters to be configured for this method are volume fractional reserve to 100% and LUN space reservation to Yes.

    `vol options <vol> fractional_reserve 100` (100% by default)

    `lun create` (LUN space reservation is on by default.)

    `lun set reservation` (to change the space reservation of a LUN)

  - Configuration of the parameters causes an amount of space equal to the size of the LUN (100%) to be excluded from Snapshot copies, thus guaranteeing that writable capacity is always available to the host.
  - Such a configuration is sometimes referred to the "two times plus Delta (2X+Δ)" overhead.

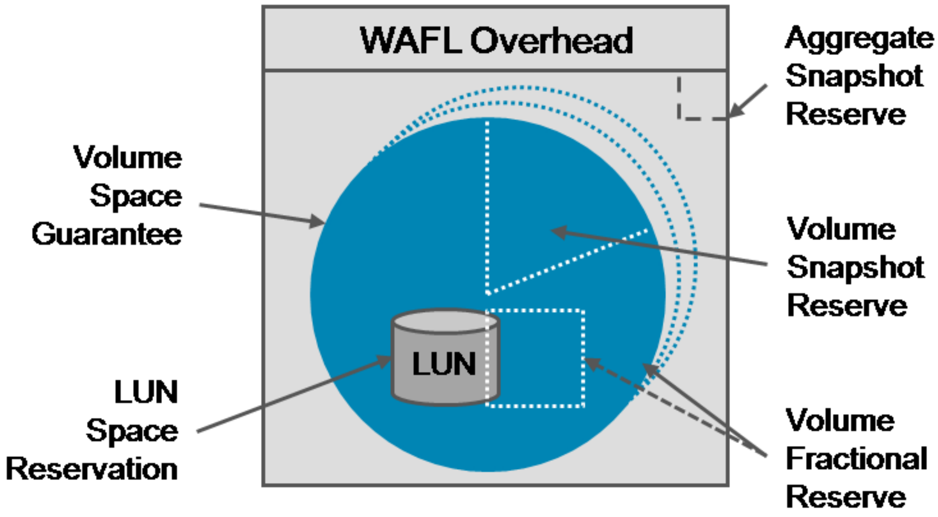


FIGURE 9 - CONFIGURING VOLUME AND LUN PROVISIONING

NOTE: Some documentation may still refer to these former best practices.

- Volume AutoGrow and Snapshot AutoDelete—current best practice
  - In the last few years, two automatic utilization thresholds were introduced, replacing the former best practice (use of the 100% fractional reserve) with the current best practice.
  - Some of the parameters to be configured for this method are Volume AutoGrow and Snapshot AutoDelete. When the parameters are configured, the Volume Fractional Reserve can be safely set to 0%.

```
vol autosize <vol-name> on        (off by default)
snap autodelete <vol-name> on     (off by default)
vol options <vol> fractional_reserve 0
```

  - This parameter configuration sets a utilization threshold at which the containing volume automatically grows and/or at which certain existing Snapshot copies are deleted. Use of the threshold ensures that space is always available for the host to write to the LUN,
  - Use of the threshold changes the capacity required to “one times plus Delta (1X+Δ).” Use of thin provisioning can produce an even better result.

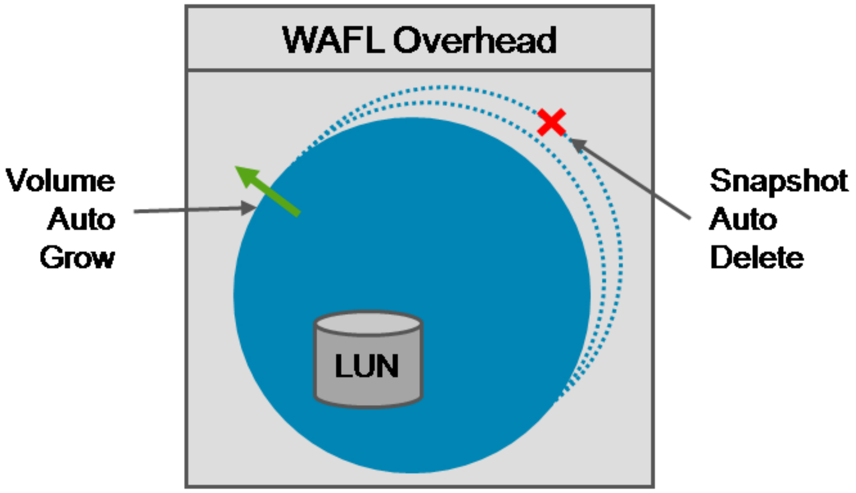


FIGURE 10 - CONFIGURING AUTOMATIC CAPACITY MANAGEMENT

NOTE: For more information about volume and LUN space management, refer to the product documentation.

ADMINISTRATION

You can use lun commands not only to create LUNs but also to perform various other LUN-related actions:

- lun offline, makes a LUN unavailable for host access
- lun move, relocates a LUN from one to another path within the same volume
- lun show, displays various types of information about LUNs
  - To display the LUN to igroup mappings: lun show -m
  - To display the LUNs operating system platform types: lun show -v
- lun clone, instantly creates an read/writable LUN as a clone of an existing LUN

  The new LUN is thin provisioned (no space is consumed until new data is written), and the two LUNs share blocks with a snapshot of the original LUN (known as the "backing" snapshot)

  ```
  lun clone create /vol/vol1/lun2 –b /vol/vol1/.snapshot/lun1
  ```
- lun clone split, splits a LUN clone from its backing snapshot

  Because a LUN clone locks the backing snapshot (that is, prevents the backing snapshot from being deleted), it is recommended that, for long-term use, the relationship be split

  ```
  lun clone split start /vol/vol1/lun2
  ```

NOTE: For more detailed descriptions of the lun command options, refer to the product documentation. Remember that the Snapshot copy of a LUN is created at the volume level. So, all data in the volume is captured in the Snapshot copy including multiple LUNs and NAS data if present in the volume. Therefore, it is best practice to simplify the Snapshot copy process by one-to-one (1:1) relationship between volume and LUN (or in other words, having the volume containing only a single LUN).

A LUN can contain a file system that is managed by the host and/or contain a database that is managed by an application on the host. In this case, when Snapshot copies are created, only the host and the application can ensure the consistency of the file system and database.

Therefore, it is often necessary to coordinate with the relevant host during the Snapshot backup of a LUN. Usually, the coordination process occurs with no disruption to service.

- On the host
  - Quiesce the application or database
  - Flush the host's file system buffers to disk (LUN)
- On the controller, create the Snapshot copy of the LUN.

  ```
  snap create /vol/vol1 <snapshot_name>
  ```

  The copy now contains a consistent image of the host's file system and application database in an idle or offline state.
- On the host, unquiesce the application or database.

NOTE: The NetApp host attachment kit includes a utility to flush the host's file system buffers to disk.

When you create a Snapshot backup of a LUN, you must consider the free capacity in the containing volume. For example, assume that you configure a 400-GB volume and then create a 300-GB LUN inside the volume. If you fill the LUN with data, then a subsequent attempt to create a Snapshot copy fails. The failure occurs because, if the Snapshot copy were created, there would be insufficient free capacity in the volume to allow the host to continue to write to the LUN.

For detailed information about the various capacity management methods that are relevant to Snapshot backup of volumes and LUNs, refer to the Snapshot section.

To restore from a Snapshot backup of a LUN, you can use either of two methods:

- Volume SnapRestore
  - This method restores the entire volume and the LUN.
  - The LUN must be offline or unmapped.
- LUN clone: You create a clone of the LUN and then mount the clone. You can then copy individual files back to the original LUN.

NOTE: If the Snapshot backup contains NAS data (rather than a LUN), then you can browse to the .snapshot directory and copy the individual files back to the active file system or, for a very large file, you can use single file SnapRestore.

Perhaps the best way to manage (and create Snapshot copies of) SAN storage is to use the NetApp host and application integration tools. For each supported host operating system platform, there is a SnapDrive package, and, for many popular business applications, there is a SnapManager package. These tools provide an easy-to-use, highly functional interface for managing the storage controller.

- SnapDrive
  - Create a LUN
  - Connect to a LUN
  - Trigger a new consistent Snapshot copy of a file system
  - Restore a Snapshot backup
  - Clone a LUN
  - Manage iSCSI connections
- SnapManager
  - Trigger a new consistent Snapshot copy of an application
  - Manage the retention of Snapshot backups
  - Restore a Snapshot backup of an application
  - Clone an application or database (only for specific applications)
  - Migrate application data onto LUNs

NOTE: You can perform Snapshot application integration without SnapManager support. To do so, you must create the required automation scripts. Many examples of the scripts can be downloaded from the NOW online support and services site.

PERFORMANCE

The performance of the SAN is dependent on the available physical resources and the host, switch, and controller configurations.

- Controller
  - Model (versus expected performance level)
  - Number of spindles supporting the LUNs
  - Network speed
  - Number of ports or paths
  - Balanced workload between high-availability controllers
  - Background tasks (replication, RAID scrub, RAID rebuild, deduplication, and so on)
- Switch
  - Network speed
  - Port oversubscription
  - Switch workload

- Host
  - Workload type (random versus sequential and read versus write)
  - Network speed
  - Multipathing configuration (high-availability or failover only)
  - Number of paths
  - HBA tuning
  - Correct partition alignment

Because all iSCSI traffic is carried over the TCP protocol, you can list connections (an iSCSI initiator-to-target relationship) and sessions (TCP) separately. The connections-to-sessions relationship can be configured in either of two modes.

- Single connection per session (1:1)
  - Creates one iSCSI connection per TCP session
  - Is supported by the Microsoft Multipath I/O (MPIO) and NetApp DSM for Windows
- Multiple connections per session (MCS)
  - Creates multiple iSCSI connections per TCP session
  - Provides a native iSCSI multipathing capability

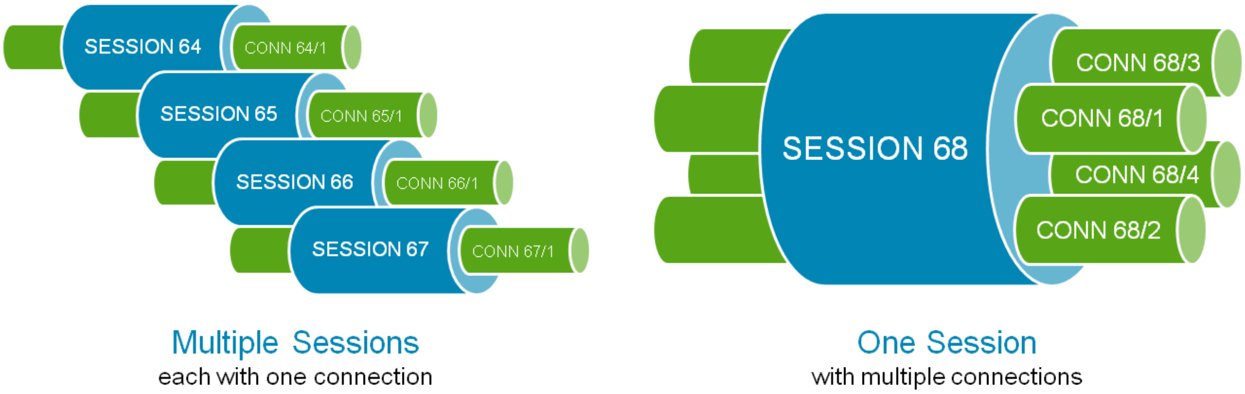


FIGURE 11- ISCSI SESSIONS AND CONNECTIONS

The commands to list the iSCSI connections and sessions are:

```
iscsi session show
iscsi connection show -v
```

NOTE: The advantages and disadvantages of the multipathing methods are beyond the scope of this document.

SECURITY

The security of the SAN is enforced at several levels—some on the controller, some in the fabric, and some in conjunction with the hosts.

- FC
  - LUN masking (igroups)
  - Port masking (portsets)
  - SAN zoning
- iSCSI
  - LUN masking (igroups)
  - Port masking (iscsi accesslist)
  - Ethernet VLANs
  - Passwords (iscsi security)
  - Encryption (ipsec)

In an FC SAN, the ability of the various devices to discover each other and communicate across the fabric is controlled by the zone definitions. Only devices that reside in the same zone can communicate. The two main types of SAN zones are:

- Soft zones
  - Is usually defined using the WWPN of the host’s and controller’s FC ports; allows communication between the device’s FC ports
  - Blocks inter-zone traffic by filtering device discovery at the fabric name service but *does not explicitly block* traffic
- Hard zones
  - Is usually defined using the physical port numbers of the FC switch; allows communication between the physical switch ports
  - Blocks inter-zone traffic by filtering device discovery at the fabric name service and by *preventing* traffic between switch ports

When designing FC SAN zones, apply the following recommendations:

- Decide on a naming convention and use only the naming convention that you decided on
- Keep disk and tape devices in separate zones
- Define a zone for every required initiator-target pairing

A concern with an iSCSI SAN is that the block-level traffic and other, less-sensitive data might be travelling over the same physical network. In this situation, the iSCSI traffic is exposed to network snooping and other attacks.

The NetApp FAS controllers support port masking, bidirectional password access, and network encryption of the iSCSI traffic.

- iSCSI port masking allows you to deny specified network adapters the ability to expose an iSCSI initiator group to the hosts:

  ```
  iscsi interface accesslist remove ig0 e0a
  ```

- iSCSI password requires an iSCSI host that is trying to access a LUN to respond to a password challenge. It can also require a controller to answer a password challenge from the host.

  ```
  iscsi security add -i <igroup_name> -s CHAP
        -p <inpassword>  -n <inname>
      [ -o <outpassword> -m <outname> ]
  ```

- IPSec encryption enables encryption of the Ethernet network traffic. Because it is not an iSCSI specific setting, it encrypts all Ethernet traffic (but the client, of course, needs to know how to decrypt the traffic).

  `options ip.ipsec.enable on` and `ipsec policy add`

NOTE: Network encryption, although effective, can negatively affect the performance of what should be a high performance iSCSI solution. Therefore, many sites choose not to encrypt iSCSI traffic and instead deploy a separate network infrastructure or a virtual local area network (VLAN) to isolate the iSCSI traffic.

TROUBLESHOOTING

Troubleshooting a SAN problem requires knowledge of the host, the SAN switches and network infrastructure, and the storage controller. Much of the information associated with this knowledge is outside the scope of this document. However, here are some actions that you can perform to troubleshoot a simple LUN access problem:

- On the controller
  - Can the storage controller see the host’s FC adapters?

    ```
    fcp show initiator
    ```

  - Is the LUN being presented to the host?

    ```
    lun map –v
    ```

    ```
    igroup show
    ```

  - If there a problem with iSCSI password access?

    ```
    iscsi security show
    ```

- On the SAN switches, are the host’s FC initiator ports and the controller’s FC target ports in the same zone?
  - Run the `zoneshow` command (Brocade).
  - Run the `show zone` command (Cisco MDS).
- On the host (Solaris 9)
  - Is the host configured to detect the LUN?
  - Is the LUN ID in the `/kernel/drv/sd.conf` file?
  - Has the host rescanned to detect the LUN?
  - Run the `devfsadm` command.

NOTE: For additional troubleshooting recommendations, refer to the product documentation.

CIFS

The Common Internet File System (CIFS) is the default NAS protocol included with Microsoft Windows. The NetApp storage controller can participate in an Active Directory domain and present files via the CIFS protocol.

CONFIGURATION

The CIFS protocol is a licensed feature that must be enabled before it can be configured and used to present files for client access.

```
license add <licnum>
```

NOTE: If the controller was ordered with the CIFS license, then CIFS license will already be installed, and the CIFS setup wizard will start automatically when the system is booted.

The easiest way to configure the CIFS protocol is to run the CIFS setup wizard. The wizard prompts you for all aspects of the CIFS configuration, such as the NetBIOS host name, the authentication mode (for example, Active Directory), and the local administrator password. To start the wizard, you run the `cifs setup` command:

The choices for CIFS user authentication are:

- Active Directory
- NT Domain
- Windows Workgroup
- Non-Windows workgroup

NOTE: If you wish to configure Active Directory mode for user authentication, then the system time on the storage controller must be within five minutes of the system time on the Active Directory server. This requirement is inherited from the Kerberos protocol, which Active Directory uses for improved security. It is recommended that both systems be configured to use the same network time server, if one is available.

You can not only use an external authentication system such as Active Directory but you can also define local users and groups. It is recommended that a local administrator be defined—to be used if the Active Directory server is unavailable.

You can add domain users (such as a Domain Admin) to local groups. This ability can be useful when you are managing the storage controller as part of a larger enterprise. For example, to add the domain user “Steven” to the local Administrators group, run the following command:

```
useradmin domainuser add steven -g Administrators
```

For information about how to manage local users and groups, refer to Figure 6 and to the text related to Figure 6.

ADMINISTRATION

To create and manage CIFS shares, you must use the `cifs` command and the `shares` parameter. Here are some examples of the use of the `cifs` command and the `shares` parameter:

- List the existing shares

  ```
  cifs shares
  ```

- Create a share

  ```
  cifs shares –add <sharename> /vol/vol1
  ```

- Change a share

  ```
  cifs shares –change <sharename> -comment
  ```

Using the CIFS protocol, you can create shares to expose the following object types for user access:

- Volume

  ```
  /vol/vol1
  ```

- Qtree

  ```
  /vol/vol1/qtree1
  ```

- Directory

  ```
   /vol/vol1/dir1
  ```

It is sometimes desirable to set a quota on the amount of disk space that an individual user or group can consume via the CIFS (or NFS) protocol. There are several types of quotas and various options to enforce each type of quota.

- Quota targets
  - User, controls how much data the specified user can store
  - Group, controls how much data the specified group of users can store
  - Qtree, controls how much data can be stored in the specified qtree (similar to a directory)

    NOTE: The Administrator and root users are exempt from the all quota types except the qtree quota type.

  - Default, applies only if the specified quota is not defined for the current user or group
- Quota objects
  - Capacity, limits the amount of disk space that can be used (sets a maximum)
  - Files, limits on the number of files that can be stored (sets a maximum)
- Quota thresholds
  - Soft: When the specified threshold is reached, the controller sends only a Simple Network Time Protocol (SNMP) trap. The user's write succeeds.
  - Hard: When the specified threshold is reached, the controller prevents the user's attempt to write more data. The client displays a "filesystem full" error.

Quotas are configured and managed via either the `quota` command or NetApp System Manager (1.1 and later) interface. Both methods store the quota configuration in the /etc/quotas file on the controller's root volume.

- List the active quotas

  ```
  quota status
  ```

- Enable quotas

  ```
  quota on <volume_name>
  ```

A file system scan is required. For a large file system, the scan can require a significant amount of time.

- Resize a quota

  `quota resize <volume_name>`

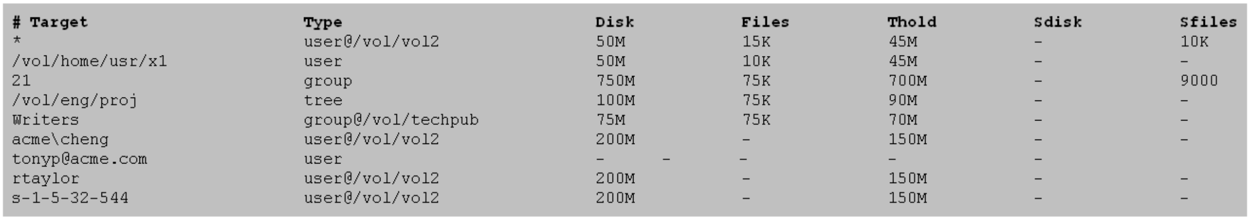
  When you modify the limits of a quota, a file system scan is not needed. However, the file system scan is not needed only if you modify the limits on an existing quota.

- Report on quota usage

  `quota report`

  A report that details the current file and space consumption for each user and group that is assigned a quota and for each qtree is printed.

Here is an example of an /etc/quotas configuration file:



```
# Target            Type                 Disk       Files     Thold     Sdisk     Sfiles
*                   user@/vol/vol2       50M        15K       45M       -         10K
/vol/home/usr/x1    user                 50M        10K       45M       -         -
21                  group                750M       75K       700M      -         9000
/vol/eng/proj       tree                 100M       75K       90M       -         -
Writers             group@/vol/techpub   75M        75K       70M       -         -
acme\cheng          user@/vol/vol2       200M       -         150M      -         -
tonyp@acme.com      user                 -     -    -         -         -
rtaylor             user@/vol/vol2       200M       -         150M      -         -
s-1-5-32-544        user@/vol/vol2       200M       -         150M      -         -
```

FIGURE 12 : AN EXAMPLE QUOTA CONFIGURATION FILE

NOTE: The second line, which begins with an asterisk (*), is a default quota.

PERFORMANCE

The FAS storage controller has numerous performance counters, statistics, and other metrics. Also, various CIFS-specific performance-analysis commands are available; for example:

- `cifs stat`, displays CIFS performance statistics
- `cifs top`, displays the most active CIFS clients, as determined by various criteria

NOTE: By default, CIFS statistics are cumulative for all clients. To collect statistics for individual clients, you must enable the `cifs.per_client_stats.enable` option.

For detailed information about the controller-performance commands that are not specific to CIFS (for example, `sysstat`, `stats`, and `statit`), refer to the Data ONTAP section.

Changing some CIFS performance parameters is disruptive. Such changes can be performed only when no clients are connected. Here is an example process:

1. `cifs terminate`

   This command halts the CIFS service and disconnects all clients.

2. `options cifs.neg_buf_size <value>`

   This command changes the buffer size.

3. `cifs restart`

   This command restarts the CIFS service (the clients can now reconnect).

One easy way to improve CIFS performance is to set Unicode mode on the volume that contains the shared files.

- `vol options <volume_name> convert_ucode on`
- `vol options <volume_name> create_ucode on`

Because the CIFS protocol always uses Unicode, if a volume is not configured for Unicode support, then the controller must continuously convert between Unicode mode and non-Unicode mode. Therefore, configuring a volume for Unicode support may increase performance.

SECURITY

To manage user access to CIFS shares, you use the `cifs` command and the `access` parameter. Here is an example of how user access rights are evaluated:

- You issue the command that identifies shares and access rights.

  ```
  cifs access <share> <user> <access rights>
  ```

  The list of share-level access rights are:
  - No Access
  - Read
  - Change
  - Full Control
- The system evaluates the user's security ID (SID), which is usually provided at login time by the Active Directory server, against the user's share permissions.

  Access to the share is either granted or denied.
- The system evaluates the user's SID against the file or directory permissions.

  On a file-by-file basis, access is either granted or denied.

If mandatory file locking is requested by the client, the CIFS protocol enforces it.

TROUBLESHOOTING

CIFS is a complex protocol that uses several TCP ports and interacts with various external systems for user authentication, Kerberos security (Active Directory), and host name resolution. Therefore, the CIFS protocol has numerous potential points of failure; but, remarkably, it is usually very reliable.

A full description of CIFS troubleshooting is outside the scope of this document. For detailed troubleshooting information, refer to the *ANCDA Boot Camp* training materials.

If you suspect that connectivity problems are the source of your CIFS problems, you can try the following commands:

- `ping`, provides a standard TCP connection test
- `testdc`, tests communication with the Active Directory server
- `ifstat`, displays a low-level view of network-interface performance data
- `netdiag`, analyzes the network protocol statistics to ensure correct operation and displays, as required, suggested remedial actions

Another potential issue is file access permissions. In this case, the access protocol is CIFS, but the containing volume or qtree may have a UNIX security style. This configuration, which is called "multiprotocol access," is discussed in the next section.

MULTIPROTOCOL

The NetApp storage controller can present files via the CIFS protocol and the NFS protocol simultaneously. Because the controller includes sophisticated mappings between Windows and UNIX user names and file system permissions, the operation is seamless.

CONFIGURATION

The only requirement for multiprotocol access is that CIFS access and NFS access be configured to the same file system. Of course, some unavoidable complexities arise, because Windows systems and UNIX systems use different security semantics.

Some security settings and user-name mappings do need to be configured. The mappings that do not need to be configured are discussed in the Security section.

ADMINISTRATION

The CIFS and NFS protocols are managed separately, even when they refer to the same file system. Therefore, the protocols should not be the source of any multiprotocol concern. For detailed information about the administration of the CIFS and NFS protocols, refer to the CIFS and NFS administration sections.

PERFORMANCE

Multiprotocol access should not be the source of any performance concern. For detailed performance information, refer to the CIFS and NFS performance sections.

SECURITY

The default security style for all new volumes is controlled by a WAFL option:

```
options wafl.default_security_style <value>
```

- Where **`<value> = unix`**

  All files and directories have UNIX permissions.

- Where **`<value> = ntfs`**

  All files and directories have New Technology File System (NTFS) permissions.

- Where **`<value> = mixed`**

  Each file and directory can have either UNIX or NTFS permissions (but not both at the same time).

NOTE: Because the use of mixed mode can complicate the troubleshooting of file-access problems, it is recommended that mixed mode be used only when it is needed to accommodate a particular requirement.

To manually set or view the security style for a volume or qtree, you use the following commands:

- qtree status, displays the list of volumes and qtrees and their security styles
- qtree security <path> [unix | ntfs | mixed], sets the security mode for a nominated volume or qtree

Although the security-style setting controls the underlying file system's security type, access via CIFS or NFS to the file data is controlled by the normal user-authorization process. The multiprotocol environment introduces additional complexity because the relationships between the security semantics of the users, the shares and exports, and the file system must be mapped. Consider the following example:

- Evaluate the user's SID or user ID against the share or export permissions.

  Access to the share or export is either granted or denied.

- Evaluate the user's SID or user ID against the file or directory permissions.

  If the ID and the permissions are of different types (for example, Windows and UNIX), then it may be necessary to map the user name to the correct type.

On a file-by-file basis, access is either granted or denied.

The following diagram identifies where user-name mapping may need to occur.

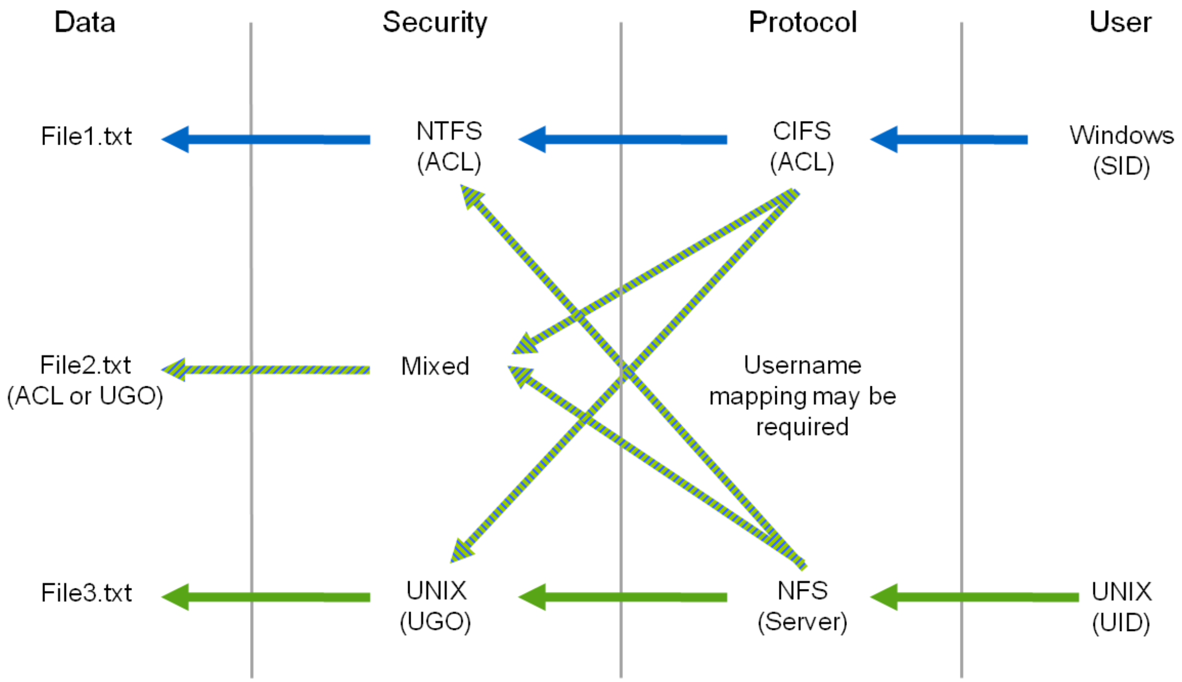


FIGURE 13: MULTIPROTOCOL ACCESS AND USER-NAME MAPPING

The user-name mapping is defined in the /etc/usermap.cfg file. The /etc/usermap.cfg file is a simple text file in the root volume of the storage controller. The process of mapping user names between Windows and UNIX contexts is reasonably straightforward:

- Automatic, if the Windows and UNIX user names match
- Specified (win_user = unix_user), if the user names are defined in /etc/username.cfg file
- Default, if the user names differ and there is no specific mapping. In this case, attempt to use the defined default UNIX or Windows user name (if any).

```
options wafl.default_nt_user <username>
```

```
options wafl.default_unix_user <username>
```

TROUBLESHOOTING

Most problems with multiprotocol access are caused by incorrect security styles or incorrect user-name mappings.

To identify how user-name mappings are being resolved, issue one of the following commands:

- wcc –u <unix_user>

  This command displays the UNIX to Windows user-name mapping.

- wcc –s <windows_user>

  This command displays the Windows to UNIX user-name mapping, for example:
  `Domain\Administrator => root`

If you are connecting as the Administrator or root user to a volume with a foreign security style, the easiest way to overcome an access problem may be to set the following options:

- `options wafl.nt_admin_priv_map_to_root on`
- `options cifs.nfs_root_ignore_acl on`

The `wafl` and `cifs` options grant superuser privileges on the foreign volume to Administrator and root users, respectively.

In some cases, due to variances in user populations and variances in CIFS and NFS file locking abilities, you may need to debug a file-access problem (for example, an NFS client can't open a file because it is locked by a CIFS client). You can use the `cifs sessions` command to list the clients that have active CIFS connections.

Another possible issue with mixed NFS and CIFS access is that NFS clients support symbolic links in the file system and CIFS clients generally do not support symbolic links. However, if you set the `cifs.symlinks.enable` option to `on` (the default value), then a CIFS client can successfully resolve any symbolic-link problem that was created by an NFS client.

NOTE: For detailed information about the interaction of CIFS clients with the various types of symbolic links, refer to the product documentation.

To resolve (better yet, to avoid) the simplest of all access problems, create both a CIFS share and an NFS export.

NFS

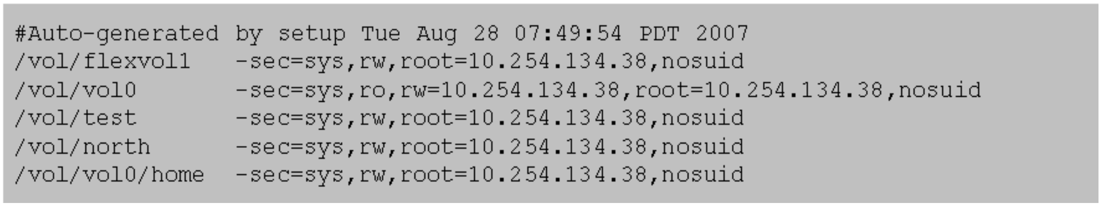
The Network File System (NFS) is the default NAS protocol that is included with all UNIX platforms. The NetApp storage controller can present files via the NFS protocol and can also participate in a Kerberos domain.

CONFIGURATION

The NFS protocol is a licensed feature that must be enabled before it can be configured and used to present files for NFS client access.

```
license add <licnum>
```

The NFS server configuration is described in the /etc/exports file. The file lists all NFS exports, specifies who can access the exports, and specifies privilege levels (for example, read-write or read-only).



```
#Auto-generated by setup Tue Aug 28 07:49:54 PDT 2007
/vol/flexvol1   -sec=sys,rw,root=10.254.134.38,nosuid
/vol/vol0       -sec=sys,ro,rw=10.254.134.38,root=10.254.134.38,nosuid
/vol/test       -sec=sys,rw,root=10.254.134.38,nosuid
/vol/north      -sec=sys,rw,root=10.254.134.38,nosuid
/vol/vol0/home  -sec=sys,rw,root=10.254.134.38,nosuid
```

FIGURE 14: AN EXAMPLE /ETC/EXPORTS FILE

NOTE: Any volume (or qtree and so on) that is to be exported is listed in the configuration file.

The /etc/exports file contains three types of information:

- Resource list
  - Exports are resources that are available to NFS clients.
  - Example: `/vol/flexvol1` identifies a resource that is to be exported.
- Identification
  - NSF clients can be identified by their host names, DNS subdomains, IP addresses, IP subnets, and so on:
  - Example: `10.254.134.38`
- Authorization
  - Exports specify access permissions for NFS clients.
  - Example: `rw` and `nosuid` specify access permissions.

ADMINISTRATION

To perform NFS server administration tasks, you use the `exportfs` command and the /etc/exports file. Here are some examples of their use:

- List the current exports (in memory)

  ```
  exportfs
  ```

- List the persistent exports (available after a reboot)

  ```
  rdfile /etc/exports
  ```

- Create an export (in memory)

  ```
  exportfs –i –o rw=host1 /vol/vol1
  ```

- Create a persistent export (available after a reboot)

  ```
  wrfile –a /etc/exports
  /vol/vol1 –rw=host1
  <CNTL>-C

  exportfs -a
  ```

With the NFS protocol, you can create exports to expose the following object types for user access:

- Volume

  ```
  /vol/vol1
  ```

- Qtree

  ```
  /vol/vol1/qtree1
  ```

- Directory

  ```
  /vol/vol1/dir1
  ```

- File

  ```
  /vol/vol1/file1.iso
  ```

NOTE: Unlike most UNIX variants, the FAS controller can successfully export nested directories (and, thus, can export ancestors and descendants). For example, both `/vol/vol1` and `/vol/vol1/qtree1` can be exported, and the NFS client must satisfy only the access controls that are associated with the mount point that was initially accessed.

For a description of file system quotas, refer to the CIFS configuration section.

PERFORMANCE

The FAS storage controller has numerous performance counters, statistics, and other metrics. Various NFS-specific performance-analysis commands are available; for example:

- `nfsstat`, displays NFS performance statistics
- `nfs_hist`, an advanced mode command that displays NFS delay time distributions (that is, the number of I/O operations per millisecond grouping)
- netapp-top.pl, a Perl script that lists the most active NFS clients. The script is run on a client system. The script can be downloaded from the NOW online support and services site at the following location: http://now.netapp.com/NOW/download/tools/ntaptop/.

NOTE: By default, the NFS statistics that are reported are cumulative for all clients. To collect statistics for individual clients, you must enable the `nfs.per_client_stats.enable` option.

For detailed information about the controller-performance commands that are not specific to NFS (for example, `sysstat`, `stats`, and `statit`), refer to the Data ONTAP section.

Performance testing is a complex topic that is beyond the scope of this document. Nevertheless, you can perform some very basic performance analysis by using the following procedure (from the NFS client):

- To write traffic, run the `time mkfile` command
- To read traffic, run the `time dd` command
- To read/write traffic, run the `time cp` command

NOTE: If you are concerned about repeatable performance testing, then you should investigate utilities such as iometer, iozone, and bonnie++ and the NetApp sio tool.

SECURITY

Traditionally, security has been seen as a weak spot for the NFS protocol, but recent versions of NFS support very strong security. The traditional security model is called “AUTH_SYS,” and the newer model is called “Kerberos.” Here is a summary of the differences between the two security models:

- AUTH_SYS
  - User authentication is performed on the remote NFS client (which is typically a UNIX server). This scenario implies that the authentication process on the NFS client is trusted and that the NFS client is not an impostor.
  - No additional authentication, data-integrity evaluation, or data encryption is performed.

- Kerberos
  - User authentication is performed on the remote NFS client, and Kerberos authenticates that the NFS client is genuine.
  - There are three levels of Kerberos security.
    - krb5: Authentication occurs with each NFS request and response.
    - krb5i: Authentication occurs with each NFS request and response, and integrity checking is performed, to verify that requests and responses have not been tampered with.
    - krb5p: Authentication occurs with each NFS request, integrity checking is performed, and data encryption is performed on each request and response.

NOTE: If you wish to configure Kerberos mode for user authentication, then the system time on the storage controller must be within five minutes of the system time on the Kerberos server. This requirement is inherited from the Kerberos protocol. It is recommended that both systems be configured to use the same network time server, if one is available.

Before a user can access an export, the NFS client (remote UNIX server) must be able to mount the export. Then, the user’s access to the shared files in the export is evaluated against the user’s UNIX user ID and group ID (UID and GID), which is usually provided at login time by the NFS client. This is a two step process; for example:

- Evaluate the server's host name against the export permissions

  Access is either granted or denied (to mount the export, r/w or r/o).

- Evaluate the user's UID/GID against the file or directory permissions

  Access is either granted or denied (on a file-by-file basis).

The NFSv2 and NFSv3 protocols use advisory file locking, and the NFSv4 protocol enforces mandatory file locking, if it is requested by the client.

TROUBLESHOOTING

NFS is a complex protocol that uses several TCP ports and interacts with various external systems for user authentication, Kerberos security, and host name resolution. As such, NFS has numerous potential points of failure, but, remarkably, it is usually very reliable.

A full description of NFS troubleshooting is outside the scope of this document. For detailed troubleshooting information, refer to the *ANCDA Boot Camp* training materials.

If you suspect that RPC problems are the source of your NFS problems, you can try the following actions:

- Verify that RCP is enabled
- Verify that NFS daemons are running
- Verify that mount points exist

If you are having problems mounting an NFS export, you can try the following commands:

- `showmount –e`

  This command, which is run from the NFS client, lists the exports that are available on the NFS server.

- `nfsstat –d`

  This command, which is run from the controller, displays low-level statistics that are useful in debugging a mount problem.

If you are experiencing "stale NFS handle" errors, you can try the following actions:

- Check the /etc/fstab file on the host for errors.
- Check connectivity between the two systems by using the `ping` command.
- List the exports that are available on the NFS server by running the `showmount –e` command on the NFS client
- Check the controller's /etc/exports file.
- Check the controller's current exports in memory by running the `exportfs` command.

EXAM NS0-154: DATA PROTECTION CONCEPTS

As a NetApp Certified Data Management Administrator, you can implement an active-active controller configuration and use SyncMirror software to ensure continuous data availability and rapid recovery of data and use the SnapMirror®, SnapRestore®, and SnapVault® products to manage and protect data.

SKILLS TESTED

- Set up and maintain Snapshot™ copies
- Configure and administer SnapRestore technology
- Configure and administer asynchronous SnapMirror product
- Configure and administer synchronous SnapMirror product
- Configure and administer Open Systems SnapVault application
- Configure and administer Operations Manager application
- Configure and administer SnapLock® technology (not applicable in Data ONTAP 8.0 7-Mode or the corresponding NS0-154 exam)
- Analyze and resolve data protection problems
- Implement high-availability (active-active) controller configuration (including SyncMirror)

RECOMMENDED COURSES

- Instructor-led: *Data ONTAP 8.0 7-Mode Administration*
- Instructor-led: *NetApp Protection Software Administration*
- Instructor-led: *Accelerated NCDA Boot Camp Data ONTAP 8.0 7-Mode*

RELATED COURSES

- Web-based: *High Availability on Data ONTAP 8.0 7-Mode*
- Web-based: *Planning and Implementing MetroCluster on Data ONTAP 8.0 7-Mode*
- Web-based: *Implementing SyncMirror on Data ONTAP8.0 7-Mode*

EXAM PREPARATION

This section describes various NetApp FAS learning points that are relevant to the NS0-163 and NS0-154 exams. These learning points focus on data protection concepts. However, the section is not limited to the exam topics. Rather, it also summarizes information about a range of NetApp technologies.

Figure 14 highlights the main subjects covered in the exam (white text) and the range of topics covered within each subject (black text).

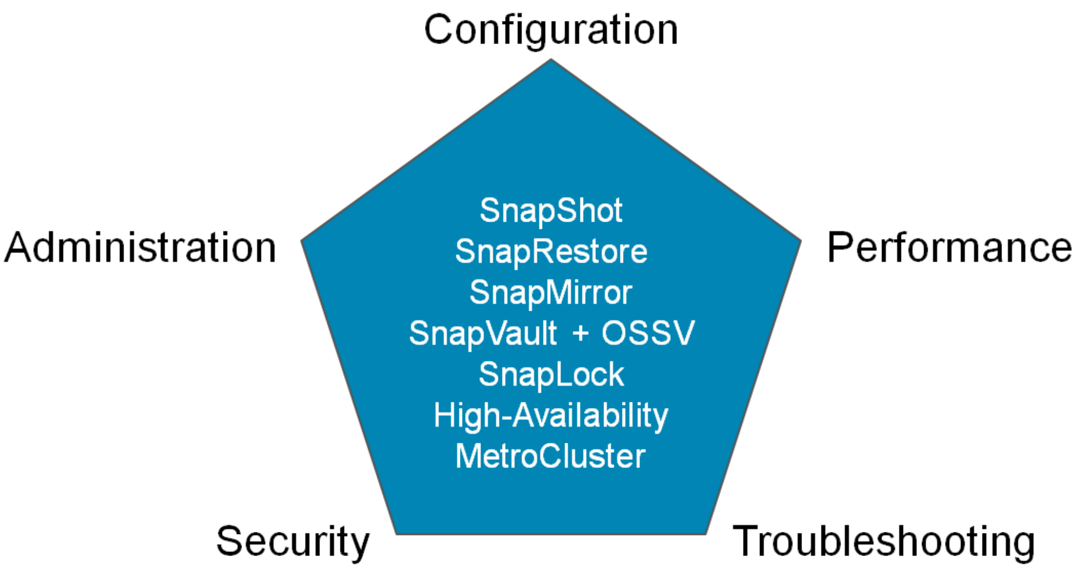


FIGURE 15 : TOPICS COVERED IN THE NS0-163 AND NS0-154 EXAMS

SNAPSHOT TECHNOLOGY

A Snapshot copy is a read-only image of a volume or an aggregate. The copy captures the state of the file system at a point in time. Many Snapshot copies may be kept online or vaulted to another system, to be used for rapid data recovery, as required.

CONFIGURATION

The Snapshot capability of the FAS storage controller is a native capability that is provided by the WAFL file system layer. Both SAN and NAS data can be captured in a Snapshot copy.

NetApp Snapshot technology is particularly efficient, providing for instant creation of Snapshot copies with near-zero capacity overhead.

This efficiency is possible because, like most UNIX file systems, the WAFL file system uses inodes to reference the data blocks on the disk. And, a Snapshot copy is a root inode that references the data blocks on the disk. The data blocks that are referenced by a Snapshot copy are locked against overwriting, so any update to the active file system (AFS) is written to other locations on the disk.

Refer to Figure 16 for an example of how the Snapshot process occurs.

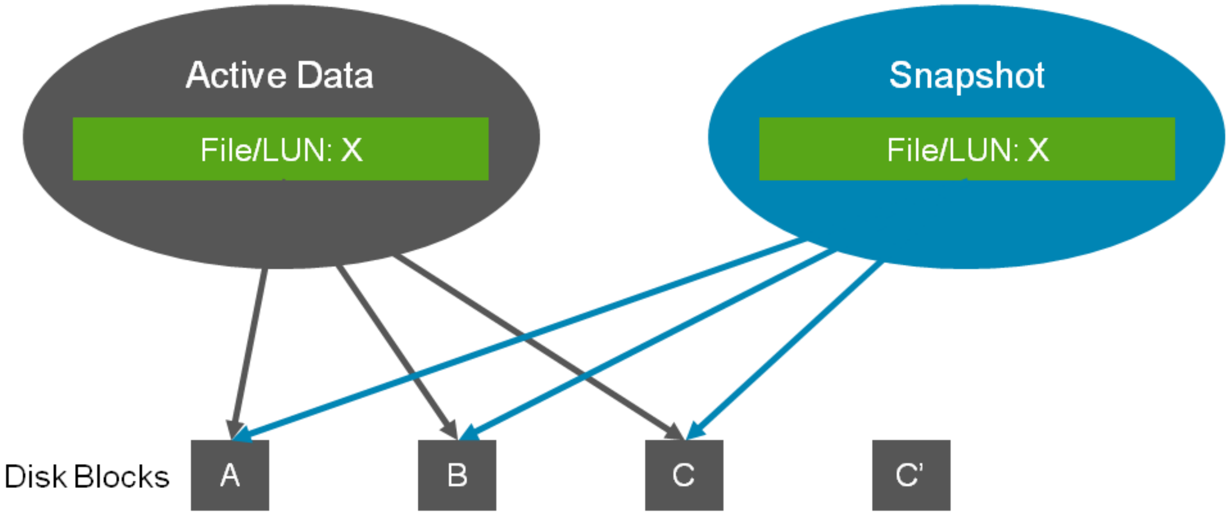


FIGURE 16: THE PROCESS OF A SNAPSHOT COPY LOCKING SOME BLOCKS IN THE AFS

NOTE: Each volume can retain up to 255 Snapshot copies.

When you create a volume, a default Snapshot schedule is created. Initially, Snapshot copies are created according to the schedule's default settings. However, you can modify or disable the default settings to satisfy your local backup requirements.

- List the default Snapshot schedule.
  - The command: `snap sched <vol>`
  - The output:

    ```
    Volume
    <vol>: 0 2 6@8,12,16,20
    ```

    The default schedule creates four hourly Snapshot copies (at 8:00, 12:00, 16:00, and 20:00) and retains 6 total, a daily Snapshot copy (at 24:00 Monday through Saturday and Sunday if a weekly Snapshot copy is not taken), retaining 2 at a time and zero weekly Snapshot copies (if created, these would occur at 24:00 on Sunday).

- Modify the Snapshot schedule by running a command similar to the following:

  ```
  snap sched <vol> weekly nightly hourly@<time>
  ```

  ```
  snap sched <vol> 2 7 6@6,9,12,15,18,21
  ```

- Disable the Snapshot schedule by running one of the following commands.

```
snap sched <vol> 0 0 0
```

```
vol options <vol> nosnap on
```

NOTE: On volumes that contain LUNs, you normally disable the controller initiated Snapshot copies because the consistency of the file system in the LUN can be guaranteed only by the host that accesses the LUN. You should then use a tool such as SnapDrive to initiate the Snapshot copies from the host.

A percentage of every new volume and every new aggregate is reserved for storing Snapshot data. is the reserved space is known as the “Snapshot reserve.” The default reserve is 5% for aggregates and 20% for volumes. You can modify the default values by running the `snap reserve` command.

```
snap reserve <vol> <percentage>
```

Refer to Figure 17 to identify where the Snapshot reserve values apply.

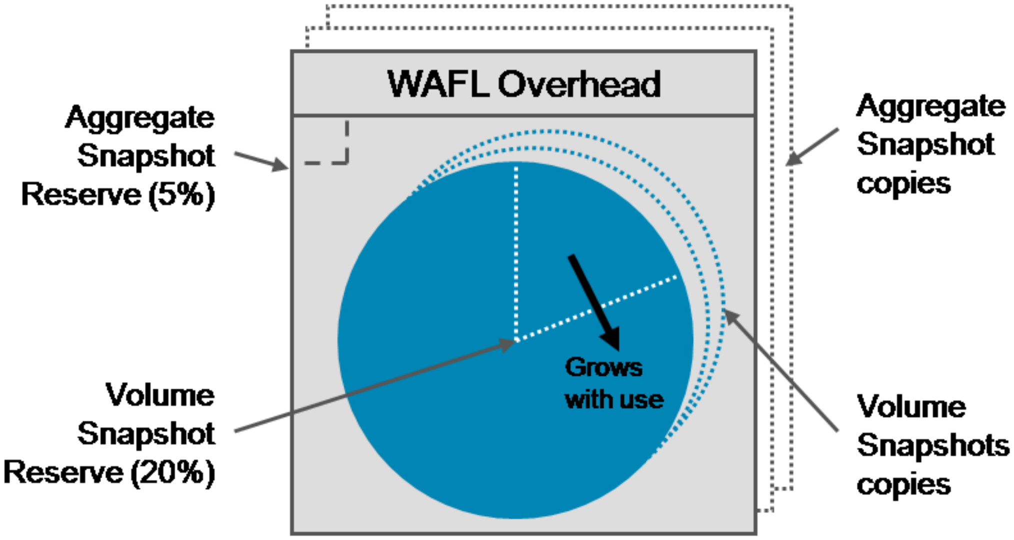

WAFL Overhead
Aggregate Snapshot Reserve (5%)
Aggregate Snapshot copies
Volume Snapshot Reserve (20%)
Grows with use
Volume Snapshots copies

FIGURE 17: AGGREGATE AND VOLUME SNAPSHOT RESERVES

NOTE: The value for the volume Snapshot reserve is the *minimum* amount of space that is reserved for Snapshot data. The Snapshot copies can consume more space than the initial reserve value specifies.

ADMINISTRATION

Almost all Snapshot management is performed by running the `snap` command; for example:

- `snap list`, shows the currently retained Snapshot copies
- `snap create <vol_name> <snap_name>,` create a volume Snapshot copy (If you want an aggregate level Snapshot copy, then specify `–A` .
- `snap delete <vol_name> <snap_name>,` deletes the specified Snapshot copy and makes its disk space available

NOTE: Some special types of Snapshot copies (for example, Snapshot copies created with SnapMirror and SnapVault software) are created and managed by the storage controller and should not be interfered with.

Snapshot copies that are created with SnapDrive and SnapManager software should not be managed from the storage controller. These Snapshot copies are created, retained, and deleted under the control of the host and application integration agents. Because the copies contain consistent backup images that are being retained by schedules and policies on the agents, they should not be deleted manually.

PERFORMANCE

Typically, because Snapshot technology is very efficient, the creation, retention, and deletion of Snapshot copies make no significant impact on performance.

For information about performance, refer to the SnapRestore performance section.

SECURITY

By definition, a Snapshot copy is a read-only view of the state of the file system at the time that the copy was created. Therefore, the contents of the copy cannot be modified by end users.

User access to the data in a Snapshot copy is controlled by the file system security settings (for example, NTFS ACLs) that were in place when the Snapshot copy was created.

NOTE: If the security style of the volume changes after the Snapshot copy is created, then the users may not be able to access the file system view in the Snapshot directory (unless their user-name mapping is configured to allow them to access the foreign security style). This problem arises because the previous security settings are preserved in the Snapshot view.

The storage administrator can configure the visibility of the Snapshot directory (for NAS clients). The following commands either enable or disable client access to the Snapshot directory:

- Per volume

  The default volume settings *do* allow the Snapshot directory to be seen by the NAS protocols. Use the following command to *disable* the Snapshot directory per volume:

  ```
  vol options <volume_name> nosnapdir on
  ```

- For CIFS access

  The default CIFS settings *do not* allow the Snapshot directory to be seen by CIFS clients. Use the following command to *enable* the Snapshot directory.

  ```
  options cifs.show_snapshot on
  ```

- For NFS access

  The default NFS settings *do* allow the Snapshot directory to be seen by NFS clients. Use the following command to *disable* the Snapshot directory per volume.

  ```
  options nfs.hide_snapshot on
  ```

TROUBLESHOOTING

Usually, there are no problems with creating Snapshot copies *per se*, but complications can arise. These complications are usually a result of incorrect scheduling or lack of disk space.

- Inconsistent file system

  If you create a Snapshot copy of a volume that contains a LUN, the file system in the LUN may or may not be in a consistent state. If the file system is corrupt, you may not be able to use the LUN Snapshot copy to recover data. The consistency of the file system in the LUN can be guaranteed only by the host that accesses the LUN. You should use a tool such as SnapDrive to initiate the LUN Snapshot copies from the host (so the tool can flush the local file system buffers to disk).

- Hosts, LUNs, and space within controller volumes

  The host that accesses a LUN assumes that it has exclusive control over the contents of the LUN and the available free space. However, as Snapshot copies are created, more and more of the space in the containing volume is consumed. If the Snapshot copies are not managed correctly, they eventually consume all of the space in the volume. If all of the space in a volume is consumed and the host attempts to write to a LUN within the volume, an “out of space” error occurs (because the host assumed that space was available). The controller then takes the LUN offline in an attempt to prevent data corruption. Refer to Figure 8 and to the text that accompanies Figure 8 for descriptions of Fractional Reserve and of the Volume AutoGrow and Snapshot Autodelete options and for an explanation of how Fractional Reserve, Volume AutoGrow, and Snapshot Autodelete ensure adequate free space and guarantee LUN availability.

SNAPRESTORE TECHNOLOGY

The SnapRestore feature enables you to use Snapshot copies to recover data quickly. Entire volumes, individual files, and LUNs can be restored in seconds, regardless of the size of the data.

CONFIGURATION

SnapRestore is a licensed feature that must be enabled before it can be configured and used.

```
license add <licnum>
```

NOTE: The SnapRestore feature is licensed system-wide. Therefore, the feature cannot be enabled or disabled at a per-volume level.

The only prerequisite for using the SnapRestore feature (other than licensing) is the existence of Snapshot copies. The SnapRestore feature restores data from Snapshot copies. Snapshot copies that you have not created or retained cannot be used to restore data.

ADMINISTRATION

The SnapRestore feature is an extension of the `snap` command. The command, when used with the `restore` parameter, can restore an entire volume or an individual file or LUN from a Snapshot copy.

- `snap restore -t vol -s <snap_name> <vol_name>`
  - This command reverts the *entire volume* back to exactly how it was when the Snapshot copy was created.
  - Be aware that all subsequent Snapshot copies are deleted.
- `snap restore -t file -s <snap_name> <file_name>`
  - This command reverts an *individual file* back to exactly how it was when the Snapshot copy was created.
  - To recover to a file name or directory location other than the original file name or directory location, add the `-r <new_path_and_file_name>` parameter.

NOTE: You can run the SnapRestore command (volume or file) only on a volume that is online.

The SnapRestore feature recovers *only* volume and file content. It does not recover the following settings:

- Snapshot copies schedule
- Volume option settings
- RAID group size
- Maximum number of files per volume

NOTE: The volume SnapRestore command reverts the entire active file system (AFS) back to the point at which the Snapshot copy was created. All Snapshot copies that were created between the time that the Snapshot backup copy was created and the time that the Snapshot backup copy was used to restore the AFS are deleted. When using the SnapRestore feature, be very careful! *You cannot back out of your changes*.

PERFORMANCE

Using the SnapRestore feature to restore one file may impact subsequent snapshot delete performance. Before a Snapshot copy is deleted, the active maps across all Snapshot copies must be checked for active blocks that are related to the restored file. This performance impact may be visible to the hosts that access the controller, depending on the workload and scheduling.

SECURITY

After you perform a SnapRestore operation (at the volume or file level), the file system metadata, such as security settings and timestamps, are reverted to exactly what they were when the Snapshot copy that was used to perform the restore was created.

- Security settings: The security settings of the file have been reverted to their earlier values. If you suspect that the revision may have created a problem, you should review the security settings.

- File timestamps: After reversion, the file timestamps are invalid for incremental backups. If you are using a third-party backup tool, so you should run a full backup.

- Virus scanning: If a virus-infected file was captured in the Snapshot copy, it is restored in its infected state (whether or not it was cleaned after the Snapshot copy was created). You should schedule a virus scan on any recovered file or volume.

TROUBLESHOOTING

Because the volume remains online and writeable during the SnapRestore activity, there is always the possibility that users may access files on the volume as the restore process is in progress. This overlap can cause file corruption and can generate NFS errors such as “stale file handle.” There are several methods of avoiding or correcting such issues:

- Disconnect the users before you begin the SnapRestore operation
- Have the users re-open files that might present a problem

The SnapRestore destination volume cannot be a SnapMirror destination volume. If you want to restore a SnapMirror destination volume, then you should use the FlexClone® feature, which must be licensed, to link the destination’s Snapshot copy to a new writable volume.

SNAPMIRROR PRODUCTS

The SnapMirror product family enables you to replicate data from one volume to another volume or, typically, from a local controller to a remote controller. Thus, SnapMirror products provide a consistent, recoverable, offsite disaster-recovery capability.

CONFIGURATION

SnapMirror is a licensed feature that must be enabled before it can be configured and used. The SnapMirror feature actually has two licenses. The first is a for-charge license that provides the asynchronous replication capability, and the second is a no-charge license that provides the synchronous and semi-synchronous capabilities. The no-charge license is available only if the for-charge license is purchased.

- First, license the SnapMirror Async function.

  ```
  license add <licnum>
  ```

- Then, license the SnapMirror Sync and SnapMirror Semi-Sync functions (if required).

  ```
  license add <licum>
  ```

  The no-charge SnapMirror Sync license code is printed in the Data ONTAP Data Protection Online Backup and Recovery Guide.

  NOTE: Some older NetApp controllers (FAS820 and prior) cannot support the SnapMirror Sync function.

NOTE: The SnapMirror feature must be licensed on both the source and the destination systems (for example, production and disaster recovery systems).

By default, the SnapMirror feature uses a TCP connection to send the replication data between the two controllers. The TCP connection is usually over an Ethernet or TCP WAN link and is usually the most cost-effective transport. However, customers with access to inter-site Fibre connections can install the model X1024 FC adapter and replicate across the optical media.

The second step (after licensing) in configuring a volume SnapMirror relationship is to create the destination volume. The source and destination volume may be located on the same controller (for data migration) or on different controllers (for disaster recovery).

- To create a restricted mode destination volume, run the following commands on the destination system:

  `vol create <vol_name>` (with parameters to suit)

  ```
  vol restrict <vol_name>
  ```

- To check the volume's status and size, run the `vol status –b` command. The volume must be online but in a restricted state to initialize a volume SnapMirror relationship.

NOTE: For a qtree SnapMirror relationship, the destination volume remains in an online and writeable state (not restricted) and the destination qtrees are created automatically when the baseline transfer is performed.

You need to know what the requirements and states of the source and destination volumes are and to understand how the requirements and states of volume SnapMirror relationships and qtree SnapMirror relationships can differ.

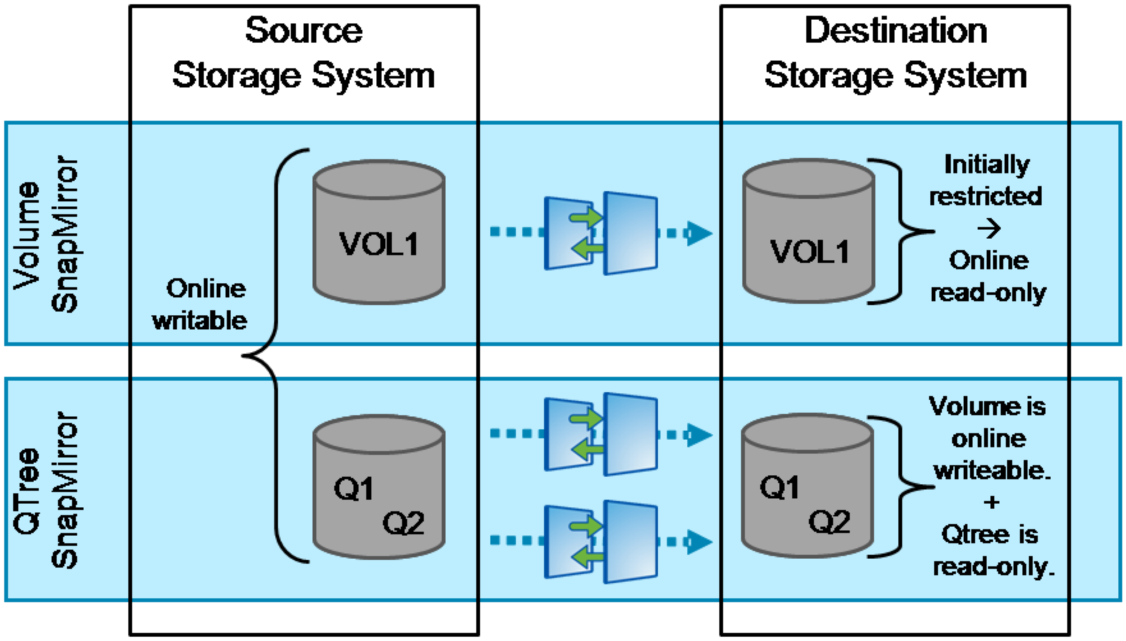


FIGURE 18 :SNAPMIRROR VOLUME AND QTREE CONFIGURATION

NOTE: If a volume SnapMirror relationship is stopped (broken or released), the destination volume changes to a writable state, and the `fs_size_fixed` parameter is enabled on the volume. These actions prevent the inadvertent resizing of the destination volume. Resizing can cause problems when (if) the relationship is resynchronized.

Before you can enable a SnapMirror relationship, you must configure the SnapMirror access control between the primary and secondary storage controllers. For a description of the required settings, refer to the Security section.

After the source and destination volumes are defined, you can configure the SnapMirror relationship. As you configure the relationship, you also perform the initial baseline transfer, copying all of the data from the source volume to the destination volume.

```
snapmirror initialize –S src:vol1 dst:vol2
```

When the baseline transfer is completed, the destination volume is an exact replica of the source volume (at that point in time).

Next you must configure the ongoing replication relationship. This relationship controls the mode and/or the schedule for replication of the changed data from the source volume to the destination volume. The SnapMirror replication parameters are defined in the snapmirror.conf file, as shown in Figure 19:

```
# Source          Destination       Options       Mode/Schedule
src:/vol/vol1/q1  dst:/vol/vol1/q1  -             15 * * *
src:vol2          dst:vol2          -             10 8,20 * *
src:/vol/vol3     dst:/vol/vol3     -             sync
src:vol4          dst:vol4          -             semi-sync
```

FIGURE 19: EXAMPLE SNAPMIRROR.CONF FILE

NOTE: The snapmirror.conf file is configured on the destination controller.

As shown in Figure 19, the SnapMirror relationship can operate in any of three modes, performing asynchronous, synchronous, or semi-synchronous replication.

- Asynchronous
  - Snapshot copies are replicated from a source volume or qtree to a destination volume or qtree.
  - The host receives acknowledgment after the write is committed to the source volume,
  - Block-level, incremental updates to the destination volume are based on schedules.
  - The following is a VSM example:
    ```
    src:vol2 dst:vol2 - 10 8,20 * *
    ```
  - The following is a QSM example:
    ```
    src:/vol/vol1/q1 dst:/vol/vol1/q1 – 15 * * *
    ```
- Synchronous
  - Writes are replicated from the source volume to the destination volume at the *same time* that they are written to the source volume.
  - The host receives acknowledgment only after the write is committed to both the source and destination volumes.
  - The following is an example command:
    ```
    src:/vol/vol1/q1 dst:/vol/vol1/q1 – sync
    ```
- Semi-synchronous
  - Writes are replicated from a source volume or qtree to a destination volume or qtree with *minimal delay*.
  - The host receives acknowledgment after the write is committed to the source volume.
  - Performance with minimal delay minimizes the performance impact on the host system.
  - The following is an example of the previously used syntax:
    ```
    src:vol1 dst:vol1 outstanding=5s sync
    ```
  - The following is an example of the current syntax:
    ```
    src:vol1 dst:vol1 – semi-sync
    ```

NOTE: For descriptions of the various replication options (such as schedule definitions or throughput throttling), refer to the product documentation.

It is possible to configure the SnapMirror feature to use two redundant data paths for replication traffic. The paths can be TCP or FC connections or a mixture of TCP and FC connections. The paths are configured in the snapmirror.conf file. The following key words are used.

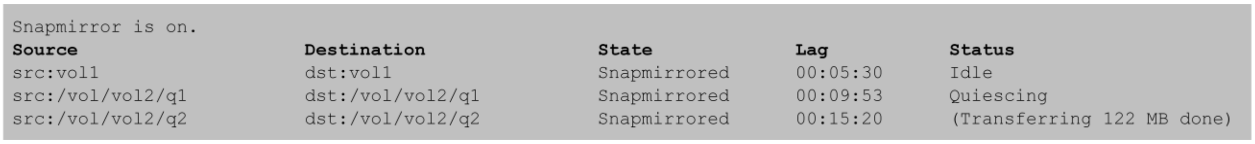
- Multiplexing: Both paths are used at the same time for load balancing.
- Failover: The first path that is specified is active. The second path is in standby mode and becomes active only if the first path fails.

NOTE: Editing the /etc/snapmirror.conf file on the destination causes an in-sync relationship to fall temporarily out-of-sync.

ADMINISTRATION

A SnapMirror relationship can be administered from either the source or the destination system, although some functions are available only on their respective systems.

You use the `snapmirror status` command to display the state of the currently defined SnapMirror relationships, as shown in Figure 20:



```
Snapmirror is on.
Source               Destination          State          Lag        Status
src:vol1             dst:vol1             Snapmirrored   00:05:30   Idle
src:/vol/vol2/q1     dst:/vol/vol2/q1     Snapmirrored   00:09:53   Quiescing
src:/vol/vol2/q2     dst:/vol/vol2/q2     Snapmirrored   00:15:20   (Transferring 122 MB done)
```

FIGURE 20: EXAMPLE OF SNAPMIRROR STATUS OUTPUT

The Lag column identifies the amount of time that has elapsed since the last successful replication of the Snapshot source-volume copy that is managed by the SnapMirror relationship.

The `snapmirror` command is used to manage all aspects of the SnapMirror relationship, such as suspending or restarting the replication or destroying the relationship. The following are examples of common `snapmirror` functions:

- `snapmirror quiesce <dst_vol>`
    - The command is executed on the *destination* system. The command temporarily pauses the replication. The destination volume remains read-only.
    - The relationship is still defined and can be resumed.
- `snapmirror resume <dst_vol>`
    - The command is executed on the *destination* system.
    - The command resumes the volume replication.
- `snapmirror break <dst_vol>`
    - The command is executed on the *destination* system. The command stops the replication and converts the destination volume to a writable state.
    - The relationship is still defined and can be resynchronized.
- `snapmirror resync <hostname:vol>`
    - The command identifies the most recently created Snapshot copy that is common to the source and destination volumes and that is managed by a SnapMirror relationship and re-synchronizes the data between the source and destination volumes.
    - The *direction* of synchronization is determined by whether the command was executed on the source or destination volume. The synchronization overwrites the new data on the controller on which the command was executed (bringing the execution-controller volume back into sync with the opposite volume).
    - If the command is executed on the *destination* system, then the relationship continues in its original direction (source → destination)
    - However, if the command is executed on the *source* system, then the relationship reverses its original direction (destination → source).
- `snapmirror release <src_vol> <dst_hostname:dst_vol>`
    - The command is executed on the *source* system. The command stops the replication and converts the destination volume to a writable state.
    - The relationship is deleted and cannot be restarted.
- `snapmirror update <dst_vol>`
    - The command is executed on the *destination* system.
    - The command performs an immediate update from the source volume to the destination volume.

The process of capturing consistent Snapshot copies on the source volume and then transferring the copies to the destination system varies, depending on your application's capabilities, the use of SnapDrive and SnapManager software, the replication mode, and the intended result. The following is one example of a process to create a consistent Snapshot copy at the destination of a qtree SnapMirror relationship:

1. Make the source volume consistent on disk.
   - Halt the application.
   - Flush the file system buffers.
2. *Quiesce* the SnapMirror relationship.
3. Create a Snapshot copy of the destination volume.
4. *Resume* the SnapMirror relationship.

NOTE: In environments that use SnapManager software, the Snapshot copy and replication process is usually automated via the SnapManager utility. In this case, the process can be performed with no disruption to the application.

PERFORMANCE

One of the challenges in a new SnapMirror configuration is the transfer of the baseline copy from the source to the destination system. Although the WAN connection may be adequate to handle the incremental synchronization traffic, it may not be adequate to complete the baseline transfer in a timely manner. In this case, you might consider using the *SnapMirror to Tape* function. This method can use physical tape media to perform the initial baseline transfer. In Data ONTAP 8.0 7-Mode, this functionality is now supported with the new `smtape` commands.

After the initial baseline transfer is completed, the incremental synchronization occurs. The initial baseline transfer is usually constrained by the bandwidth of the connection, and the incremental synchronization is usually constrained by the latency of the connection.

The appropriate choice of SnapMirror mode (synchronous, semi-synchronous, or asynchronous) is often driven by the latency of the WAN connection. Because latency increases over distance, latency effectively limits the synchronous mode to a range of less than 100 km. If you require a “sync-like” replication feature beyond 100 km or want to reduce the performance impact on the source system, then you should consider using the semi-synchronous mode.

In contrast, the asynchronous mode uses scheduled replication and is not affected by connection latency. One way to improve asynchronous performance is to increase the interval between the replication times. This increase allows for “file system churn.” Data is rewritten throughout the day, but only the latest version is included in the less frequent replication schedules.

In contrast to flexible volumes, the physical characteristics of traditional volumes affect SnapMirror performance. When you use traditional volumes, for best SnapMirror performance, you should configure the source and destination volumes with the same RAID size, RAID group size, and number of RAID groups.

The visibility_interval parameter controls the *apparent* performance of the SnapMirror synchronization. The parameter controls the view of the data on the destination system. Even after the data is received, the destination file-system view is not updated until the visibility interval elapses. The default visibility interval is three minutes, with a minimum setting of 30 seconds. Reducing the internal is not recommended because deviation from the default value can have a detrimental impact on controller performance.

NOTE: It is possible to throttle the SnapMirror traffic so as to reduce its impact on other applications that use the WAN connection.

SECURITY

Before you can enable the replication relationship, you must configure the SnapMirror access control between the source and destination storage controllers.

The source controller needs to grant access to the destination controller so that the destination controller can “pull” updates from the source. And the destination controller needs to grant access to the source controller so that the replication relationship can be reversed after a disaster event is resolved (synchronizing back from the disaster recovery site to the production site).

There are two ways to configure SnapMirror access control on the storage controllers:

- Method 1

  `options snapmirror.access host=legacy`

  – Edit the /etc/snapmirror.allow file.

  – Add the other storage controller’s host name.

- Method 2

  `options snapmirror.access host=<other controller>`

  – By default, this option is set to the keyword “`legacy`.” Use of the keyword causes the system to refer to the snapmirror.allow file for access control.

  – Alternatively, you can set the option to `host=<hostname>`, to enable the SnapMirror relationship to be accessed from the remote controller.

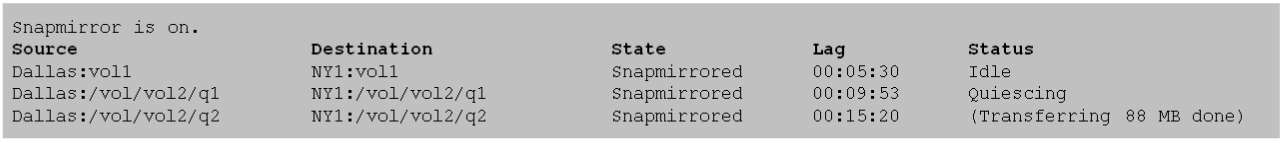
NOTE: Method 2 is the preferred way to enable the remote access.

The traffic between the source and destination controllers is not encrypted. In a security-conscious environment, it may be necessary to implement some type of network-level encryption for the replication traffic (for example, to use the NetApp DataFort™ encryption devices).

The DataFort security system is designed to encrypt data-at-rest, not data-in-transit. An encrypting Ethernet switch is used to encrypt data-at-rest.

TROUBLESHOOTING

Comprehensive logging of all SnapMirror activity is enabled by default. The log file is saved to the **/etc/log/snapmirror.[0-5]** file(s). The log can be disabled by executing the following command: `options snapmirror.log.enable [on|off].` The `snapmirror status` command displays the current status of all SnapMirror relationships. Some status information is available only on the destination controller.



```
Snapmirror is on.
Source               Destination          State          Lag        Status
Dallas:vol1          NY1:vol1             Snapmirrored   00:05:30   Idle
Dallas:/vol/vol2/q1  NY1:/vol/vol2/q1     Snapmirrored   00:09:53   Quiescing
Dallas:/vol/vol2/q2  NY1:/vol/vol2/q2     Snapmirrored   00:15:20   (Transferring 88 MB done)
```

FIGURE 21: SAMPLE OUTPUT OF SNAPMIRROR STATUS

A SnapMirror relationship passes through several defined stages as it initializes the baseline transfer (level-0), synchronizes data (level-1), and possibly reestablishes a broken mirror. The details of the process of troubleshooting and rectification are determined by the stage that the relationship was in when the failure occurred. For example, if communications failed during the initial baseline transfer, then the destination is incomplete. In this case, you must rerun the initialization, rather than trying to re-establish synchronization to the incomplete mirror.

SNAPVAULT FEATURE

The SnapVault feature enables you to create and archive Snapshot copies from one volume to another volume or, typically, from a local controller to a remote controller. The feature provides a consistent, recoverable, offsite, long-term backup and archive capability.

CONFIGURATION

SnapVault is a licensed feature that must be enabled before it can be configured and used. The SnapVault feature has two licenses. One license is for the primary controller (the backup source), and the other license is for the secondary controller (the archive destination).

- License the primary controller (sv_ontap_pri).

  ```
  license add <licnum>
  ```

- License the secondary controller (sv_ontap_sec).

  ```
  license add <licnum>
  ```

NOTE: The two licenses enable different functionalities, and the correct license must be enabled on the appropriate controller (for example, the production or the disaster recovery controller).

The second step in configuring a SnapVault relationship (after licensing) is to create the destination volume. Typically, you create the destination volume on a remote controller that provides lower-cost storage (for example, Serial Advanced Technology Attachment or SATA disks).

- Create a normal destination volume by running the following command on the destination system:

  `vol create <vol_name>` (with parameters to suit)

- Check the volume’s status and size by running the following command:

  ```
  vol status -b
  ```

  NOTE: The volume must be online and in a writable state.

- *Do not* create the destination qtrees. The destination qtrees are created automatically when the SnapVault relationship is initialized.

NOTE: Although the destination volume remains writable, the individual destination qtrees are read-only.

It is important that you know the requirements and states of the source and destination volumes and that you understand how SnapMirror requirements for the source and destination volumes differ.

Figure 22 illustrates the requirements of the source and destination volumes.

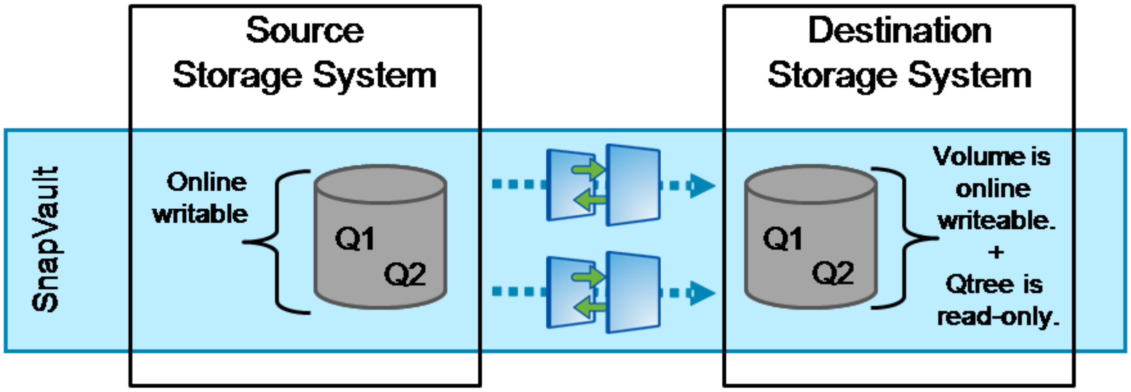


FIGURE 22: SNAPVAULT VOLUME AND QTREE CONFIGURATION

The technology behind the SnapVault feature is based on the qtree SnapMirror function. This function determines many of the features and limitations of the SnapVault feature. For example, the basic unit of SnapVault backup is the qtree, and all SnapVault transfers are based on schedules for asynchronous mode.

Before you can enable the SnapVault relationship, you must configure the SnapVault access control between the source and destination storage controllers. For descriptions of the access control settings, refer to the Security section.

After the source and destination volumes are defined, you can configure the SnapVault schedules on the primary and secondary controllers and start the incremental backups. You can also perform the initial baseline transfer, copying the data from the source qtree to the destination qtree.

1. Configure the primary controller and define a SnapVault schedule.
   ```
   snapvault snap sched vol1 sv_hourly 5@mon-fri@9-19
   ```
2. Configure the secondary controller and perform the baseline transfer.
   ```
   snapvault start –S pri:/vol/vol1/q1 sec:/vol/vol1/q1
   ```
   When the baseline transfer is completed, the destination volume is an exact replica of the source volume.
3. Define a SnapVault schedule.
   ```
   snapvault snap sched –x vol1 sv_hourly 5@mon-fri@9-19
   ```
   The `–x` parameter instructs the secondary controller to request a resynchronization with the primary controller. This request reports the current file system state and then creates a Snapshot copy to retain the data.

NOTE: The SnapVault schedule definition is in the following format:
`<snapshots_to_retain>@<day_of_the_week><@hour_of_the_day>`

The schedule can be specified in more than one way. Because the day of the week is specified as a mnemonic expression, you can define the schedule as the following:
`<snapshots_to_retain><@hour_of_the_day>@<day_of_the_week>`

ADMINISTRATION

The administration of a SnapVault relationship can be performed from either the primary or the secondary system, although some functions are available only on their respective systems.

You use the `snapvault status` command to display the state of the currently defined SnapVault relationships, as shown in Figure 23:

```
SnapVault secondary is ON.
Source                   Destination              State           Lag           Status
pri:/vol/vol1/q1         sec:/vol/vol1/q1         SnapVaulted     00:09:53      Quiescing
pri:/vol/vol1/q2         sec:/vol/vol1/q2         SnapVaulted     00:15:20      Quiescing
```

FIGURE 23: EXAMPLE OF SNAPMIRROR STATUS OUTPUT

NOTE: You can use the `snapvault status –c` option to display the SnapVault qtree configuration parameters.

You can use the `snapvault` command to manage all aspects of the SnapVault relationship, such as updating the secondary system or restoring the backup. The following are examples of frequently used `snapvault` functions:

- `snapvault update sec_hostname:/vol/vol_name/qtree`

  When executed on the *secondary* system, this command triggers a manual (unscheduled) update of the specified qtree destination.

- `snapvault release <path> <other_hostname>:<path>`

  When executed on *either* system, this command deletes the SnapVault relationship.

- `snapvault restore –S sec_hostname:/vol/vol_name/qtree pri_hostname:/vol/vol_name/qtree`

  - When executed on the *primary* system, this command restores the qtree contents from the backup.
  - To restore to the original qtree location on the primary system, you must break the SnapVault relationship or restore to a new qtree (and rename later).
  - To restore a small amount of data (like one file), you may prefer to copy the files from a CIFS share on the secondary qtree.

- `snapvault start –r <path>`

  When executed on the *secondary* system, this command resynchronizes the relationship and resumes backup operations after the SnapVault restore is completed.

NOTE: For information about the other SnapVault commands, refer to the product manual.

Some third-party backup applications use SnapVault integration. To enable these applications to communicate with the controller, you must enable the NDMP protocol and define the user name and password for the application.

PERFORMANCE

One of the challenges of a new SnapVault configuration is the transfer of the baseline copy from the primary to the secondary system. Although the WAN connection may be adequate to handle the incremental backup traffic, it may not be adequate to complete the baseline transfer in a timely manner. In this case, you should consider using the Logical Replication *(*LREP*)* function. The LREP function can perform the initial baseline transfer by using external disk media, such as a USB drive connected to a laptop computer.

Because SnapVault backups are scheduled activities (asynchronous), they are constrained only by the bandwidth of the connection and are not significantly affected by the link latency.

Similar to the qtree SnapMirror process, the SnapVault process accesses the primary qtree at the file-system level and therefore sees (and backs up) the original version of any deduplicated data. Although this process may cause more data to be sent across the WAN than is expected, the secondary qtree is written in the original capacity. Further deduplication can be scheduled on the secondary system.

SECURITY

By default, no access is granted for SnapVault traffic, and specific access must be granted to any remote controller in a backup relationship.

The primary controller must grant access to the secondary controller so that the secondary controller can “pull” backups from the source. And the secondary controller must grant access to the primary controller so that the primary controller can request restores of the backups.

SnapVault access can be configured on the primary and secondary controllers by using the following command:

```
   options snapvault.access host=<other controller>
```

TROUBLESHOOTING

Comprehensive logging of all SnapVault activity is enabled by default. Because the SnapVault function is based on the qtree SnapMirror function, all log information for the SnapVault and qtree SnapMirror functions is stored to the same file.

The log information is saved to the /etc/log/snapmirror.[0-5] files. The log can be disabled by executing the following command:

```
options snapmirror.log.enable [on|off]
```

OPEN SYSTEMS SNAPVAULT (OSSV)

The OSSV agent is a software application that allows SnapVault-like backups (block-level, incremental forever) to be performed by a non-NetApp platform (a Windows, UNIX, or Linux host). The OSSV software is installed on the host (primary). It uses the SnapVault protocol to send the SnapVault backup data to a NetApp controller (secondary). The backup data is then retained using normal SnapVault schedules on the NetApp controller.

CONFIGURATION

OSSV is a licensed feature that must be enabled before it can be configured and used. The OSSV feature requires two licenses. One license is for the OSSV primary server type (the Windows or UNIX backup source). The other license is for the NetApp ONTAP secondary controller (the archive destination).

- License the primary host (sv_windows_pri)

  ```
  license add <licnum>
  ```

  NOTE: The OSSV primary license key must be enabled on the secondary controller and not on the Windows or UNIX host.

- License the secondary controller (sv_ontap_sec)

  ```
  license add <licnum>
  ```

The installation details for the OSSV agent are operating system dependant. For example, a setup EXE is used for the Windows platform, and installation scripts are used for the various UNIX platforms.

Some of the OSSV utilities:

- svconfigpackager

  Unattended installation utility

- svinstallcheck

  Automatic post-installation check

- svconfigurator (GUI)

  The primary configuration tool, used to start and stop the OSSV service, set the NDMP password, enable debugging, enable file tracing, and so on

- svsetstanza

  Command line alternative to the `svconfigurator` GUI tool

NOTE: To read configuration changes, you must restart the OSSV service.

In the OSSV configuration tool (on the primary host), you must configure the access controls and the NDMP user name and password and specify which directories to include in the SnapVault backup.

In general, the configuration of the secondary controller is identical to the configuration of a SnapVault relationship.

ADMINISTRATION

The mechanism that the OSSV agent uses to determine which changed blocks to back up differs greatly from the method used by the SnapVault feature on a primary NetApp controller. However, the administration of the OSSV agent and the SnapVault feature is very similar.

- To back up the OSSV client:
  - Perform an initial baseline backup to the destination
  - Schedule block-level incremental backups
- To restore the OSSV client, use the `snapvault restore` command, similar to how you restore a SnapVault client

One difference between an OSSV client and a SnapVault client is that the clients use different mechanisms to identify which changed blocks to back up. This block-level incremental (BLI) mechanism requires free space on the OSSV client to store a database of previously backed-up files and their block checksum values. Subsequent backups are compared to this database to determine which changed data to back up.

You can use the Free Space Estimator Utility to determine whether there is sufficient disk space on the OSSV primary client to store the database and to perform a BLI backup.

PERFORMANCE

Many of the considerations for SnapVault performance apply to OSSV performance, as the two features perform almost identical functions.

OSSV may differ from a controller based-SnapVault primary by the scale of the backup environment. Even a modest OSSV implementation can have tens (if not hundreds) of OSSV agents installed on Windows and UNIX hosts. Because the OSSV backup traffic concentrates on the SnapVault primary controller, some thought must be given to the number of concurrent backup streams.

Each model of NetApp controller supports a specified number of concurrent backup streams. The number varies according to the model type and, sometimes, according to the software version. The maximum number of concurrent backup streams can be increased by enabling the NearStore® Personality License (NPL) on the secondary controller. The NPL feature is available for all NetApp controller models. Previously, NetApp marketed a purpose-built NearStore appliance, which combined SATA-only storage and the NPL function.

Because the backup traffic is a non-latency sensitive, sequential workload, it is generally recommended that the secondary backup data be stored on a SATA disk.

NOTE: The NearStore feature was originally available only on the dedicated NearStore controller, but it is now available as a licensed feature on all NetApp FAS controllers. Some documentation may refer to the original NearStore hardware requirement. A difference between the NearStore hardware and software implementations is that the hardware implementation provides a primary system *or* a secondary system, but the software implementation provides both a primary *and* a secondary system.

SECURITY

By default, no access is granted for OSSV traffic, and specific access must be granted for any remote controller in a backup relationship.

The OSSV primary host (Windows or UNIX) must grant access for the secondary controller so that the secondary controller can “pull” backups from the source. The secondary controller must be granted access to the OSSV primary host so that the host can request restores of the backups.

OSSV and SnapVault access can be configured as follows:

- On the primary (Windows or UNIX) host, use the client configuration tool to edit the QSM Access List field and enter the host name of the secondary controller.
- On the secondary (NetApp) controller , run the following command:

  ```
  options snapvault.access host=<other controller>
  ```

You must set the NDMP user name and password on the OSSV client. You can use either the client configuration GUI or the `svpassword` command.

TROUBLESHOOTING

The log-file locations for the OSSV agents are operating-system dependant.

- Primary Windows OSSV client

  ```
  c:\Program Files\netapp\snapvault\etc\snapvault.yyyymmdd
  ```

- Primary UNIX and Linux OSSV client

  ```
  /usr/snapvault/snapvault.yyyymmdd
  ```

- Secondary NetApp SnapVault controller

  ```
  /etc/log/snapmirror.[0-5]
  ```

HIGH-AVAILABILITY CONFIGURATION

In the past, NetApp used the term “active-active” to describe the high-availability (HA) controller failover configuration. Two controller heads are configured as a pair, with each node providing failover support for its partner.

CONFIGURATION

High-availability is a licensed feature that must be enabled before it can be configured and used. The high-availability feature must be licensed on both HA nodes.

- License the first node.
  ```
  license add <licnum>
  ```
- License the second node.
  ```
  license add <licnum>
  ```

NOTE: You must then reboot and start the controller failover feature.

After the high-availability feature is enabled, you can unlicense it only when the HA pair is in a normal state and the controller failover services are manually disabled.

Before you enable the high-availability feature, you must configure various settings. For example, the HA pair interconnect cable must be attached, both controllers must be provided Fibre connections to all expansion drawers, and both controllers must be provided access to the same IP subnets.

The high-availability feature activates numerous high availability capabilities, such as NVRAM mirroring, which enables the controllers to provide failover support for each other.

For example, if a controller failure occurs:

- The surviving node spawns a virtual instance of the failed node
- The virtual node accesses its mirrored NVRAM to complete any interrupted write
- The local network interface assumes the IP address of both the local and partner interfaces (for Ethernet traffic)
- The local FC interfaces retain their original WWPN addresses, and the host-based MPIO drivers direct all FC traffic via the interfaces (assuming that single-image `cfmode` is being used)

The process of removing a high-availability configuration is as follows:

1. Disable the controller failover feature (`cf disable`).
2. Delete the controller failover license (`license delete` …).
3. Remove the partner’s network entries from the /etc/rc file.
4. Halt, and make sure the partner-sysid is blank.
5. Power down and remove or relocate the controller failover interconnect card.
6. Repeat the process on the other controller

ADMINISTRATION

In a high-availability configuration, all disks are visible to both controllers. Before a disk can be used in an aggregate or a spare, it must be assigned to one or the other controller. This process is known as “software disk assignment.” If a disk is not assigned to a controller (in order words, if it is listed as “not owned”), then it cannot be used by either controller for any purpose.

Use the following commands to manage disk ownership:

- Assign disk ownership

  ```
  disk assign
  ```

- List disk ownership (several methods)

  ```
  disk show –v
  storage show disk
  sysconfig –r
  ```

NOTE: Earlier versions of Data ONTAP supported *hardware disk assignment*, where ownership was determined by the Fibre Channel cabling topology. This mode is not supported on any current-generation controller.

Generally, administration of a high-availability configuration and administration of two non-clustered controllers are identical. The clustered controllers are managed separately, although some configuration settings must be synchronized between the two controllers. One of the features that you must master is the process of HA pair failover and failback.

For example, after a failed controller is rebooted and ready to assume its old identity and workload, it displays a “waiting for giveback” or "waiting for mb giveback" message. At this point, the administrator enters the `cf giveback` command on the operational controller to return the failed controller back to the normal state.

NOTE: For more information about controller failover management, refer to the product manuals.

PERFORMANCE

Optimum performance is usually achieved when the controllers in a high-availability configuration share the client workload evenly. An even distribution of the workload is usually attributable to good solution planning and to automatic load balancing in the host-based MPIO drivers (for FC traffic).

In an FC SAN environment, ensure that the host-based multipathing support is correctly configured. Where appropriate, use Asymmetric Logical Unit Access (ALUA) support.

In most other ways, the performance concerns of a high-availability configuration and of a non-clustered configuration are identical.

SECURITY

In almost all aspects of security, a high-availability configuration and a non-clustered configuration are identical.

TROUBLESHOOTING

In a high-availability configuration, both controllers require connectivity to all of the disk expansion shelves. It is not possible to have a shelf connected to one controller and not to the other controller. If a controller loses access to one of the disk shelves, a negotiated (clean, but automatic) failover is triggered.

It is recommended that multipath HA cabling be used for the disk-expansion shelf connections. The cabling prevents unnecessary controller failover for non-critical reasons, such as SFP failure, loop breaks, ESH module failure, and so on.

METROCLUSTER FEATURE AND SYNCMIRROR SOFTWARE

You can use MetroCluster technology to split two controllers of a high-availability pair and position them in two physical locations. This process involves the use of a SyncMirror configuration to mirror the data between the two locations. Then, if a site disaster occurs, you can fail over the HA pair, restore operations at the remote site, and minimize disruption.

CONFIGURATION

The MetroCluster feature combines two hardware configurations and several licensed features. The two MetroCluster modes, stretch mode and fabric-attached mode, require different hardware configurations. All of the components must be connected and enabled before they can be configured and used.

- Hardware requirements
  - High-availability controller heads
  - In fabric mode only
    - MetroCluster FC/VI adapter
    - FC switches
- Software license requirements
  - Cf license
  - Cf_Remote license
  - SyncMirror license

There are a number of prerequisites and limitations that you should be aware of before you attempt to configure a MetroCluster environment.

- **Distance / Latency limitations**
  - Stretch mode
    - The maximum cable distance between the two controllers is 500 m.
    - The 500-m limitation is primarily a limitation of the MultiMode Fibre cable that is used to connect the two controllers. If cable type and patches are not adequate, the limitation may be less than 500 m.
  - Fabric-attached mode
    - The maximum supported distance between the two controllers is 100 km.
    - The 100-km limitation is primarily a limitation of the latency across the distance. If additional latency is present, the limitation may be less than 100 km.
    - Other factors, such as the type of laser small form-factor pluggable (SFP) that is used or whether fiber optic repeaters or wavelength-division multiplexing (WDM) devices are used, also affect the maximum supported distance.
- Disk types and expansion shelves
  - Both controllers require connectivity to all expansion shelves.
  - Stretch mode supports both FC and SATA expansion shelves.
  - Fabric-attached mode supports only FC type expansion shelves (although SATA shelves can be present if they are not mirrored).
  - In fabric mode, disk ownership is determined by where the controller's HBAs connect to the switch and where the disk shelves connect to the switch.
- Networking
  - Prior to Data ONTAP 7.3.2, both controllers (at their individual sites) require connectivity to the same IP network subnet ranges or VLANs. With Data ONTAP 7.3.2 and later, the MetroCluster pair maybe on a separate subnet and use the /etc/mcrc file to configure the interfaces appropriately.

- SyncMirror configuration
  - An even number of disks are required. The disks must be divided evenly between the two controller locations.
  - All disks must be the same size. If the disks are not the same size, they are resized.
  - All disks must be of the same type. In other words, you cannot mirror FC disks to SATA disks.

NOTE: This section summarizes configuration requirements and capabilities. Before attempting to design or install a MetroCluster environment, refer to the product manuals.

The following diagrams illustrate the main components of MetroCluster and the differences between stretch mode and fabric-attached mode.

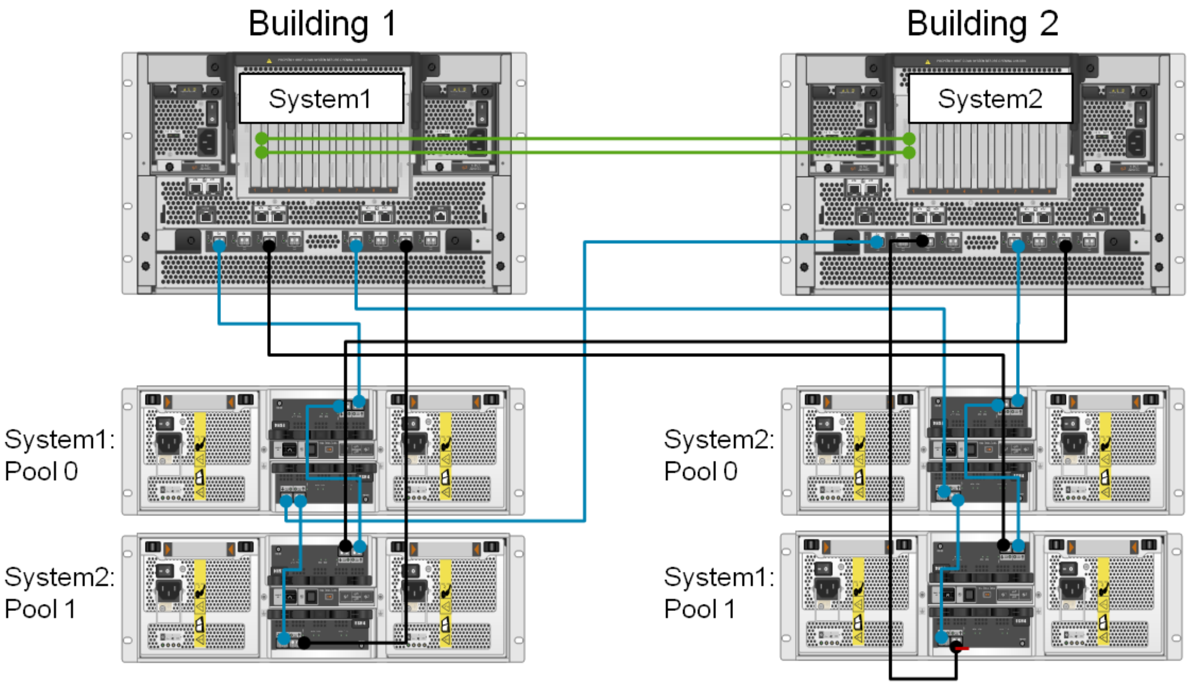


FIGURE 24:STRETCH MODE METROCLUSTER DIAGRAM

NOTE: The number and ownership of the disk expansion shelves is indicative only. For more information about the supported expansion shelf configurations in a MetroCluster environment, refer to the product documentation.

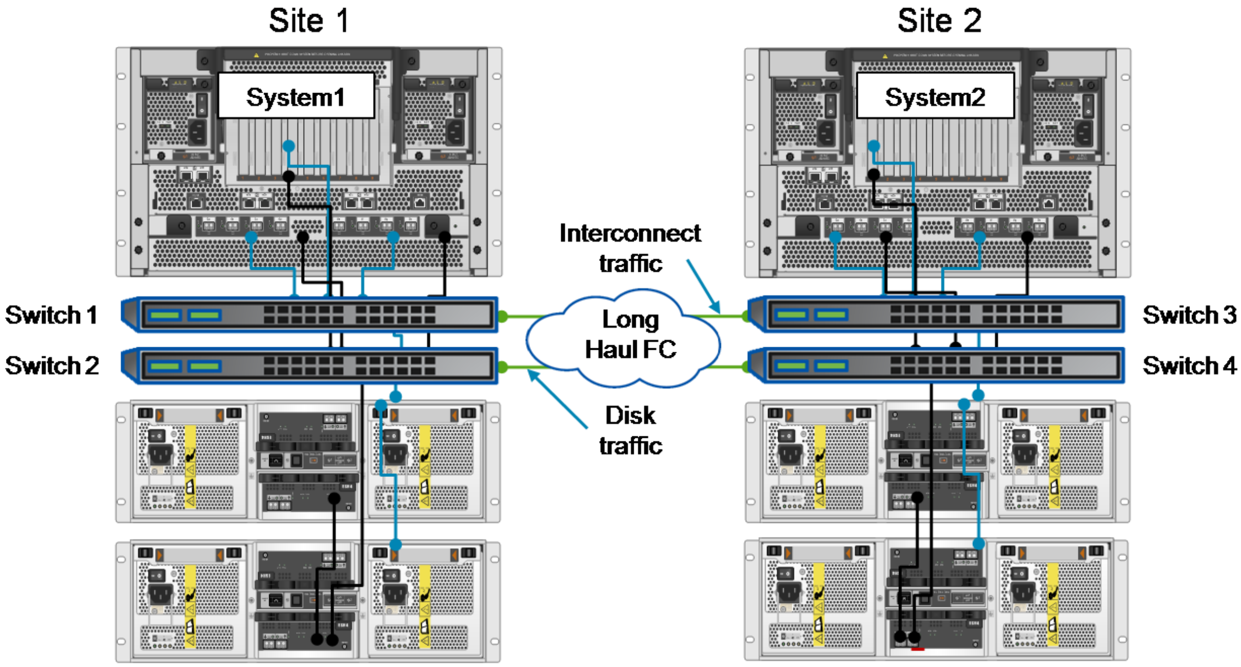


FIGURE 25: FABRIC-ATTACHED MODE METROCLUSTER DIAGRAM

ADMINISTRATION

After the MetroCluster-specific configuration is completed (disk pool assignment, SyncMirror configuration, and so on), the storage administration is generally equivalent to the storage administration for a high-availability controller.

An exception to this equivalency concerns failover management of HA pairs. Normally, failover occurs automatically. However, with MetroCluster, failover requires administrative intervention. Intervention protects against the split-brain scenario. In this scenario, a communications outage causes both controllers to assume that the other controller has failed, and each controller attempts to assume the partner's role.

In a MetroCluster environment, the state of the partner controller determines which takeover command is used. In extreme cases, where a site disaster has made the partner controller unresponsive (offline or destroyed), you may need to use the `cf forcetakeover -d` command. The command allows HA pair failover to proceed even when configuration problems might otherwise prevent the implementation of a takeover command. The command also splits the SyncMirror relationship.

The use of the `forcetakeover` option is *very dangerous*. If the partner controller is operational and able to access its storage, use of the `forcetakeover` option can corrupt the data. Use the `forcetakeover` command *only* if the remote MetroCluster partner node is powered off and inaccessible.

NOTE: For more information about MetroCluster takeover and giveback procedures, refer to the product documentation.

If you are using SyncMirror software to provide local mirroring (not in a MetroCluster configuration), then at some point you may wish to split a mirrored volume and disable the SyncMirror software.

The following is an example of the process that you use to split a mirror and disable SyncMirror software:

1. Check the current status.

   Before you split a SyncMirror volume, both plexes should be online and operational.

2. Split the mirrored volume by running the `vol split vol0/plex0 vol0new` command.

   You should now have two unmirrored volumes.

3. Disable the SyncMirror license by running the `license delete <licnum>` command.

NOTE: If one or more mirrored volumes exist, you cannot disable the SyncMirror license.

PERFORMANCE

Except for the potential for additional latency in the SyncMirror configuration, the performance of the MetroCluster feature and the performance of a standard high-availability controller are identical.

The 100-km maximum-distance limitation for a fabric MetroCluster environment is intended to limit the additional latency to a maximum of 1 ms. The 1-ms maximum is usually seen as the practical limit for synchronous replication.

If you need to replicate across a greater distance or where the latency may be higher than 1 ms, then you should consider a standard SnapMirror (semi-synchronous or asynchronous) configuration.

SECURITY

In almost all aspects of security, a MetroCluster configuration and a standard high-availability controller configuration are identical.

Usually, both controllers in a high-availability and MetroCluster configuration are enabled. Therefore, storage service is provided to both sites, and, if a disaster occurs, the simplest and fastest site failover is provided. However, in some environments, you may want to limit or prevent access to the remote controller. This result can be achieved in several ways. For example, you can configure the appropriate SAN and NAS security (for example, manual NFS fencing), or you can isolate the remote controller by powering off the remote node (If a disaster occurs, the node must be manually powered on to enable a takeover to be performed.).

TROUBLESHOOTING

A MetroCluster configuration, typically spanning two sites, is more vulnerable to communications disruptions than a standard high-availability controller. The two controller heads must maintain communication by way of at least one of the two interconnect ports (over fiber) and over IP (over the WAN). Both controllers need connectivity to all expansion shelves (over fiber). Systems uses multipath HA cabling.

It is good design to ensure that the various redundant inter-site connections (assuming that they exist) are located in physically separate conduits. Always remember the maxim—men with backhoes are irresistibly drawn to network cabling.

A prolonged disruption of the communications between the two controllers is classified as a disaster. Such a disruption requires administrative action to resolve the MetroCluster site failover.

If you specified multipath HA cabling to the disk expansion shelves, then you should double-check the connectivity by running the following command:

```
storage show disk –p
```

In a multipath HA configuration, the output should list two paths to each disk.

In a MetroCluster configuration, each site is assigned its own disks, including its own spare disks. The spare disks assigned to a site are, as a group, known as a pool. Disk-site assignment ensures that adequate spare disks are located at each site.

If a disk in a RAID-4 SyncMirror volume fails (or if two disks in a RAID-DP volume fail) and the pool contains no spare disks, then the system generates a warning and continues normal operation. Even though there is no disk redundancy in the volume in which the failure occurred, the volume does not go into degraded mode or shut down after the RAID timeout. This result (lack of result) occurs because the volume’s data integrity is guaranteed by the mirror plex at the other site (or pool).

ADDITIONAL MATERIALS

The following resources are recommended as preparation for the NCDA certification exams.

PRACTICE EXAM

NS0-154—Data ONTAP 8.0 7-Mode Administrator

http://now.netapp.com/NOW/products/education/public/certification/NS0-154-PRACTICE/index.html

FURTHER READING

- Product documentation

  http://now.netapp.com/NOW/knowledge/docs/ontap/ontap_index.shtml

  - System Administration Guide
  - Storage Administration Guide
  - High Availability Configuration Guide
  - Network Management Guide
  - File Access and Protocols Management Guide
  - Block Access Management Guide
  - Data Protection Online Backup and Recovery Guide
  - Data Protection Tape Backup and Recovery Guide
  - Commands: Manual Page Reference, Volume 1
  - Commands: Manual Page Reference, Volume 2
  - Core Command Quick Reference

- Technical reports

  http://www.netapp.com/us/library/technical-reports.html

- Best-practice guides

  http://now.netapp.com/NOW/knowledge/docs/docs.cgi