IBM System Storage N series
Data ONTAP 7.3 Storage Management Guide

GC52-1277-04
NA 210-05202_A0

Contents

Copyright and trademark information

Copyright information

Trademark information

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. A complete and current list of other IBM trademarks is available on the Web at http://www.ibm.com/legal/copytrade.shtml

Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.

Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

UNIX is a registered trademark of The Open Group in the United States and other countries.

NetApp, the NetApp logo, Network Appliance, the Network Appliance logo, ApplianceWatch, ASUP, AutoSupport, Bycast, Campaign Express, ComplianceClock, Cryptainer, CryptoShred, Data ONTAP, DataFabric, DataFort, Decru, Decru DataFort, FAServer, FilerView, FlexCache, FlexClone, FlexScale, FlexShare, FlexSuite, FlexVol, FPolicy, GetSuccessful, gFiler, Go further, faster, Imagine Virtually Anything, Lifetime Key Management, LockVault, Manage ONTAP, MetroCluster, MultiStore, NearStore, NetCache, NOW (NetApp on the Web), ONTAPI, OpenKey, RAID-DP, ReplicatorX, SANscreen, SecureAdmin, SecureShare, Select, Shadow Tape, Simulate ONTAP, SnapCopy, SnapDirector, SnapDrive, SnapFilter, SnapLock, SnapManager, SnapMigrator, SnapMirror, SnapMover, SnapRestore, Snapshot, SnapSuite, SnapValidator, SnapVault, StorageGRID, StoreVault, the StoreVault logo, SyncMirror, Tech OnTap, The evolution of storage, Topio, vFiler, VFM, Virtual File Manager, VPolicy, WAFL, and Web Filer are trademarks or registered trademarks of NetApp, Inc. in the United States, other countries, or both.

All other brands or products are 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.

NetApp NetCache is certified RealSystem compatible.

Notices

This information was developed for products and services offered in the U.S.A.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe on any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing to:

IBM Director of Licensing
IBM Corporation
North Castle Drive
Armonk, N.Y. 10504-1785
U.S.A.

For additional information, visit the web at:
http://www.ibm.com/ibm/licensing/contact/

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law:

INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM web sites are provided for convenience only and do not in any manner serve as an endorsement of those web sites. The materials at those web sites are not part of the materials for this IBM product and use of those web sites is at your own risk.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Any performance data contained herein was determined in a controlled environment. Therefore, the results obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurement may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

If you are viewing this information in softcopy, the photographs and color illustrations may not appear.

About this guide

You can use your product more effectively when you understand this document's intended audience and the conventions that this document uses to present information.

This document describes how to configure, operate, and manage the storage resources for storage systems that run Data ONTAP software. It covers disks, RAID groups, plexes, and aggregates, and how file systems, or volumes, and qtrees are used to organize and manage data.

Note: In this document, the term *gateway* describes IBM N series storage systems that have been ordered with gateway functionality. Gateways support various types of storage, and they are used with third-party disk storage systems—for example, disk storage systems from IBM, HP®, Hitachi Data Systems®, and EMC®. In this case, disk storage for customer data and the RAID controller functionality is provided by the back-end disk storage system. A gateway might also be used with disk storage expansion units specifically designed for the IBM N series models.

The term *filer* describes IBM N series storage systems that either contain internal disk storage or attach to disk storage expansion units specifically designed for the IBM N series storage systems. Filer storage systems do not support using third-party disk storage systems.

Next topics

Audience

This document is written with certain assumptions about your technical knowledge and experience.

This document is for system administrators and service personnel who are familiar with storage system equipment and who need to perform the following tasks:

- Create and maintain aggregates and volumes
- Remove and replace disks
- Organize or limit access to storage space using qtrees and quotas

Supported features

IBM® System Storage™ N series storage systems are driven by NetApp® Data ONTAP® software. Some features described in the product software documentation are neither offered nor supported by IBM. Please contact your local IBM representative or reseller for further details. Information about supported features can also be found at the following Web site:

www.ibm.com/storage/support/nas/

A listing of currently available N series products and features can be found at the following Web site:

www.ibm.com/storage/nas/

Getting information, help, and services

If you need help, service, or technical assistance or just want more information about IBM products, you will find a wide variety of sources available from IBM to assist you. This section contains information about where to go for additional information about IBM and IBM products, what to do if you experience a problem with your IBM N series product, and whom to call for service, if it is necessary.

Next topics

Before you call

Before you call, make sure that you have taken these steps to try to solve the problem yourself:

- Check all cables to make sure that they are connected properly.
- Check the power switches to make sure that the system is turned on.
- Use the troubleshooting information in your system documentation and use the diagnostic tools that come with your system.

Using the documentation

Information about N series hardware products is available in printed documents and a documentation CD that comes with your system. The same documentation is available as PDF files on the IBM NAS support Web site:

www.ibm.com/storage/support/nas/

Data ONTAP software publications are available as PDF files on the IBM NAS support Web site:

www.ibm.com/storage/support/nas/

Web sites

IBM maintains pages on the World Wide Web where you can get the latest technical information and download device drivers and updates.

- For NAS product information, go to the following Web site:
  www.ibm.com/storage/nas/
- For NAS support information, go to the following Web site:
  www.ibm.com/storage/support/nas/
- For AutoSupport information, go to the following Web site:
  www.ibm.com/storage/support/nas/
- For the latest version of publications, go to the following Web site:
  www.ibm.com/storage/support/nas/

Accessing online technical support

For online Technical Support for your IBM N series product, visit the following Web site:

www.ibm.com/storage/support/nas/

Hardware service and support

You can receive hardware service through IBM Integrated Technology Services. Visit the following Web site for support telephone numbers:

www.ibm.com/planetwide/

Supported servers and operating systems

IBM N series products attach to many servers and many operating systems. To determine the latest supported attachments, follow the link to the Interoperability Matrices from the following Web site:

www.ibm.com/systems/storage/network/interophome.html

Firmware updates

As with all devices, it is recommended that you run the latest level of firmware, which can be downloaded by visiting the following Web site:

www.ibm.com/storage/support/nas/

Verify that the latest level of firmware is installed on your machine before contacting IBM for technical support. See the *Data ONTAP Upgrade Guide* for your version of Data ONTAP for more information on updating firmware.

Accessing Data ONTAP man pages

You can use the Data ONTAP manual (man) pages to access technical information.

About this task

Data ONTAP manual pages are available for the following types of information. They are grouped into sections according to standard UNIX naming conventions.

Types of information	Man page section
Commands	1
Special files	4
File formats and conventions	5
System management and services	8

Step

1. View man pages in the following ways:
   - Enter the following command at the storage system command line:

     `man` ***`command_or_file_name`***
   - Click the manual pages button on the main Data ONTAP navigational page in the FilerView user interface.
   - Use the *Commands: Manual Page Reference*, Volumes 1 and 2

     Note: All Data ONTAP man pages are stored on the storage system in files whose names are prefixed with the string "na_" to distinguish them from client man pages. The prefixed names are used to distinguish storage system man pages from other man pages and sometimes appear in the NAME field of the man page, but the prefixes are not part of the command, file, or services.

Terminology

To understand the concepts in this document, you might need to know how certain terms are used.

Storage terms

array LUN — The storage that third-party storage arrays provide to storage systems running Data ONTAP software. One array LUN is the equivalent of one disk on a native disk shelf.

LUN (logical unit number) — A logical unit of storage identified by a number.

native disk — A disk that is sold as local storage for storage systems that run Data ONTAP software.

native disk shelf — A disk shelf that is sold as local storage for storage systems that run Data ONTAP software.

storage controller — The component of a storage system that runs the Data ONTAP operating system and controls its disk subsystem. Storage controllers are also sometimes called *controllers*, *storage appliances*, *appliances*, *storage engines*, *heads*, *CPU modules*, or *controller modules*.

storage system — The hardware device running Data ONTAP that receives data from and sends data to native disk shelves, third-party storage, or both. Storage systems that run Data ONTAP are sometimes referred to as *filers*, *appliances*, *storage appliances*, *gateways*, or *systems*.

> **Note:** The term *gateway* describes IBM N series storage systems that have been ordered with gateway functionality. Gateways support various types of storage, and they are used with third-party disk storage systems—for example, disk storage systems from IBM, HP®, Hitachi Data Systems®, and EMC®. In this case, disk storage for customer data and the RAID controller functionality is provided by the back-end disk storage system. A gateway might also be used with disk storage expansion units specifically designed for the IBM N series models.
>
> The term *filer* describes IBM N series storage systems that either contain internal disk storage or attach to disk storage expansion units specifically designed for the IBM N series storage systems. Filer storage systems do not support using third-party disk storage systems.

third-party storage — The back-end storage arrays, such as IBM, Hitachi Data Systems, and HP, that provide storage for storage systems running Data ONTAP.

Cluster and high-availability terms

active/active configuration

- In the Data ONTAP 7.2 and 7.3 release families, a pair of storage systems or gateways (sometimes called *nodes*) configured to serve data for each other if one of the two systems stops functioning. Also sometimes referred to as *active/active pairs*.
- In the Data ONTAP 8.*x* release family, this functionality is referred to as a *high-availability (HA) configuration* or an *HA pair*.
- In the Data ONTAP 7.1 release family, this functionality is referred to as a *cluster*.

cluster

- In Data ONTAP 8.*x*, a group of connected nodes (storage systems) that share a global namespace and that you can manage as a single virtual server or multiple virtual servers, providing performance, reliability, and scalability benefits.
- In the Data ONTAP 7.1 release family, a pair of storage systems (sometimes called *nodes*) configured to serve data for each other if one of the two systems stops functioning.
- In the Data ONTAP 7.3 and 7.2 release families, this functionality is referred to as an *active/active configuration*.
- For some storage array vendors, *cluster* refers to the hardware component on which host adapters and ports are located. Some storage array vendors refer to this component as a *controller*.

HA (high availability)

- In Data ONTAP 8.*x*, the recovery capability provided by a pair of nodes (storage systems), called an *HA pair*, that are configured to serve data for each other if one of the two nodes stops functioning.
- In the Data ONTAP 7.3 and 7.2 release families, this functionality is referred to as an *active/active configuration*.

HA pair

- In Data ONTAP 8.*x*, a pair of nodes (storage systems) configured to serve data for each other if one of the two nodes stops functioning.
- In the Data ONTAP 7.3 and 7.2 release families, this functionality is referred to as an *active/active configuration*.

Where to enter commands

You can use your product more effectively when you understand how this document uses command conventions to present information.

You can perform common administrator tasks in one or more of the following ways:

- You can enter commands either at the system console or from any client computer that can obtain access to the storage system using a Telnet or Secure Shell (SSH) session.
  In examples that illustrate command execution, the command syntax and output shown might differ from what you enter or see displayed, depending on your version of the operating system.
- You can use the FilerView graphical user interface.
  For information about accessing your system with FilerView, see the *Data ONTAP System Administration Guide*.
- You can enter Windows, ESX, HP-UX, AIX, Linux, and Solaris commands at the applicable client console.
  In examples that illustrate command execution, the command syntax and output shown might differ from what you enter or see displayed, depending on your version of the operating system.

Keyboard and formatting conventions

You can use your product more effectively when you understand how this document uses keyboard and formatting conventions to present information.

Keyboard conventions

Convention	What it means
The IBM NAS support site	Refers to *www.ibm.com/storage/support/nas/*.
Enter, *enter*	• Used to refer to the key that generates a carriage return; the key is named Return on some keyboards. • Used to mean pressing one or more keys on the keyboard and then pressing the Enter key, or clicking in a field in a graphical interface and then typing information into the field.
hyphen (-)	Used to separate individual keys. For example, Ctrl-D means holding down the Ctrl key while pressing the D key.
type	Used to mean pressing one or more keys on the keyboard.

Formatting conventions

Convention	What it means
Italic font	• Words or characters that require special attention. • Placeholders for information that you must supply. For example, if the guide says to enter the `arp -d` *`hostname`* command, you enter the characters "arp -d" followed by the actual name of the host. • Book titles in cross-references.
`Monospaced` font	• Command names, option names, keywords, and daemon names. • Information displayed on the system console or other computer monitors. • Contents of files. • File, path, and directory names.
`Bold monospaced` font	Words or characters you type. What you type is always shown in lowercase letters, unless your program is case-sensitive and uppercase letters are necessary for it to work properly.

Special messages

This document might contain the following types of messages to alert you to conditions that you need to be aware of.

Note: A note contains important information that helps you install or operate the system efficiently.

Attention: An attention notice contains instructions that you must follow to avoid a system crash, loss of data, or damage to the equipment.

How to send your comments

Your feedback is important in helping us provide the most accurate and high-quality information. If you have comments or suggestions for improving this document, send us your comments by e-mail to *starpubs@us.ibm.com*. Be sure to include the following:

- Exact publication title
- Publication form number (for example, GC26-1234-02)
- Page, table, or illustration numbers
- A detailed description of any information that should be changed

Data ONTAP storage architecture overview

Storage architecture refers to how Data ONTAP provides data storage resources to host or client systems and applications. Data ONTAP distinguishes between the physical layer of data storage resources and the logical layer.

- The physical layer includes disks, array LUNs, RAID groups, plexes, and aggregates.

  Note: A disk is the basic unit of storage for storage systems that use Data ONTAP to access native disk shelves. An array LUN is the basic unit of storage that a third-party storage array provides to a storage system that runs Data ONTAP.

- The logical layer includes the file systems— volumes, qtrees, logical unit numbers (LUNs)— and the directories and files that store data.

  Note: LUNs are storage target devices in iSCSI and FC networks.

Aggregates provide storage to volumes. Aggregates can be composed of either disks or array LUNs, but not both. Data ONTAP organizes the disks or array LUNs in an aggregate into one or more RAID groups. Aggregates have one or two plexes, depending on whether RAID-level mirroring (SyncMirror), is in use.

Volumes are data containers. Clients can access the data in volumes through the access protocols supported by Data ONTAP. These protocols include Network File System (NFS), Common Internet File System (CIFS), HyperText Transfer Protocol (HTTP), Web-based Distributed Authoring and Versioning (WebDAV), Fibre Channel Protocol (FCP), and Internet SCSI (iSCSI).

You can partition volumes and control resource usage using qtrees. You can create LUNs for use in a SAN environment, using the FCP or iSCSI access protocols. Volumes, qtrees, and LUNs contain directories and files.

Note: Starting in Data ONTAP 7.3, gateways also support native disk shelves. See the *Gateway Implementation Guide for Native Disk Shelves* for more information.

Related concepts

How Data ONTAP works with disks on page 31
Managing array LUNs through Data ONTAP on page 87
How Data ONTAP uses RAID to protect your data and data availability on page 101
How you use aggregates to provide storage to your volumes on page 125
How volumes work on page 155
About qtrees on page 285

Related information

IBM NAS documentation and support site - www.ibm.com/storage/support/nas

How Data ONTAP works with disks

Disks provide the basic unit of storage for storage systems running Data ONTAP. Understanding how Data ONTAP uses and classifies disks will help you manage your storage more effectively.

Next topics

What disk types Data ONTAP supports

Data ONTAP supports five disk types: Fibre Channel (FC), Advanced Technology Attachment (ATA), Serial Advanced Technology Attachment (SATA), Serial Attached SCSI (SAS), and Bridged SAS (BSAS).

BSAS disks are SATA disks with added hardware (the bridge) that allows them to appear as SAS drives to the disk shelf.

For a specific configuration, the disk types supported depend on the storage system model, the disk shelf type, and the I/O modules installed in the system. For more information about the types of disks supported by your configuration, see the appropriate hardware and service guide.

For information about best practices for working with different types of disks, see *Technical Report 3437: Storage Best Practices and Resiliency Guide*.

Note: This technical report contains information about NetApp products that IBM licenses and in some cases customizes. Technical reports might contain information about models and features that are not supported by IBM.

Related concepts

Rules for mixing disk types in aggregates on page 132

Related information

TR 3437: Storage Best Practices and Resiliency Guide

Disk connection protocols, topologies, and types

Data ONTAP supports two disk connection protocols: serial-attached SCSI (SAS) and Fibre Channel (FC). The Fibre Channel protocol supports three topologies: arbitrated loop, switched, and point-to-point.

- SAS, BSAS, and SATA disks use the SAS disk connection protocol.
- FC and ATA disks use the Fibre Channel protocol with an arbitrated loop topology, or FC-AL.
- Array LUNs use the FC protocol, with either the point-to-point or switched topology.

You cannot combine different disk connection types in the same loop or stack.

Next topics

Fibre Channel-Arbitrated Loop (FC-AL) disk connection type on page 32
Serial-attached SCSI (SAS) disk connection type on page 32

Fibre Channel-Arbitrated Loop (FC-AL) disk connection type

For the FC-AL disk connection type, disk shelves are connected to the controller in a loop.

Devices must arbitrate for the chance to communicate over the loop to avoid collisions on the loop. If connectivity is lost somewhere along the loop and a redundant path is not available, the controller loses the ability to communicate with some devices on the loop.

You cannot combine disk shelves containing FC disks and disk shelves containing ATA disks in the same loop.

Serial-attached SCSI (SAS) disk connection type

The SAS disk connection type is a point-to-point architecture. This means that the controller can communicate with more than one device at once.

Disk shelves are connected to the controller on a daisy chain called a *stack*.

For information about combining different disk types within a stack, see the *Hardware and Service Guide* for your SAS disk shelf.

Available disk capacity by disk size

To maintain compatibility across brands of disks, Data ONTAP rounds down ("right-sizes") the amount of space available for user data.

Because of right-sizing, informational commands such as `sysconfig` show a lower number for available space than the disk's rated capacity (you use rated capacity if you specify disk size when creating an aggregate). The available disk space is rounded down as shown in the following table.

Note: For this table, GB = 1,000 MB.

The capacity numbers in this table do not take into account the 10 percent of disk space that Data ONTAP reserves for its own use.

Disk size	Right-sized capacity	Available blocks
FC disks		
36 GB	34.5 GB	70,656,000
72 GB	68 GB	139,264,000
144 GB	136 GB	278,528,000
300 GB	272 GB	557,056,000
450 GB	418 GB	856,064,000
600 GB	560 GB	1,146,880,000
ATA disks—FC connected		
250 GB	211 GB	432,901,760
320 GB	274 GB	561,971,200
500 GB	423 GB	866,531,584
750 GB	635 GB	1,301,618,176
1 TB	847 GB	1,735,794,176
2 TB	1,695 GB	3,472,315,904
SAS disks—SAS connected		
144 GB	136 GB	278,528,000
300 GB	272 GB	557,056,000
450 GB	418 GB	856,064,000

Disk size	Right-sized capacity	Available blocks
600 GB	560 GB	1,146,880,000
SATA disks—SAS connected		
250 GB	211 GB	432,901,760
500 GB	423 GB	866,531,584
750 GB	635 GB	1,301,618,176
1 TB	847 GB	1,735,794,176
2 TB	1,695 GB	3,472,315,904

Disk speeds supported by Data ONTAP

For hard disk drives, which use rotating media, speed is measured in revolutions per minute (RPM). Faster disks provide more disk input/output operations per second (IOPS) and faster response time.

It is best to use disks of the same speed in an aggregate.

Data ONTAP supports the following rotational speeds for disks:

- FCAL disks (FC-AL connected)
  - 10K RPM
  - 15K RPM
- ATA disks (FC-AL connected)
  - 5.4K RPM
  - 7.2K RPM
- SAS disks (SAS-connected)
  - 15K RPM
- SATA disks (SAS-connected)
  - 7.2K RPM
- BSAS disks (SAS-connected)
  - 7.2K RPM

For more information about supported disk speeds, see the appropriate hardware and service guide. For information about optimizing performance with 15K RPM FC disks, see *Technical Report 3285: 15,000 RPM Fibre Channel Disk Drives: A Best-Practice Guide for Optimizing System Performance*.

Note: This technical report contains information about NetApp products that IBM licenses and in some cases customizes. Technical reports might contain information about models and features that are not supported by IBM.

Related concepts

How you can use disks with mixed speeds in the same aggregate on page 130
How you use aggregates to provide storage to your volumes on page 125

Related information

TR 3285: 15,000 RPM Fibre Channel Disk Drives: A Best-Practice Guide for Optimizing System Performance

Disk formats supported by Data ONTAP

The disk format determines how much of the disk's raw capacity can be used for data storage.

All disks used in storage systems are block checksum disks (BCS disks).

The amount of space available for data depends on the bytes per sector (bps) of the disk:

- Disks that use 520 bps provide 512 bytes per sector for data. 8 bytes per sector are used for the checksum.
- Disks that use 512 bps use some sectors for data and others for checksums. For every 9 sectors, 1 sector is used for the checksum, and 8 sectors are available for data.

The disk formats by Data ONTAP disk type are as follows:

- FCAL and SAS BCS disks use 520 bps.
- ATA, SATA, and BSAS BCS disks use 512 bps.

How to interpret disk names

Each disk has a name that differentiates it from all other disks for a storage system. Disk names have different formats, depending on the disk connection type (FC-AL or SAS) and whether the disk is directly attached to the storage system or attached to a switch.

The following table shows the various formats for disk names, depending on how they are connected to the storage system.

Note: For internal disks, the slot number is zero, and the internal port number depends on the system model.

Disk connection	Disk name	Example
FC-AL, direct-attached	<slot><port>.<loopID>	The disk with loop ID 19 (bay 3 of shelf 1) connected to onboard port 0a has an address of 0a.19. The disk with loop ID 34 connected to an HBA in slot 8, port c has an address of 8c.34.
FC-AL, switch-attached	<switch_name>.<port>.<loopID>	The disk with loop ID 51 connected to port 3 of switch SW7 has an address of SW7.3.51.
SAS, direct-attached	<slot><port>.<shelfID>.<bay>	The internal SAS-connected disk in bay 9 for a N3400 has an address of 0c.0.9. The disk in shelf 2, bay 11, connected to onboard port 0a has an address of 0a.2.11. The disk in shelf 6, bay 3, connected to an HBA in slot 1, port c, has an address of 1c.6.3.

Loop IDs for FC-AL connected disks

For disks connected using Fibre Channel-Arbitrated Loop (FC-AL or FC), the loop ID is an integer between 16 and 126. The loop ID identifies the disk within its loop, and is included in the disk name, which identifies the disk uniquely for the entire system.

The loop ID corresponds to the disk shelf number and the bay in which the disk is installed. The lowest loop ID is always in the far right bay of the first disk shelf. The next higher loop ID is in the next bay to the left, and so on. You can view the device map for your disk shelves with the `fcadmin device_map` command.

For more information about the loop ID map for your disk shelf, see the hardware guide for the disk shelf.

Understanding RAID disk types

Data ONTAP classifies disks as one of four types for RAID: data, hot spare, parity, or dParity. The RAID disk type is determined by how RAID is using a disk.

Data disk Holds data stored on behalf of clients within RAID groups (and any data generated about the state of the storage system as a result of a malfunction).

Spare disk	Does not hold usable data, but is available to be added to a RAID group in an aggregate. Any functioning disk that is not assigned to an aggregate but is assigned to a system functions as a hot spare disk.
Parity disk	Stores data reconstruction information within RAID groups.
dParity disk	Stores double-parity information within RAID groups, if RAID-DP is enabled.

How disk sanitization works

Disk sanitization is the process of physically obliterating data by overwriting disks with specified byte patterns or random data so that recovery of the original data becomes impossible. You use the `disk sanitize` command to ensure that no one can recover the data on the disks.

The `disk sanitize` command uses three successive default or user-specified byte overwrite patterns for up to seven cycles per operation. Depending on the disk capacity, the patterns, and the number of cycles, the process can take several hours. Sanitization runs in the background. You can start, stop, and display the status of the sanitization process.

After you enter the `disk sanitize start` command, Data ONTAP begins the sanitization process on each of the specified disks. The process consists of a disk format operation, followed by the specified overwrite patterns repeated for the specified number of cycles.

Note: The formatting phase of the disk sanitization process is skipped on ATA disks.

If the sanitization process is interrupted by power failure, system panic, or a user-invoked `disk sanitize abort` command, the `disk sanitize` command must be re-invoked and the process repeated from the beginning in order for the sanitization to take place.

When the sanitization process is complete, the specified disks are in a sanitized state. You designate the sanitized disks as spare disks by using the `disk sanitize release` command.

Note: You must install the disk sanitization license before you can perform disk sanitization.

Next topics

Disk sanitization limitations on page 38
What happens if disk sanitization is interrupted on page 38
How selective disk sanitization works on page 39

Related tasks

Removing data from disks using disk sanitization on page 78

Disk sanitization limitations

Installing the disk sanitization license disables certain Data ONTAP commands. In addition, disk sanitization cannot be used with all configurations, models and disk drives.

Installing the disk sanitization license prohibits the following commands from being used on the storage system with that license:

- `dd` (to copy blocks of data)
- `dumpblock` (to print dumps of disk blocks)
- `setflag wafl_metadata_visible` (to allow access to internal WAFL files)

The disk sanitization process has the following limitations:

- It is not supported in takeover mode for systems in an active/active configuration. (If a storage system is disabled, it remains disabled during the disk sanitization process.)
- It cannot be carried out on disks that were failed due to readability or writability problems.
- It cannot be carried out on disks that belong to an SEC 17a-4-compliant SnapLock volume until the expiration periods on all files have expired--that is, all of the files have reached their retention dates.
- It does not perform its formatting phase on ATA drives.
- If you are using the random pattern, it cannot be performed on more than 100 disks at one time.
- It is not supported on array LUNs.
- If you sanitize both SES disks in the same ESH shelf at the same time, you see errors on the console about access to that shelf, and shelf warnings are not reported for the duration of the sanitization. However, data access to that shelf is not interrupted.

What happens if disk sanitization is interrupted

Disk sanitization can take time to complete. If disk sanitization is interrupted by user intervention or an unexpected event such as a power outage, Data ONTAP takes certain actions to prevent corrupted disks if necessary.

If the sanitization process is interrupted by power failure, system panic, or a user-invoked `disk sanitize abort` command, the `disk sanitize` command must be re-invoked and the process repeated from the beginning in order for the sanitization to take place.

If the formatting phase of disk sanitization is interrupted, Data ONTAP attempts to reformat any disks that were corrupted by the interruption. After a system reboot and once every hour, Data ONTAP checks for any sanitization target disk that did not complete the formatting phase of its sanitization. If such a disk is found, Data ONTAP attempts to reformat that disk, and writes a message to the console informing you that a corrupted disk has been found and will be reformatted. After the disk is reformatted, it is designated as a hot spare. You can then rerun the `disk sanitize` command on that disk.

How selective disk sanitization works

Selective disk sanitization consists of physically obliterating data in specified files or volumes while preserving all other data located on the affected aggregate for continued user access. Because a file can be stored on multiple disks, there are three parts to the process.

To selectively sanitize data contained in an aggregate, you must carry out three general tasks:

1. Delete the files, directories or volumes that contain the data you want to sanitize from the aggregate that contains them.
2. Migrate the data that you want to preserve to a new set of disks in a destination aggregate on the same storage system.
   You can migrate data using the `ndmpcopy` command or qtree SnapMirror.
3. Destroy the original aggregate and sanitize all the disks that were RAID group members in that aggregate.

Related tasks

Removing data from disks using selective disk sanitization on page 81

Tips for creating and backing up aggregates containing data that will be sanitized

If you are creating or backing up aggregates to contain data that might need to be sanitized, following some simple guidelines will reduce the time it takes to sanitize your data.

- Make sure your aggregates containing sensitive data are not larger than they need to be.
  If they are larger than needed, sanitization requires more time, disk space, and bandwidth.
- When you back up aggregates containing sensitive data, avoid backing them up to aggregates that also contain large amounts of nonsensitive data.
  This will reduce the resources required to move nonsensitive data before sanitizing sensitive data.

Commands to display disk and array LUN information

You can see information about your disks and array LUNs using several commands, including the `aggr`, `disk`, `fcstat`, `sasadmin`, `storage`, `sysconfig`, and `sysstat` commands.

Use this Data ONTAP command...	**To display information about..**
`aggr status -f`	Disks or array LUNs in your storage system that have failed, or that have been preemptively failed by Data ONTAP.

Use this Data ONTAP command...	To display information about..
`aggr status -m`	Disks in your storage system that are currently in the maintenance center, that have been or are being sanitized, and that are being checked by Data ONTAP due to poor response time.
`aggr status -r`	All disks and array LUNs available in your storage system.
`aggr status -s`	Hot spare disks and spare array LUNs available in your storage system.
`disk maint status`	The status of disk maintenance tests that are in progress.
`disk sanitize status`	The status of the disk sanitization process, after the `disk sanitize start` command has been executed.
`disk shm_stats`	SMART (Self-Monitoring, Analysis, and Reporting Technology) data, disk error information, and log sense information for disks.
`disk show`	List of disks and array LUNs owned by a storage system, or unowned disks and array LUNs. This command is available only for systems using software-based disk ownership.
`fcstat device_map`	A physical representation of where FC-AL attached disks reside in a loop and a mapping of the disks to the disk shelves.
`fcstat fcal_stats`	Error and exceptions conditions, and handler code paths executed.
`fcstat link_stats`	Link event counts.
`sasadmin devstats`	Statistics for SAS-connected disks: command completion counts, frame in and out counts, error and timeout counts.
`sasadmin shelf [short]`	Logical view of SAS shelf (long and short view).
`storage show acp`	The Alternate Control Path (ACP) module. Specifies whether the mode is enabled and displays connectivity and configuration information.
`storage show disk -a`	Detailed information about disks presented in a report form that is easily interpreted by scripts. This content also appears in the STORAGE section of an AutoSupport report.
`storage show disk -p`	Primary and secondary paths to all disks and array LUNs.
`storage show disk -T -x`	The disk type (FCAL, LUN, SATA, and so on) along with the disk and array LUN information.

Use this Data ONTAP command...	**To display information about..**
`storage show disk -x`	The disk ID, shelf, bay, serial number, vendor, model, and revision level of all disks and array LUNs.
`sysconfig -d`	Disk name in the Device column, followed by the expansion slot number, shelf, bay, channel, and serial number.
`sysconfig -h`	Each disk, along with the size displayed in appropriate units (KB, GB, or TB) as calculated using the powers of two. (GB = 1024 × 1024 × 1024)
`sysstat`	The number of kilobytes per second (kB/s) of data being read and written.

Commands to display disk space information

You can see information about how disk space is being used in your aggregates and volumes and their Snapshot copies.

Use this Data ONTAP command...	**To display information about...**
`aggr show_space`	Disk space usage for aggregates
`df`	Disk space usage for volumes or aggregates
`snap delta`	The estimated rate of change of data between Snapshot copies in a volume
`snap reclaimable`	The estimated amount of space freed if you delete the specified Snapshot copies

For more information about the `snap` commands, see the *Data ONTAP Data Protection Online Backup and Recovery Guide*. For more information about the `df` and `aggr show_space` commands, see the appropriate man page.

How Data ONTAP monitors disk performance and health

Data ONTAP continually monitors disks to assess their performance and health. When Data ONTAP encounters certain errors or behaviors from a disk, it takes the disk offline temporarily or takes the disk out of service to run further tests.

Next topics

When Data ONTAP takes disks offline temporarily

Data ONTAP temporarily stops I/O activity to a disk and takes a disk offline when Data ONTAP is updating disk firmware in background mode or when disks become non-responsive. While the disk is offline, Data ONTAP performs a quick check on it to reduce the likelihood of forced disk failures.

While the disk is offline, Data ONTAP reads from other disks within the RAID group while writes are logged. When the offline disk is ready to come back online, Data ONTAP re-synchronizes the RAID group and brings the disk online. This process generally takes a few minutes and incurs a negligible performance impact.

Note: The disk offline feature is only supported for spares and data disks within RAID-DP and mirrored-RAID4 aggregates. A disk can be taken offline only if its containing RAID group is in a normal state and the plex or aggregate is not offline.

How Data ONTAP reduces disk failures using Rapid RAID Recovery

When Data ONTAP determines that a disk has exceeded its error thresholds, Data ONTAP can perform Rapid RAID Recovery by removing the disk from its RAID group for testing and, if necessary, failing the disk. Spotting disk errors quickly helps prevent multiple disk failures and allows problem disks to be replaced.

By performing the Rapid RAID Recovery process on a suspect disk, Data ONTAP avoids three problems that occur during sudden disk failure and the subsequent RAID reconstruction process:

- Rebuild time
- Performance degradation
- Potential data loss due to additional disk failure during reconstruction

During Rapid RAID Recovery, Data ONTAP performs the following tasks:

1. Places the suspect disk in pre-fail mode.

2. Selects a hot spare replacement disk.

   Note: If no appropriate hot spare is available, the suspect disk remains in pre-fail mode and data continues to be served. However, a suspect disk performs less efficiently. Impact on performance ranges from negligible to worse than degraded mode. For this reason, make sure hot spares are always available.

3. Copies the suspect disk's contents to the spare disk on the storage system before an actual failure occurs.
4. After the copy is complete, attempts to put the suspect disk into the maintenance center, or else fails the disk.

Note:

Tasks 2 through 4 can only occur when the RAID group is in normal (not degraded) mode.

If the suspect disk fails on its own before copying to a hot spare disk is complete, Data ONTAP starts the normal RAID reconstruction process.

Related concepts

About degraded mode on page 110
When Data ONTAP can put a disk into the maintenance center on page 44
How Data ONTAP works with hot spare disks on page 109

How the maintenance center works

When a disk is in the maintenance center, it is subjected to a number of tests. If the disk passes all of the tests, it is redesignated as a spare. Otherwise, Data ONTAP fails the disk.

The maintenance center is controlled by the `disk.maint_center.enable` option. It is on by default.

Data ONTAP puts disks into the maintenance center only if there are two or more spares available for that disk.

You can control the number of times a disk is allowed to go to the maintenance center using the `disk.maint_center.allowed_entries` option. The default value for this option is `1`, which means that if the disk is ever sent back to the maintenance center, it is automatically failed.

Data ONTAP informs you of these activities by sending messages to the following destinations:

- The console
- A log file at `/etc/maintenance.log`

Note: When Data ONTAP puts a drive into the maintenance center, and that drive is housed in a disk shelf that supports automatic power cycling, power to that drive might be turned off for a short period of time. If the drive returns to a ready state after the power cycle, the maintenance center tests the drive. Otherwise, the maintenance center fails the drive immediately.

You can see the power-cycle status for ESH4 disk shelves by using the `environment shelf_power_status` command.

For information about best practices for working with the maintenance center, see *Technical Report 3437: Storage Best Practices and Resiliency Guide*.

Note: This technical report contains information about NetApp products that IBM licenses and in some cases customizes. Technical reports might contain information about models and features that are not supported by IBM.

Related information

TR 3437: Storage Best Practices and Resiliency Guide

When Data ONTAP can put a disk into the maintenance center

When Data ONTAP detects certain disk errors, it tries to put the disk into the maintenance center for testing. Certain requirements must be met for the disk to be put into the maintenance center.

If a disk experiences more errors than are allowed for that disk type, Data ONTAP takes one of the following actions:

- If the `disk.maint_center.spares_check` option is set to `on` (the default) and two or more spares are available, Data ONTAP takes the disk out of service and assigns it to the maintenance center for data management operations and further testing.
- If the `disk.maint_center.spares_check` option is set to `on` and fewer than two spares are available, Data ONTAP does not assign the disk to the maintenance center. It simply fails the disk and designates the disk as a broken disk.
- If the `disk.maint_center.spares_check` option is set to `off`, Data ONTAP assigns the disk to the maintenance center without checking the number of available spares.

Note: The `disk.maint_center.spares_check` option has no effect on putting disks into the maintenance center from the command-line interface.

How Data ONTAP uses continuous media scrubbing to prevent media errors

The purpose of the continuous media scrub is to detect and correct media errors in order to minimize the chance of storage system disruption due to a media error while a storage system is in degraded or reconstruction mode.

By default, Data ONTAP runs continuous background media scrubbing for media errors on all storage system disks. If a media error is found, Data ONTAP uses RAID to reconstruct the data and repairs the error.

Media scrubbing is a continuous background process. Therefore, you might observe disk LEDs blinking on an apparently idle storage system. You might also observe some CPU activity even when no user workload is present.

Note: You can disable continuous media scrubbing for disks in use in aggregates by using the `raid.media_scrub.enable` option. In addition, you can disable continuous media scrubbing

for spare disks by using the `raid.media_scrub.spares.enable` option. However, you are advised not to disable continuous media scrubbing, especially for SATA or ATA disks and disks used in RAID4 aggregates.

For more information about the `raid.media_scrub` options, see the na_options(1) man page.

Next topics

How continuous media scrub impacts system performance on page 45
Why continuous media scrubbing should not replace scheduled RAID-level disk scrubs on page 45

How continuous media scrub impacts system performance

Because continuous media scrubbing searches only for media errors, the impact on system performance is negligible. In addition, the media scrub attempts to exploit idle disk bandwidth and free CPU cycles to make faster progress. However, any client workload results in aggressive throttling of the media scrub resource.

If needed, you can further decrease the CPU resources consumed by a continuous media scrub under a heavy client workload by increasing the maximum time allowed for a media scrub cycle to complete. You can do this by using the `raid.media_scrub.rate` option.

Why continuous media scrubbing should not replace scheduled RAID-level disk scrubs

Because the continuous media scrub process scrubs only media errors, you should continue to run the storage system's scheduled complete RAID-level scrub operation. The RAID-level scrub finds and corrects parity and checksum errors as well as media errors.

Related concepts

How you schedule automatic RAID-level scrubs on page 114

Increasing storage availability by using ACP

ACP, or Alternate Control Path, is a protocol that enables Data ONTAP to manage and control a SAS disk shelf storage subsystem. It uses a separate network (alternate path) from the data path, so management communication is not dependent on the data path being intact and available.

You do not need to actively manage the SAS disk shelf storage subsystem. Data ONTAP automatically monitors and manages the subsystem without operator intervention. However, you must provide the required physical connectivity and configuration parameters to enable the ACP functionality.

Note: You can install SAS disk shelves without configuring ACP. However, for maximum storage availability and stability, you should always have ACP configured and enabled.

After you enable ACP, you can use the `storage show acp` and `acpadmin list all` commands to display information about your ACP subsystem.

Because ACP communication is on a separate network, it does not affect data access in any way.

Enabling ACP

ACP can increase your storage availability when you use SAS disk shelves. If your storage system model has a dedicated port for ACP, then ACP is enabled by default, and you do not need to explicitly enable ACP.

Before you begin

- Is the ACP subnet cabled on an isolated network, with no switches or hubs?
  For more information, see the *Hardware and Service Guide* for your disk shelf.
- Have you identified a port that is not in use by any other subsystem?
- If you are configuring ACP for disk shelves attached to an active/active configuration, have you recorded the domain name and network mask to ensure that they are the same for both nodes?

About this task

The ACP subnet is a private Ethernet network that enables the ACP processor in the SAS module to communicate both with Data ONTAP and with the SAS IOMs in the disk shelves.

The ACP subnet is separate from the I/O data path that connects the disk shelves to the HBA on the storage controller. When you configure ACP on one of the system's network interfaces, you must supply a private domain name that conforms to the standard for private internet addresses (RFC1918). You can use the system default domain, 198.15.1.0, or another network name (that is, an IP address ending in 0) that conforms to the standard.

Steps

1. Ensure that the port you are assigning to ACP is not in use by any other subsystem by entering the following command:

   `7-mode command here`

   You should not see blah blah....need more info here....

2. At the Data ONTAP command line, enter the following command:

   `options acp.enabled on`

   If you have not previously configured the networking information for ACP, you are prompted for that information. When you select a domain name and network mask for the ACP interface, Data ONTAP automatically assigns IP addresses for the ACP interface on the storage controller and both I/O modules on each disk shelf on the ACP subnet.

3. You can verify your ACP connectivity by entering the following command:

`sysconfig -v`

The ACP Connectivity Status should show "Full Connectivity".

Example

For example, if you select e0b as the interface for ACP traffic, 198.15.1.0 as the ACP domain, and 255.255.255.0 as the network mask for the ACP subnet, the `storage show acp` command output looks similar to the following example:

```
my-sys-1> storage show acp

 Alternate Control Path:  enabled
 Ethernet Interface:      e0b
 ACP Status:              Active
 ACP IP address:          198.168.1.16
 ACP domain:              198.168.0.1
 ACP netmask:             255.255.252.0
 ACP Connectivity Status: Full Connectivity

Shelf Module      Reset Cnt    IP address      FW Version    Module
Type   Status
-------------------------------------------------------------------------
-----------
7a.001.A          002          198.15.1.145    01.05
IOM6          active
7a.001.B          003          198.15.1.146    01.05
IOM6          active
7c.002.A          000          198.15.1.206    01.05
IOM6          active
7c.002.B          001          198.15.1.204    01.05
IOM6          active
```

After you finish

If you want to change your ACP configuration values later, you can use the `setup` command.

How ownership for disks and array LUNs works

Disk and array LUN ownership determines which node owns a disk or array LUN and what pool a disk or array LUN is associated with. Understanding how ownership works enables you to maximize storage redundancy and manage your hot spares effectively.

In a stand-alone storage system that does not use SyncMirror, ownership is simple—each disk or array LUN is assigned to the single controller and is in pool0. However, the following situations are more complicated:

- Active/active configurations (two controllers are involved)
- SyncMirror is in use (two pools are involved)

Disk ownership can be hardware-based or software-based, depending on your storage system model.

Next topics

How software-based ownership works on page 49
How hardware-based disk ownership works on page 56
How disks are assigned to spare pools when SyncMirror is enabled on page 57
Storage system models that support hardware-based ownership or both types on page 57

Related concepts

Managing array LUNs through Data ONTAP on page 87

Related references

Storage system models that support hardware-based ownership or both types on page 57

How software-based ownership works

Software-based ownership information is stored on the disk or array LUN rather than determined by the topology of the storage system's physical connections. It gives you increased flexibility and control over your storage configuration.

Software-based ownership requires that you take a more active role in managing ownership. For example, when you add disks or disk shelves to an existing storage system that uses software-based disk ownership, you might need to explicitly assign ownership of the new disks if Data ONTAP is unable to do so. When you make array LUNs available to Data ONTAP you must explicitly assign ownership.

Next topics

Why you assign ownership of disks and array LUNs on page 50

Related references

Why you assign ownership of disks and array LUNs

Storage system ownership must be assigned for disks and array LUNs before they become an effective part of your system. You must explicitly assign ownership for array LUNs. Disks can be automatically or manually assigned.

You assign ownership of a disk or array LUN to accomplish the following actions:

- Associate the disk or array LUN with a specific storage system.
  For a stand-alone system, all disks and array LUNs are owned by that system. In an active/active configuration, the disks and array LUNs could be owned by either system.
- Enable the disk or array LUN to be used and managed by the system that owns it.
  Unowned disks cannot be used as spares and do not receive the automatic firmware updates that owned disks do.
- Associate the disk or array LUN with a specific SyncMirror pool (when SyncMirror is in use).
  If SyncMirror is not in use, all disks and array LUNs are in pool0.

What it means for Data ONTAP to own an array LUN

Data ONTAP cannot use an array LUN presented to it by a storage array until you have configured a logical relationship in Data ONTAP that identifies a specific system running Data ONTAP as the *owner* of the array LUN.

A storage array administrator creates array LUNs and makes them available to specified FC initiator ports of storage systems running Data ONTAP. (The process for how to do this varies among storage array vendors.) When you assign an array LUN to a system running Data ONTAP, Data ONTAP writes data to the array LUN to identify that system as the *owner* of the array LUN. Thereafter, Data ONTAP ensures that only the owner can write data to and read data from the array LUN.

From the perspective of Data ONTAP, this logical relationship is referred to as *disk ownership* because Data ONTAP considers an array LUN to be a virtual disk. From the perspective of Data ONTAP, you are assigning disks to a storage system.

An advantage of the disk ownership scheme is that you can make changes through the Data ONTAP software that, on typical hosts, must be done by reconfiguring hardware or LUN access controls. For example, through Data ONTAP you can balance the load of requests among a group of systems running Data ONTAP by moving data service from one system to another, and the process is transparent to most users. You do not need to reconfigure hardware or the LUN access controls on

the storage array to change which system running Data ONTAP is the owner and, therefore, servicing data requests.

Attention: The Data ONTAP software-based scheme provides ownership control only for storage systems running Data ONTAP; it does not prevent a different type of host from overwriting data in an array LUN owned by a system running Data ONTAP. Therefore, if multiple hosts are accessing array LUNs through the same storage array port, be sure to use LUN security on your storage array to prevent the systems from overwriting each other's array LUNs.

Array LUN reconfiguration, such as resizing the array LUN, must be done from the storage array. Before such activities can occur, you must release Data ONTAP ownership of the array LUN.

Why you might assign array LUN ownership after installation

For a gateway ordered with disk shelves, you are not required to set up third-party storage during initial installation. For a gateway using only third-party storage, you need to assign only two array LUNs during initial installation.

If you ordered your gateway with disk shelves, you do not need to assign any array LUNs initially because the factory installs the root volume on a disk for you. If you are using only third-party storage, you must configure one array LUN for the root volume and one array LUN as a spare for core dumps during initial installation. In either case, you can assign ownership of additional array LUNs to your system at any time after initial installation.

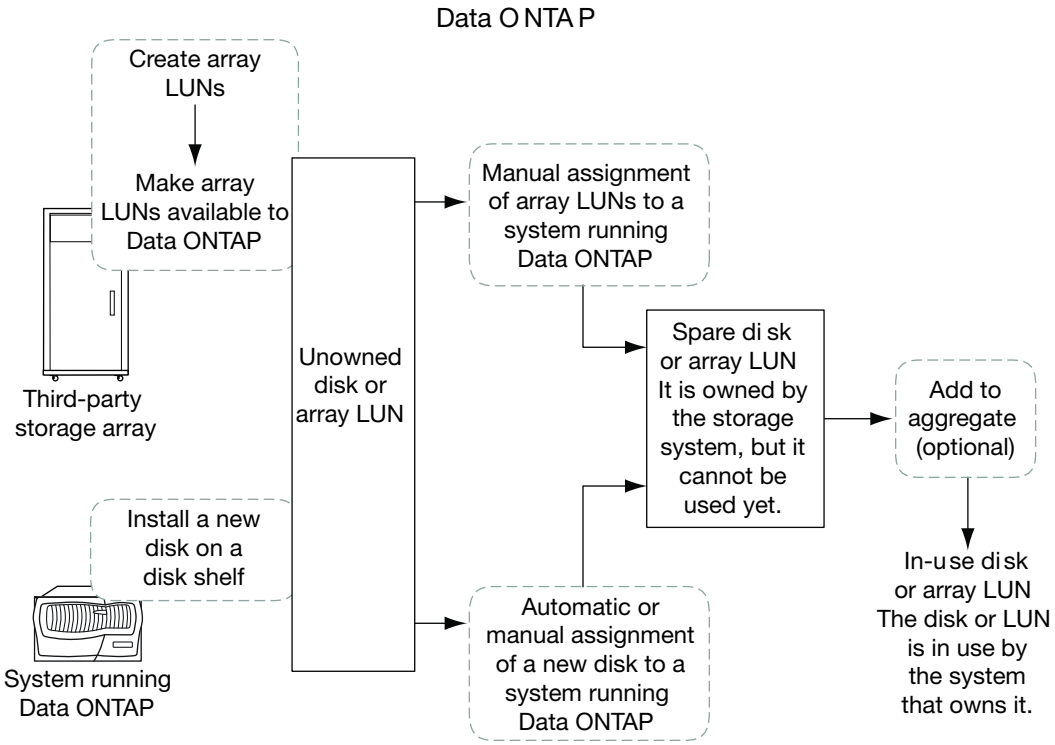
After initial configuration of your system, you might assign ownership of an array LUN in circumstances such as the following:

- You ordered your gateway with native disk shelves and you did not set up your system to work with third-party storage initially
- You left some LUNs that the storage array presented to Data ONTAP unowned and you now need to use the storage
- Another system released ownership of a particular array LUN and you want this system to be able to use the LUN
- The storage array administrator had not made the LUNs available to Data ONTAP when you initially configured your system and you now want to use the storage

How disks and array LUNs become available for use

When you add a disk or array LUN to a system running Data ONTAP, the disk or array LUN goes through several stages before it can be used by Data ONTAP to store data or parity information.

The process for making a disk available for use differs slightly from the process for making an array LUN available for use. Both processes are shown in the following diagram.



The process for disks includes the following actions:

1. The administrator physically installs the disk into a disk shelf.
   Data ONTAP can see the disk but the disk is still unowned.
2. If the system is configured to support disk autoassignment, Data ONTAP assigns ownership for the disk. Otherwise, the administrator must use the `disk assign` command to assign ownership for the disk manually.
   The disk is now a spare disk.
3. The administrator or Data ONTAP adds the disk to an aggregate.
   The disk is now in use by that aggregate. It could contain data or parity information.

The process for array LUNs includes the following actions:

1. The administrator uses the third-party storage array to create the array LUN and make it available to Data ONTAP.
   Data ONTAP can see the array LUN but the array LUN is still unowned.
2. The administrator uses the `disk assign` command to assign ownership for the array LUN.
   The array LUN is now a spare array LUN.
3. The administrator adds the array LUN to an aggregate.

The array LUN is now in use by that aggregate and is used to contain data.

How ownership autoassignment works for disks

If your configuration follows some basic rules to avoid ambiguity, Data ONTAP can automatically assign ownership and pool membership for disks. Autoassignment is not available for array LUNs.

If you decide to change the way Data ONTAP has assigned the disks, you can do so at any time.

Note: You can disable disk autoassignment using the `disk.auto_assign` option. For more information, see the na_option(1) man page.

Next topics

What autoassignment does on page 53
When autoassignment is invoked on page 53

What autoassignment does

When disk autoassignment runs, Data ONTAP looks for any unassigned disks and assigns them to the same system and pool as all other disks on their loop or stack.

Note: If a single loop or stack has disks assigned to multiple systems or pools, Data ONTAP does not perform autoassignment on that loop or stack. To avoid this issue, always follow the disk assignment guidelines.

Related concepts

Guidelines for assigning ownership for disks on page 59
How Data ONTAP works with disks on page 31

When autoassignment is invoked

Disk ownership autoassignment does not happen immediately after disks are introduced into the storage system.

Disk autoassignment is invoked at the following times:

- Every five minutes during normal system operation
- Ten minutes after the initial system initialization

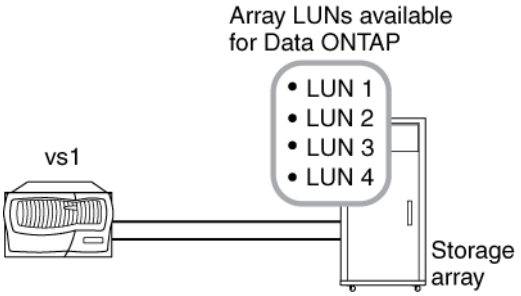
  This delay allows the person configuring the system enough time to finish the initial disk assignments so that the results of the autoassignment are as expected.
- Whenever you enter the `disk assign auto` command.

Examples showing when Data ONTAP can use array LUNs

After an array LUN has been assigned to a storage system, it can be added to an aggregate and used for storage or it can remain a spare LUN until it is needed for storage.

No storage system owns the LUNs yet

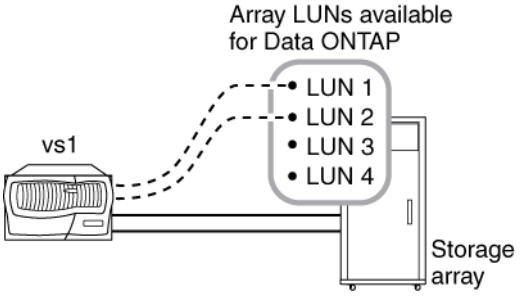
In this example, the storage array administrator made the array LUNs available to Data ONTAP. However, system vs1 has not yet been configured to "own" any of the LUNs. Therefore, it cannot read data from or write data to any array LUNs on the storage array.



Only some array LUNs are owned

In this example, vs1 was configured to own array LUNs 1 and 2, but not array LUNs 3 and 4. LUNs 3 and 4 are still available to Data ONTAP, however, and can be assigned to a storage system later.

Data ONTAP used the smallest of the two array LUNs, LUN 1, for the root volume. System vs1 can read data from and write data to LUN 1, because LUN 1 is in an aggregate. LUN 2 remains a spare LUN because it has not yet been added to an aggregate. System vs1 cannot read data from and write data to LUN 2 while it is a spare.

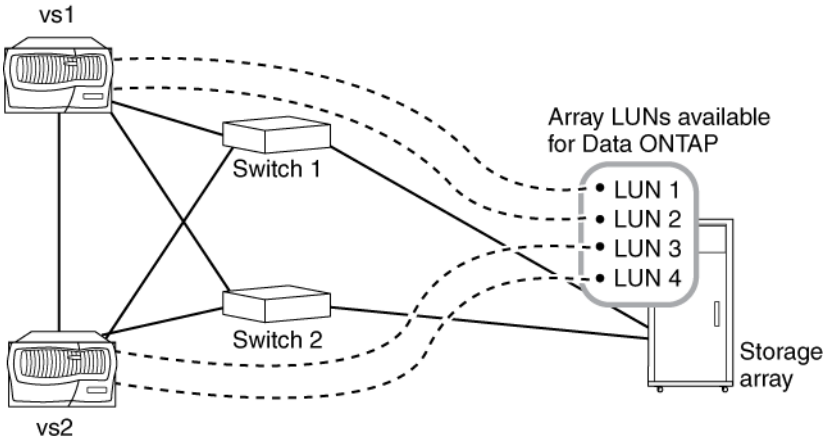


After you perform initial setup of the storage system, you could configure vs1 to also own LUN 3, LUN 4, both, or neither, depending on your storage needs.

Ownership of LUNs in an active/active configuration

In this example, two storage systems running Data ONTAP are configured in an active/active configuration. In an active/active configuration, only one node can be the owner of a particular LUN, but both nodes must be able to see the same LUNs so that the partner can take over if the owning node becomes unavailable.

LUN 1 through LUN 4 were created on the storage array and mapped to the ports on the storage array to which the storage systems are connected. All four LUNs are visible to each node in the active/active configuration.



Assume that during initial setup vs1 was assigned ownership of LUN 1 and LUN 2. LUN 1 was automatically added to the root volume, so LUN 1 is now "in use" by vs1. LUN 2 remains a spare until it is explicitly added to an aggregate on vs1. Similarly, assume that during initial setup vs2 was assigned ownership of LUN 3 and LUN 4, with LUN 3 assigned to the root volume. LUN 4 remains a spare LUN until it is explicitly added to an aggregate.

The key points of this example are as follows:

- By deploying the storage systems in an active/active configuration, one system can take over services for its partner if the partner becomes unavailable.
- Only one storage system can own a specific array LUN. However, all array LUNs assigned to a node in an active/active configuration must be visible to—but not assigned to or owned by—the other node in the active/active configuration.
- By deploying two switches, if one switch fails, the other switch provides the alternate path to the storage array.
- Both switches must be zoned correctly so that each storage system in the active/active configuration can see the array LUNs owned by its partner.

How hardware-based disk ownership works

If your storage system is configured to use hardware-based disk ownership, disk ownership is determined by your hardware configuration. You do not need to assign disk ownership.

Hardware-based disk ownership is determined by two conditions: how a storage system is configured and how the disk shelves are attached to it:

- If the storage system is a stand-alone system, it owns all of the disks directly attached to it.
- If the storage system is part of an active/active configuration, the local node owns all direct-attached disks connected to the local node on the A channel (the loop or stack attached to the A module on the disk shelf) and its partner owns the disks connected to the local node on the B channel.

Note:

The storage system is considered to be part of an active/active configuration if one or more of the following conditions applies:

- An InterConnect card is installed in the system.
- The system has a partner-sysid environment variable.
- The system has the cluster license installed and enabled.

Fabric-attached MetroClusters can use either hardware-based or software-based disk ownership, depending on their hardware configuration. When they use hardware-based disk ownership, the rules for determining disk ownership are different from the rules in a standard active/active configuration. For more information, see the *Data ONTAP Active/Active Configuration Guide*.

Related references

Storage system models that support hardware-based ownership or both types on page 57

Data ONTAP automatically recognizes and assigns disks for hardware-based disk ownership

Because disk ownership and pool membership is determined by the hardware configuration, Data ONTAP automatically performs certain tasks at boot time and when new disks are added.

For all hardware-based disk ownership storage systems, Data ONTAP performs the following functions:

- Recognizes all of the disks at bootup or when they are inserted into a disk shelf.
- Initializes all new disks as spare disks.
- Automatically puts all new disks into a pool.

Note: Spare disks remain spare disks until they are used to create aggregates and are designated as data disks or as parity disks by you or by Data ONTAP.

How disks are assigned to spare pools when SyncMirror is enabled

All spare disks are in pool0 unless the SyncMirror software is enabled. If SyncMirror is enabled on a hardware-based disk ownership storage system, all spare disks are divided into two pools, pool0 and pool1.

For hardware-based disk ownership storage systems, disks are automatically placed in spare pools based on their location in the disk shelves, as follows:

- Pool0—Onboard ports 0a, 0b, and host adapters in expansion slots 1 and 2
- Pool1—Onboard ports 0c, 0d, and host adapters in expansion slots 3 and 4

Storage system models that support hardware-based ownership or both types

Most storage system models support software-based ownership. A few only support hardware-based disk ownership, and a few others support both types.

The following table shows which models support only hardware-based disk ownership or both types of ownership.

Note: The following systems support software-based ownership *only*:

- Systems that contain array LUNs.
- Systems that contain EXN3000 and EXN3500 disk shelves
- Any system model not listed in the following table.

Storage system	Hardware-based	Software-based
N5200	X	X
N5500	X	X

Related concepts

How software-based ownership works on page 49
How hardware-based disk ownership works on page 56

When a storage system uses software-based disk ownership

Some storage systems can use either software-based or hardware-based disk ownership.

If a storage system supports both hardware-based and software-based disk ownership, it uses software-based disk ownership if any of the following criteria are met:

- The storage system is configured to use software-based disk ownership.
- The storage system has the SnapMover license enabled.
- The storage system has disks with software ownership information on them.

Managing ownership for disks and array LUNs

You can display, assign, and modify ownership information for disks and array LUNs.

Next topics

Guidelines for assigning ownership for disks

When you assign ownership for disks, follow these guidelines to keep autoassignment working and to maximize fault isolation.

- Always assign all disks on the same loop or stack to the same system and pool.
- Always assign all loops or stacks connected to the same adapter to the same pool.
- For systems using SyncMirror, pool0 is typically assigned to the local pool and pool1 is assigned to the remote pool.

  For more information about configuring SyncMirror with disks or array LUNs, see the *Data ONTAP Data Protection Online Backup and Recovery Guide*.

Note: You can configure your system to have both pools on a single loop or stack. On storage system models that only support one loop or stack, this configuration cannot be avoided. However, in this configuration, a shelf failure would cause a data service outage.

Displaying ownership information

You use ownership information on systems that use software-based disk ownership to ensure that your hot spares are correctly assigned, or to troubleshoot ownership problems. You view this information with the `disk show` command.

About this task

The `disk show` command is available only for systems using software-based disk ownership.

For more information about the `disk show` command and its options, see the na_disk(1) man page.

The `sysconfig` command can be used to display information about disks and array LUNs, but it does not display disks and array LUNs that are unassigned.

Step

1. Enter the following command to display a list of all the disks and array LUNs visible to the storage system, whether they are owned or not:

   `disk show -v`

   Note: You can display ownership information for a particular disk or array LUN by specifying its name. You can also use the wildcard character (*) to specify multiple disks or array LUNs.

Example ownership display

The following example shows sample output of the `disk show -v` command on an active/active configuration using software-based disk ownership. Disks 0b.16 through 0b.29 are assigned to the system controllers sh1 and sh2. Odd-numbered disks are assigned to sh1 and even-numbered disks are assigned to sh2. The fourteen disks on the add-on disk shelf are still unassigned to either system controller.

```
sh1> disk show -v

 DISK      OWNER            POOL   SERIAL NUMBER
 --------- ---------------  -----  -------------
 0b.43     Not Owned        NONE   41229013
 0b.42     Not Owned        NONE   41229012
 0b.41     Not Owned        NONE   41229011
 0b.40     Not Owned        NONE   41229010
 0b.39     Not Owned        NONE   41229009
 0b.38     Not Owned        NONE   41229008
 0b.37     Not Owned        NONE   41229007
 0b.36     Not Owned        NONE   41229006
 0b.35     Not Owned        NONE   41229005
 0b.34     Not Owned        NONE   41229004
 0b.33     Not Owned        NONE   41229003
 0b.32     Not Owned        NONE   41229002
 0b.31     Not Owned        NONE   41229001
 0b.30     Not Owned        NONE   41229000
 0b.29     sh1 (84165672)   Pool0  41226818
 0b.28     sh2 (84165664)   Pool0  41221622
 0b.27     sh1 (84165672)   Pool0  41226333
 0b.26     sh2 (84165664)   Pool0  41225544
 0b.25     sh1 (84165672)   Pool0  41221700
 0b.24     sh2 (84165664)   Pool0  41224003
 0b.23     sh1 (84165672)   Pool0  41227932
 0b.22     sh2 (84165664)   Pool0  41224591
 0b.21     sh1 (84165672)   Pool0  41226623
 0b.20     sh2 (84165664)   Pool0  41221819
 0b.19     sh1 (84165672)   Pool0  41227336
 0b.18     sh2 (84165664)   Pool0  41225345
```

```
0b.17     sh1 (84165672)   Pool0  41225446
0b.16     sh2 (84165664)   Pool0  41201783
```

Related concepts

How ownership for disks and array LUNs works on page 49
How you use the wildcard character with the disk command on page 64

Assigning ownership for disks and array LUNs

On systems using software-based disk ownership, disks and array LUNs must be owned by a storage system before they can be used in an aggregate. If your system is not configured for ownership autoassignment, or if your system contains array LUNs, you must assign ownership manually.

Before you begin

If you plan to use SyncMirror with third-party storage, you should install the SyncMirror license before assigning the array LUNs you plan to mirror. If you install the SyncMirror license after the array LUNs are assigned to a system, you must unassign the LUNs you want to use in the second plex, then assign them to the system again and specify that they are in pool1.

About this task

Use this procedure to assign ownership of disks and array LUNs that are currently unowned. If you want to change the ownership of disks or array LUNs that are already owned by a system, use the procedure for changing ownership for disks and array LUNs.

Steps

1. Use the `disk show -n` command to view all disks and array LUNs that do not have assigned owners.

   Note: You must make array LUNs available to Data ONTAP before they can be assigned to a system.

2. Use the following command to assign the disks and array LUNs that are labeled `Not Owned` to a storage system.

   `disk assign {`*`disk_list`*` | all | [-T `*`storage_type`*`] -n `*`count`*`|auto} [-c block | zoned] [-o `*`owner_name`*`] [-s `*`sysid`*`] [-f] [-p `*`pool`*`]`

   You can specify the disks and array LUNs to be assigned in the following ways:

   - Use the *`disk_list`* parameter to specify one or more individual disk or array LUN names.
     This is the most specific way to specify disks and array LUNs. However, you have to manually enter each disk name.

- Use the *disk_list* parameter with the wildcard character (*) to specify a group of disks or array LUN names.
- Use the `all` keyword to specify all unowned disks and array LUNs.
- Use the `-n` *count* option to specify a number of unassigned disks and array LUNs to be assigned
- Use the `auto` option to initiate autoassignment.

  Note: Only disks installed in loops or stacks that conform to the autoassignment guidelines will be affected by autoassignment. Array LUNs are not affected by autoassignment.

You use the following options to further qualify which disks and array LUNs Data ONTAP assigns:

- The `-T` option specifies a specific type of disk or array LUN to be assigned: `ATA`, `FCAL`, `SAS`, `SATA`, or `LUN`. The `LUN` disk type is used for array LUNs.

  Note:

  If you have different disk types or disks and array LUNs on your system, always use the `-T` option to ensure that Data ONTAP uses the disks or array LUNs that you expect. Without this option, Data ONTAP uses the type of disk or array LUN with the most spares.

  This option cannot be used with a list of disk or array LUN names. You must use the `-n` option with the `-T` option.

- The `-c` option specifies the checksum type for the array LUNs to be assigned, block or zoned. The default checksum type is block. For more information about checksums, see the *Gateway Installation Requirements and Reference Guide*.

  This option is not used for disks.

You use the following options to specify the system to own the disks and array LUNs you are assigning.

Note: If you do not specify a system to own the disks and array LUNs, they are assigned to the local system.

- The `-o` *owner_name* option specifies the name of the system to which you want to assign the disks and array LUNs.
- The `-s` *sysid* option specifies the ID of the system that the disks and array LUNs are assigned to. This is an alternative to specifying the system name using the `-o` option.
- The `-f` option is used only for changing ownership for a disk or array LUN that is already owned by a system.

You use the `-p` option to specify which SyncMirror pool the disks and array LUNs are assigned to. Its value is either 0 or 1.

Note: If you do not specify a pool, the disks and array LUNs will be assigned to pool0. You need to specify the pool only if SyncMirror is in use on your system.

3. You can use the `disk show -v` command to verify the assignments that you have just made.

After you finish

If you assigned ownership for array LUNs, you should verify that two paths exist for each array LUN and verify path failover to ensure that you have path redundancy.

Related concepts

How ownership for disks and array LUNs works on page 49
How disks and array LUNs become available for use on page 51
What disk types Data ONTAP supports on page 31
How you use the wildcard character with the disk command on page 64

Modifying assignment of spare disks or array LUNs

You can change the ownership of a *spare* disk or array LUN to another storage system.

Before you begin

A disk or array LUN that is a spare has been assigned to a specific system, but it has not yet been added to an aggregate. If the disk or array LUN whose ownership you want to change is in an aggregate, you must do the following before you can change ownership of the disk or array LUN:

- For an array LUN that is part of an aggregate, you must first remove the LUN from the aggregate, which changes the state of the array LUN to spare. To remove an array LUN from an aggregate, you must destroy the aggregate.
- For a disk that is part of an aggregate, you must first perform a disk replace and make the disk a spare.

About this task

You can change ownership of disks only between nodes in an active/active configuration. You can change ownership of array LUNs among the systems in a gateway neighborhood.

Steps

1. At the console of the storage system that owns the disk or array LUN that you want to reassign, enter the following to see a list of spare disks or spare array LUNs on the system:

   `aggr status -s`

2. On the system that owns the spare disk or array LUN you want to reassign, enter either of the following commands to reassign ownership of the disk or array LUN:

   `disk assign LUN-or-disk-name -o *new_owner_name* -f`
   or

disk assign LUN-or-disk-name -s *sysID-of-receiving_system* -f

`-o` is the name of the system that you want to be the new owner of the disk or array LUN.

`-s` is the ID of the system that you want to be the new owner of the disk or array LUN. You can obtain the system ID of the destination system by running `sysconfig` on the destination system.

`-f` is required to force the change.

3. Enter the following command to verify that the ownership of the spare disk or array LUN moved to the other system:

   aggr status -s

   The spare disk or array LUN that you moved should no longer appear in the list of spares.

4. On the destination system, enter the following command to verify that the spare disk or spare array LUN whose ownership you changed is listed as a spare owned by the destination system:

   aggr status -s

After you finish

If you changed ownership for array LUNs, you should verify that two paths exist for each array LUN and verify path failover to ensure that you have path redundancy. You must add the disk or array LUN to an aggregate before you can use it for storage.

Related concepts

How ownership for disks and array LUNs works on page 49

How you use the wildcard character with the disk command

You can use the wildcard character ("*") when you use certain commands to manage disk ownership. However, you need to be sure you understand how Data ONTAP expands the wildcard character.

You can use the wildcard character with the following commands:

- `disk show`
- `disk assign`
- `disk remove_ownership`

When you use the wildcard character with these commands, Data ONTAP expands it with zero or more characters to create a list of disk names that will be operated on by the command. This can be very useful when you want to assign all of the disks attached to a particular port or switch, for example.

Note: Be careful when you use the wildcard character. It is accepted anywhere in the disk name string, and is a simple string substitution. You might get unexpected results.

For example, to assign all disks on port 1 of the switch brocade23 to pool0, you would use the following command:

```
disk assign brocade23:1.* -p 0
```

However, if you left off the second ".", as in the following command, you would assign all disks attached to ports 1, 10, 11, 12, and so on:

```
disk assign brocade23:1* -p 0
```

Assigning multiple disks attached to an HBA

To assign all of the disks attached to the B port of the HBA in expansion slot 5 to pool0, use the following command:

```
disk assign 5b.* -p 0
```

Determining whether a system has hardware-based or software-based disk ownership

You need to know which type of disk ownership your storage system is using to properly administer your hot spare disks. You find this out using the `storage show` command.

About this task

Some storage system models can use either software-based or hardware-based disk ownership.

Step

1. Enter the following command:

   ```
   storage show
   ```

   If the system is using hardware-based disk ownership, the last line of the output is:
   `SANOWN not enabled.`
   Otherwise, the system is using software-based disk ownership.

Related concepts

How ownership for disks and array LUNs works on page 49

Related references

Storage system models that support hardware-based ownership or both types on page 57

Changing between hardware-based and software-based ownership

Hardware-based and software-based ownership have different strengths and tradeoffs. If your storage system supports both, you can change the type of disk ownership. In some cases you can do so without affecting data availability.

About this task

The steps you use to change between hardware-based and software-based disk ownership depend on whether you are going from hardware-based to software-based or vice-versa, and whether your storage system is in an active/active configuration or is a stand-alone system.

Note: Array LUNs support only software-based ownership. If your storage system includes a third-party storage array, the system must be configured to use software-based ownership.

Next topics

Changing from hardware-based to software-based disk ownership nondisruptively on page 66
Changing from hardware-based to software-based disk ownership using the standard method on page 68
Changing from software-based to hardware-based disk ownership for stand-alone systems on page 68
About changing from software-based to hardware-based disk ownership for active/active configurations on page 69

Related concepts

How ownership for disks and array LUNs works on page 49

Changing from hardware-based to software-based disk ownership nondisruptively

You can change an active/active configuration from hardware-based disk ownership to software-based disk ownership without having to take both nodes of an active/active configuration offline at the same time. Future versions of Data ONTAP will not support hardware-based disk ownership.

Before you begin

Familiarize yourself with the procedure for a nondisruptive upgrade for an active/active configuration. Even though the procedure to change to software-based disk ownership does not include upgrading your Data ONTAP software, it requires a reboot for the changes to take effect. You use the same method for rebooting the nodes as you do for a nondisruptive upgrade.

Attention: This procedure is disruptive for CIFS clients.

This procedure works only for storage systems in an active/active configuration.

About this task

You must complete several of the steps in this procedure on both systems for the procedure to complete successfully.

Steps

1. Disable failover on the system by entering the following command:

   ```
   cf disable
   ```

2. Enter advanced mode *on both controllers* by entering the following command at both system prompts:

   ```
   priv set advanced
   ```

3. Write the software ownership information *on both controllers* to the disk by entering the following command at both system prompts:

   ```
   disk upgrade_ownership
   ```

4. Return to normal mode on both controllers by entering the following command at both system prompts:

   ```
   priv set
   ```

5. On the system that you disabled failover on, re-enable failover by entering the following command:

   ```
   cf enable
   ```

6. Follow the steps to reboot an active/active configuration nondisruptively as described in the documentation for nondisruptive upgrades.

   For more information about nondisruptive upgrades, see the *Data ONTAP Upgrade Guide*.

Related information

IBM NAS documentation and support site - www.ibm.com/storage/support/nas

Changing from hardware-based to software-based disk ownership using the standard method

If you can schedule downtime for an active/active configuration or if you are updating a single system, you can use the standard method to change disk ownership.

About this task

If you are converting an active/active configuration to software-based disk ownership, you must convert both nodes at the same time.

Steps

1. Boot the storage system into maintenance mode.

   For more information about maintenance mode, see the *Data ONTAP System Administration Guide*.

2. Enter the following command:

   ```
   disk upgrade_ownership
   ```

   The system is converted to use software-based disk ownership. In addition, Data ONTAP assigns all the disks to the same system and pool they were assigned to for the hardware-based disk ownership.

3. Halt the system and reboot to normal mode.

Changing from software-based to hardware-based disk ownership for stand-alone systems

If a system is using software-based ownership, and it supports hardware-based disk ownership, you can convert the system to use hardware-based disk ownership.

About this task

Attention:

Do not use this procedure for systems in an active/active configuration.

Do not use this procedure for systems that include third party storage arrays. They do not support hardware-based disk ownership.

Do not use this procedure for systems that include EXN3000 or EXN3500 disk shelves.

Steps

1. If you are using SyncMirror, use the `disk show` command to determine whether your physical cabling conforms to the pool rules for your system.

2. If you found discrepancies between the software ownership and the physical cabling configuration, note those discrepancies and what disks or HBAs you need to move or recable to conform to the hardware-based pool rules.

   Note: Do not make any changes to your cabling configuration yet.

3. Boot the system into maintenance mode.

   For more information, see the section on booting the storage system in the *System Administration Guide*.

4. Enter the following commands to disable software-based ownership:

   ```
   storage release disks
   ```

   ```
   disk remove_ownership all
   ```

5. If you determined previously that any cabling or configuration changes needed to be made to conform to the pool rules, make those changes now.

6. Boot the system into normal mode and verify your configuration using the `aggr status -r` command.

   Data ONTAP will automatically recognize and assign disks.

Related concepts

How disks are assigned to spare pools when SyncMirror is enabled on page 57

Related tasks

Displaying ownership information on page 59

About changing from software-based to hardware-based disk ownership for active/active configurations

Converting an active/active configuration from software-based disk ownership to hardware-based disk ownership is a complicated process. If done incorrectly, your system might not boot. You are advised to contact technical support for assistance with this conversion.

The complexities arise because Data ONTAP cannot automatically provide the correct configuration the way it can when converting from hardware to software. If you do not make all of the required cabling or configuration changes correctly, you could be unable to boot your configuration.

Reusing disks configured for software-based disk ownership

If you plan to reuse disks from storage systems that have been configured for software-based disk ownership, you should remove the software information from the disks first.

Attention: If disks with unerased software-based ownership information are installed in an unbooted storage system that does not use software-based disk ownership, the storage system will not boot.

You do so by transferring the disks to the target storage system while that storage system is in operation, thus automatically erasing their disk ownership information.

Note: If you accidentally cause a boot failure by installing software-assigned disks, you can recover by running the `disk remove_ownership` command in maintenance mode.

Next topics

Automatically erasing disk ownership information on page 70
Recovering from accidental conversion to software-based disk ownership on page 71

Related concepts

How ownership for disks and array LUNs works on page 49

Automatically erasing disk ownership information

If you physically transfer disks from a storage system that uses software-based disk ownership to a running storage system that does not, you can do so without using the `disk remove_ownership` command .

Steps

1. Do not shut down the target storage system.
2. Enter the following command for each of the disks you plan to remove to spin down the disks:

   `disk remove disk_name`
3. Remove the disks from their original storage system and physically install them in the running target storage system.

Result

The running target storage system automatically erases any existing software-based disk ownership information on the transferred disks.

After you finish

On the target storage system, you can use the `aggr status -r` command to verify that the disks you have added are successfully installed.

Recovering from accidental conversion to software-based disk ownership

If you move disks from a system using software-based disk ownership to one using hardware-based disk ownership without taking the appropriate steps, you can inadvertently cause the target system to start using software-based disk ownership.

About this task

Accidental conversion to software-based disk ownership can occur in the following circumstances:

- You do not remove software-based disk ownership information from the target disks before you remove them from their original storage system.
- You add the disks to a target storage system that does not use software-based disk ownership while the target storage system is off.

Under these circumstances, if you boot the target storage system in normal mode, the remaining disk ownership information causes the target storage system to convert to a misconfigured software-based disk ownership setup. It will fail to boot.

Steps

1. Boot the system into maintenance mode

   For more information, see the *System Administration Guide*.

2. In maintenance mode, enter the following command:

   `disk remove_ownership all`

   The software-based disk ownership information is erased from all disks.

3. Return the storage system to normal mode.

Managing disks

You can add and remove disks, sanitize them, and display information about them. These tasks help you use your disks efficiently.

Next topics

Adding disks to a storage system

You add disks to a storage system to increase the number of hot spares, to add space to an aggregate, or to replace disks.

Before you begin

Before adding new disks to the storage system, confirm that the storage system supports the type of disk you want to add. For the latest information on supported disk drives, see the appropriate hardware and service guide.

Steps

1. Install one or more disks according to the hardware guide for your disk shelf or the hardware and service guide for your storage system.

   For storage systems using software-based disk ownership, the new disks are not recognized until they are assigned to a system and pool. You can assign the new disks manually, or you can wait for Data ONTAP to automatically assign the new disks if your system follows the rules for disk autoassignment.

   For storage systems using hardware-based disk ownership, Data ONTAP displays a message confirming that one or more disks were added and then recognizes the disks as hot spare disks.

2. After the new disks have all been recognized, verify their addition, and (if your system is using software-based disk ownership) their ownership information, by entering the following command:

   ```
   disk show -v
   ```

You should see the new disks, owned by the correct system and in the correct pool, listed as hot spare disks.

3. (Optional) You can zero the newly added disks now, if needed, by entering the following command:

   `disk zero spares`

   Note: Disks that have been used previously in a Data ONTAP aggregate must be zeroed before they can be added to another aggregate. Zeroing the disks now can prevent delays in case you need to quickly increase the size of an aggregate. The `disk zero` command runs in the background and can take hours to complete, depending on the size of the unzeroed disks in the system.

Result

The new disks are ready to be added to an aggregate, replace an existing disk, or remain available as hot spares.

Related concepts

Guidelines for assigning ownership for disks on page 59
How ownership autoassignment works for disks on page 53
How Data ONTAP works with disks on page 31

Replacing disks that are currently being used in an aggregate

You can use the `disk replace` command to replace disks that are part of an aggregate without disrupting data service. You do this to swap out mismatched disks from a RAID group. Keeping your RAID groups homogenous helps optimize storage system performance.

Before you begin

You should already have an appropriate hot spare disk of the correct type, size, speed and checksum type installed in your storage system. This spare must be assigned to the same system and pool as the disk it will replace.

About this task

If you need to replace a disk—for example a mismatched data disk in a RAID group—you use the `disk replace` command. This command uses Rapid RAID Recovery to copy data from the specified old disk in a RAID group to the specified spare disk in the storage system. At the end of the process, the spare disk replaces the old disk as the new data disk, and the old disk becomes a spare disk in the storage system.

Note: If you replace a smaller disk with a larger disk, the capacity of the larger disk is downsized to match that of the smaller disk; the usable capacity of the aggregate is not increased.

Step

1. Enter the following command:

   `disk replace start [-m] old_disk_name new_spare_name`

   If you need to use a disk that does not match the speed or pool of the other disks in the aggregate, you can use the `-m` option.

   If you need to stop the disk replace operation, you can use the `disk replace stop` command. If you halt a disk replace operation, the target spare disk needs to be zeroed before it can be used as a data disk in another aggregate.

Related concepts

How Data ONTAP works with hot spare disks on page 109
How Data ONTAP works with disks on page 31
Guidelines for assigning ownership for disks on page 59
How ownership autoassignment works for disks on page 53

Related tasks

Adding disks to a storage system on page 73
Assigning ownership for disks and array LUNs on page 61

Converting a data disk to a hot spare

Data disks can be converted to hot spares by destroying the aggregate that contains them. You must convert a data disk to a hot spare before moving it to another storage system.

About this task

Converting a data disk to a hot spare does not change the ownership information for that disk.

Step

1. Destroy the aggregate the contains the disk by entering the following command:

   `aggr destroy aggr_name`

   All disks in use by that aggregate are converted to hot spare disks.

Removing disks from a storage system

How you remove a disk from your storage system depends how the disk is being used. By using the correct procedure, you can prevent unwanted AutoSupport notifications from being generated and ensure that the disk will function correctly if it is reused in another storage system.

About this task

Remember that if you are removing a disk because it has failed or because it is producing excessive error messages, you should not use the disk again in this or any other storage system.

If you are removing a spare disk, and you might use the disk in a storage system running an earlier version of Data ONTAP, be sure you erase the disk ownership information from the disk before removing it from the storage system.

Note: You cannot reduce the number of disks in an aggregate by removing data disks. The only way to reduce the number of data disks in an aggregate is to copy the data and transfer it to a new aggregate that has fewer data disks.

Next topics

Removing a failed disk on page 76
Removing a hot spare disk on page 77
Removing a data disk on page 77

Removing a failed disk

A disk that has already failed is no longer counted by Data ONTAP as a usable disk. You can just physically disconnect the disk from the disk shelf.

Steps

1. Find the disk ID of the failed disk by entering the following command:

   ```
   aggr status -f
   ```

   The ID of the failed disk is shown next to the word `failed`. The location of the disk is shown to the right of the disk ID, in the columns labeled HA, SHELF, and BAY.

2. Remove the disk from the disk shelf, following the instructions in the hardware guide for your disk shelf model.

Removing a hot spare disk

Removing a hot spare disk requires you to remove ownership information and notify Data ONTAP that you are removing the disk to avoid unwanted AutoSupport messages.

Steps

1. Find the disk name of the hot spare disk you want to remove by entering the following command:

   aggr status -s

   The names of the hot spare disks appear next to the word spare. The locations of the disks are shown to the right of the disk name.

2. If the storage system is using software-based disk ownership, remove the software ownership information from the disk by entering the following commands in the specified order:

   priv set advanced

   disk remove_ownership *disk_name*

   priv set

3. Enter the following command to spin down the disk:

   disk remove *disk_name*

4. Wait for the disk to stop spinning.

   See the hardware guide for your disk shelf model to learn about how to tell when a disk stops spinning.

5. Remove the disk from the disk shelf, following the instructions in the hardware guide for your disk shelf model.

Removing a data disk

The only time you should remove a data disk from a storage system is if the disk is not functioning correctly. If you want to remove a data disk so that it can be used in another system, you must convert it to a hot spare disk first.

Steps

1. Determine the name of the disk you want to remove.

   If the disk is reporting errors, you can find the disk name in the log messages that report disk errors. The name is prepended with the word "Disk".

2. Determine the location of the disk you want to remove by entering the following command:

   aggr status -r

   The location of the disk appears to the right of its name, in the columns HA, SHELF, and BAY.

3. If you do not need to remove the disk immediately, enter the following command to pre-fail the disk:

   disk fail -f *disk_name*

   Attention: You must wait for the disk copy to complete before physically removing the disk.

   Data ONTAP pre-fails the specified disk and attempts to create a replacement disk by copying the contents of the pre-failed disk to a spare disk.

   Note: This copy might take several hours, depending on the size of the disk and the load on the storage system.

   If the copy operation is successful, then Data ONTAP fails the disk and the new replacement disk takes its place. If the copy operation fails, the pre-failed disk fails and the storage system operates in degraded mode until the RAID system reconstructs a replacement disk.

4. If you need to remove the disk immediately, enter the following command:

   disk fail -i -f *disk_name*

   `-i` fails the disk immediately.

   Attention: Do not immediately fail a disk unless it is causing immediate performance or availability issues for your storage system. Depending on your storage system configuration, additional disk failures could result in data loss.

   The disk fails and the storage system operates in degraded mode until the RAID system reconstructs a replacement disk.

5. Remove the failed disk from the disk shelf, following the instructions in the hardware guide for your disk shelf model.

Related concepts

About degraded mode on page 110
How Data ONTAP works with disks on page 31

Removing data from disks using disk sanitization

Disk sanitization enables you to erase data from a disk or set of disks so that the data can never be recovered.

Before you begin

Before you can use the disk sanitization feature, you must install the disk sanitization license.

Attention:

After the license for disk sanitization is installed on a storage system, it is permanent, and it prevents certain Data ONTAP commands from being run.

For more information about licenses, see the *System Administration Guide*.

About this task

You can sanitize any disk that has *spare* status.

If your storage system is using software-based disk ownership, you must ensure that the disks you want to sanitize have been assigned ownership. You cannot sanitize unowned disks.

Steps

1. Verify that the disks that you want to sanitize do not belong to a RAID group in any existing aggregate by entering the following command:

   sysconfig -r

   The disks that you want to sanitize should be listed with spare status.

   Note: If the expected disks are not displayed, they have not been assigned ownership. You must assign ownership to a disk before you can sanitize it.

2. Sanitize the specified disk or disks of all existing data by entering the following command:

   disk sanitize start [-p *pattern1*|-r [-p *pattern2*|-r [-p *pattern3*|-r]]] [-c *cycle_count*] *disk_list*

   Attention:

   Do not turn off the storage system, disrupt the storage connectivity, or remove target disks while sanitizing. If sanitizing is interrupted while target disks are being formatted, the disks must be reformatted before sanitizing can finish.

   If you need to abort the sanitization process, you can do so by using the `disk sanitize abort` command. If the specified disks are undergoing the disk formatting phase of sanitization, the abort will not occur until the disk formatting is complete. After the sanitizing is stopped, Data ONTAP displays a message informing you that sanitization was stopped.

   `-p` *`pattern1`* `-p` *`pattern2`* `-p` *`pattern3`* specifies a cycle of one to three user-defined hex byte overwrite patterns that can be applied in succession to the disks being sanitized. The default pattern is three passes, using 0x55 for the first pass, 0xaa for the second pass, and 0x3c for the third pass.

   `-r` replaces a patterned overwrite with a random overwrite for any or all of the passes.

   `-c` *`cycle_count`* specifies the number of times the specified overwrite patterns will be applied. The default value is one cycle. The maximum value is seven cycles.

   `disk_list` specifies a space-separated list of the IDs of the spare disks to be sanitized.

3. To check the status of the disk sanitization process, enter the following command:

   disk sanitize status [*disk_list*]

4. To make sanitized disks available for reuse as spare disks, enter the following command:

   `disk sanitize release` *`disk_list`*

   Data ONTAP designates the specified disks as hot spares.

   Note: Rebooting the storage system or removing and reinserting a disk that has been sanitized causes that disk to be designated as a broken disk.

Result

The specified disks are sanitized and designated as hot spares. The serial numbers of the sanitized disks are written to /etc/sanitized_disks.

Examples

The following command applies the default three disk sanitization overwrite patterns for one cycle (for a total of 3 overwrites) to the specified disks, 8a.6, 8a.7, and 8a.8:

```
disk sanitize start 8a.6 8a.7 8a.8
```

The following command would result in three disk sanitization overwrite patterns for six cycles (for a total of 18 overwrites) to the specified disks:

```
disk sanitize start -c 6 8a.6 8a.7 8a.8
```

After you finish

You can monitor the status of the sanitization process by using the `/etc/sanitized_disks` and `/etc/sanitization.log` files:

- Status for the sanitization process is written to the `/etc/sanitization.log` file every 15 minutes.
- The `/etc/sanitized_disks` file contains the serial numbers of all drives that have been successfully sanitized. For every invocation of the `disk sanitize start` command, the serial numbers of the newly sanitized disks are appended to the file.

You can verify that all of the disks were successfully sanitized by checking the `/etc/sanitized_disks` file.

Related concepts

How disk sanitization works on page 37
How Data ONTAP works with disks on page 31

Removing data from disks using selective disk sanitization

The procedure you use to selectively sanitize data depends on whether your data is contained in FlexVol or traditional volumes.

Next topics

Selectively sanitizing data contained in FlexVol volumes on page 81
Selectively sanitizing data contained in traditional volumes on page 84

Related concepts

How selective disk sanitization works on page 39
How Data ONTAP works with disks on page 31

Selectively sanitizing data contained in FlexVol volumes

To selectively sanitize data contained in FlexVol volumes, you need to migrate any data you want to preserve in the *entire aggregate*, because every disk used by that aggregate must be sanitized.

Before you begin

- You must install a disk sanitization license on your storage system.
- You need enough free space to duplicate the data you want to preserve, plus extra space for overhead. If you have a limited amount of free space, you can decrease the size of the FlexVol volumes after you delete the data you do not want to preserve and before migrating the volume.

Steps

1. Stop any applications that write to the aggregate you plan to sanitize.
2. From a Windows or UNIX client, delete the directories or files whose data you want to selectively sanitize from the active file system. Use the appropriate Windows or UNIX command, for example:

   `rm /nixdir/nixfile.doc`
3. Remove NFS and CIFS access to all volumes in the aggregate.
4. From the Data ONTAP command line, enter the following command to delete all volume Snapshot copies of the FlexVol volumes that contained the files and directories you just deleted:

   `snap delete -V -a` *`vol_name`*

   `vol_name` is the FlexVol volume that contains the files or directories that you just deleted.
5. Note the names of the volumes that contain data you want to preserve.
6. Enter the following command for each volume you want to preserve, noting the total size and space used:

df -g *vol_name*

7. If you do not have sufficient free space to create an aggregate to contain the migrated volumes at their current size, and the volumes have free space, enter the following command for each volume to decrease its size:

   vol size *vol_name new_size*

   Note: The new size must be larger than the used space in the volume.

8. Create an aggregate to which you will migrate the data you did not delete by entering the following command:

   aggr create *dest_vol disks*

   Example

   aggr create nixdestaggr 8@72G

   This new aggregate provides a migration destination that is absolutely free of the data that you want to sanitize.

9. For each FlexVol volume that contains data you want to preserve, enter the following command to create a corresponding FlexVol volume in the new aggregate:

   vol create *dest_vol dest_aggrsize*

   dest_vol is the name of the new FlexVol volume. Use a different name for the new FlexVol volume.

   dest_aggr is the aggregate you just created.

   size must be at least as large as the current size of the FlexVol volume in the aggregate you will sanitize.

   Example
   To create a FlexVol volume to preserve the data in the nixsrcvol volume, which is a little more than 19 GB, you could use the following command:

   vol create nixsrcvol_1 nixdestaggr 20G

   You now have the volumes into which you will copy the data you want to preserve.

10. For each FlexVol volume that contains data you want to preserve, enter the following command to copy the data to the new aggregate:

    ndmpcopy /vol/*src_vol* /vol/*dest_vol*

    src_vol is the FlexVol volume in the aggregate you want to sanitize.

    dest_vol is the new FlexVol volume that you just created that corresponded to the *src_vol* volume.

    Attention: Be sure that you have deleted the files or directories that you want to sanitize from the source volume before you run the `ndmpcopy` command.

Example

`ndmpcopy /vol/nixsrcvol /vol/nixsrcvol_1`

For information about the `ndmpcopy` command, see the *Data ONTAP Data Protection Tape Backup and Recovery Guide*.

All of the data you want to preserve is now contained in the new aggregate.

11. List the disk IDs used by the source aggregate by entering the following command:

`aggr status src_aggr -r`

Example

`aggr status nixsrcaggr -r`

The disks that you will sanitize are listed in the Device column of the `aggr status -r` output.

12. Record the disk IDs you listed in the previous step.

13. For each FlexVol volume in the aggregate you are sanitizing, enter the following commands to take the volume offline and destroy it:

`vol offline src_vol`

`vol destroy src_vol`

14. Enter the following commands to take the source aggregate offline and destroy it:

`aggr offline src_aggr`

`aggr destroy src_aggr`

The volumes and aggregate that housed the data you want to sanitize have been destroyed. The disks used in this aggregate are now hot spares.

15. Enter the following command to rename the new aggregate, giving it the name of the aggregate that you just destroyed:

`aggr rename dest_aggr old_src_aggr_name`

Example

`aggr rename nixdestaggr nixsrcaggr`

16. For each FlexVol volume in the new aggregate, enter the following command to rename the FlexVol volume to the name of the original FlexVol volume:

`vol rename dest_vol old_src_vol_name`

Example

`vol rename nixsrcvol_1 nixsrcvol`

17. Reestablish your CIFS or NFS services.

- If the original volume supported CIFS services, restart the CIFS services on the volumes in the destination aggregate after migration is complete.
- If the original volume supported NFS services, enter the following command:

```
exportfs -a
```

Users who were accessing files in the original volume will continue to access those files in the renamed destination volume with no remapping of their connections required.

18. Follow the procedure for sanitizing disks on the disks that belonged to the source aggregate.

Related tasks

Removing data from disks using disk sanitization on page 78

Selectively sanitizing data contained in traditional volumes

To selectively sanitize data contained in traditional volumes, you migrate any data you want to preserve to a new volume, and then sanitize the disks that contained the old volume.

Before you begin

- You must install a disk sanitization license on your storage system.
- You need enough free space to duplicate the entire traditional volume you are performing the selective sanitization on, regardless of how much data you are deleting before migrating the data.

Steps

1. Stop any applications that write to the volume you plan to sanitize.
2. From a Windows or UNIX client, delete the directories or files whose data you want to selectively sanitize from the active file system. Use the appropriate Windows or UNIX command, such as

   ```
   rm /nixdir/nixfile.doc
   ```

3. Remove NFS and CIFS access to the volume you plan to sanitize.
4. Create a traditional volume to which you will migrate the data you did not delete by entering the following command:

   ```
   aggr create dest_vol -v disks
   ```

   Note: This traditional volume must have a storage capacity equal to or greater than the volume from which you are migrating. It must have a different name; later, you will rename it to have the same name as the volume you are sanitizing.

   #### Example

   ```
   aggr create nixdestvol -v 8@72G
   ```

   This new volume provides a migration destination that is absolutely free of the data that you want to sanitize.

5. From the Data ONTAP command line, enter the following command to delete all volume Snapshot copies of the traditional volume that contained the files and directories you just deleted:

snap delete -V -a *vol_name*

vol_name is the traditional volume that contained the files or directories that you just deleted.

Example

snap delete -V -a nixdestvol

6. Copy the data you want to preserve to the destination volume from the volume you want to sanitize by entering the following command:

   ndmpcopy /vol/*src_vol* /vol/*dest_vol*

   Attention: Confirm that you have deleted the files or directories that you want to sanitize from the source volume before you run the `ndmpcopy` command.

   src_vol is the volume you want to sanitize.

   dest_vol is the destination volume.

   For information about the `ndmpcopy` command, see the *Data ONTAP Data Protection Tape Backup and Recovery Guide*.

   Example

   ndmpcopy /vol/nixsrcvol /vol/nixdestvol

7. List the disks used in the source volume by entering the following command:

   aggr status *src_vol* -r

   Example

   aggr status nixsrcvol -r

   The disks that you will sanitize are listed in the Device column of the `aggr status -r` output.

8. Record the IDs of the disks used in the source volume.

   After that volume is destroyed, you will sanitize these disks.

9. Take the volume you are sanitizing offline and destroy it by entering the following commands:

   aggr offline *src_vol*

   aggr destroy *src_vol*

   Example

   aggr offline nixsrcvol

   aggr destroy nixsrcvol

10. Rename the new volume, giving it the name of the volume that you just destroyed, by entering the following command:

    aggr rename *dest_vol old_src_vol_name*

    Example

    aggr rename nixdestvol nixsrcvol

11. To confirm that the new volume is named correctly, list your volumes by entering the following command:

    aggr status *old_src_vol_name*

12. Reestablish your CIFS or NFS services.

    - If the original volume supported CIFS services, restart the CIFS services on the volumes in the destination aggregate after migration is complete.
    - If the original volume supported NFS services, enter the following command:

      exportfs -a

    Users who were accessing files in the original volume will continue to access those files in the renamed destination volume.

13. Follow the procedure for sanitizing disks to sanitize the disks that belonged to the source volume.

Result

After sanitizing, the data that you removed from the source volume no longer exists anywhere on your storage system and cannot be restored.

Related tasks

Removing data from disks using disk sanitization on page 78

Stopping disk sanitization

You can use the `disk sanitize abort` command to stop an ongoing sanitization process on one or more specified disks. If you use the `disk sanitize abort` command, the specified disk or disks are redesignated as spares.

Step

1. Enter the following command:

   disk sanitize abort *disk_list*

   If the specified disks are undergoing the disk formatting phase of sanitization, the abort will not occur until the disk formatting is complete.

   Data ONTAP displays the message `Sanitization abort initiated.` After the process is stopped, Data ONTAP displays another message for each disk to inform you that sanitization is no longer in progress.

Managing array LUNs through Data ONTAP

Before a storage array administrator can reconfigure an array LUN that was assigned to a gateway, you must remove the information that Data ONTAP wrote to that LUN when it was assigned.

Certain storage management tasks must always be done on the storage array—for example, creating the LUNs, mapping them to Data ONTAP, and reconfiguring them (for example, to resize them). Other storage management tasks are done through Data ONTAP—for example, creating volumes and aggregates. Depending on what you need to do, you might need to coordinate storage management activities with the storage array administrator.

For example, you need to remove information that Data ONTAP has written to a LUN before the storage array administrator can reconfigure the LUN on the storage array to resize it or use it for a different host. The reason is that Data ONTAP disk ownership information still exists in the disk label.

Next topics

Array LUN name format

The array LUN name is a path-based name that includes the devices in the path between the gateway and the storage array.

By looking at the array LUN name as it is displayed in Data ONTAP output, you can identify devices in the path between the storage system and the storage array, ports used, and the LUN identifier that the storage array presents externally for mapping to hosts. The format of the array LUN name depends on whether the system that runs Data ONTAP connects directly to the storage array or whether it connects through a switch.

The format for an array LUN name for a direct-attached configuration is as follows:

adapter.idlun-id

- *adapter* is the adapter on the storage system that runs Data ONTAP.
- *id* is the channel adapter port on the storage array.
- *lun-id* is the array LUN number that the storage array presents to hosts.

- Example: 0a.0L0

The format for an array LUN name for a switch-attached configuration is as follows:

switch-name:port.idlun-id

- *switch-name* is the name of the switch.
- *port* is the switch port that is connected to the target port (the end point).
- *id* is the device ID.
- *lun-id* is the array LUN number that the storage array presents to hosts.
- Example: mcdata3:6.127L0

These names consist of a path component and the SCSI LUN id on that path. For example, in the array LUN name example for a fabric-attached configuration, mcdata3:6.127 is the path component and L0 is the SCSI LUN ID.

On a gateway, there are two names for each LUN because there are two paths to each LUN—for example, mcdata3:6.127L0 and brocade15:6.127L0.

See the *Gateway Installation Requirements and Reference Guide* for details about how to use the array LUN names when you are checking paths to array LUNs.

Why you might change the checksum type of an array LUN

All array LUNs in an aggregate must be the same checksum type. If necessary, you can change the checksum type of an array LUN to be able to add it to an aggregate.

Data ONTAP formats array LUNs in a special way to store checksum information that is used for data integrity checking on READs. The major factor that determines the usable space in an array LUN is the checksum type. For array LUNs, Data ONTAP supports both block (BCS) checksum and zoned (ZCS) checksum types. You specify a checksum type when you assign ownership of an array LUN to a storage system (or accept the default of BCS).

You might need to change the checksum type associated with an array LUN after you have assigned the LUN to a system running Data ONTAP, for example, because your remaining array LUNs are BCS and you want to add them to an aggregate that is ZCS type. Before changing the checksum type of an array LUN, you should review the tradeoffs between performance in certain types of workloads and storage capacity utilization of each checksum type.

- Block checksums

  With block checksums, Data ONTAP reserves 12.5 percent of the space of the array LUN is used for checksum. Data ONTAP uses BCS by default because it provides better performance in certain workloads.
- Zoned checksums

  Zoned checksums have better storage capacity utilization. However, at certain workloads ZCS array LUNs have a performance impact. Random-read intensive workloads are affected the most.

See the *Gateway Installation Requirements and Reference Guide* for more information about checksums. Contact your Sales Engineer for more details about using checksums.

Changing the checksum type of an array LUN

Sometimes you need to change the checksum type that you assigned to an array LUN, for example, because an array LUN that you want to add to an aggregate is a different checksum type than the aggregate.

About this task

For array LUNs, you can change the checksum type of an array LUN from block checksum type (BCS) to zoned checksum type (ZCS), or the reverse. For example, if your remaining array LUNs are BCS type and the aggregate that you need to add them to is ZCS type, you would need to change the checksum type of those LUNs before you can add them to the aggregate.

Note: Data ONTAP automatically assigns a BCS type to native disks. You cannot change the checksum type of native disks.

Steps

1. Enter the following command:

   disk remove -w *LUN-name*

   LUN-name is the name of the array LUN whose checksum type you want to change.

2. Enter the following command:

   disk assign *LUN-name* -c *new_checksum_type*

   LUN-name is the name of the array LUN whose checksum type you want to change.

   new_checksum_type can be `block` or `zoned`.

   The checksum type of the array LUN is changed to the new checksum type you specified.

Prerequisites to reconfiguring a LUN on the storage array

If an array LUN has already been assigned (through Data ONTAP) to a particular storage system, you must ensure that the information Data ONTAP wrote to the LUN is removed before the storage administrator attempts to reconfigure the LUN on the storage array.

When the storage array presents a LUN to Data ONTAP, Data ONTAP collects information about the LUN (for example, its size) and writes that information to the LUN. Data ONTAP cannot dynamically update information that it wrote to an array LUN. Therefore, before the storage array administrator reconfigures a LUN, you must use Data ONTAP to change the state of the LUN to *unused*. (The LUN is unused from the perspective of Data ONTAP.)

While changing the state of the LUN to unused, Data ONTAP does the following:

- Terminates I/O operations to the LUN.
- Removes the label for RAID configuration information and the persistent reservations from the LUN, which makes the array LUN unowned by any gateway.

After you run `disk remove -w` on a LUN, you can do the following on the storage array:

- Remove the mapping of the LUN to Data ONTAP and make the LUN available to other hosts. No Data ONTAP information remains in the LUN.
- Resize the LUN or change its composition.
- If you want Data ONTAP to use the LUN again, present the LUN to Data ONTAP again.
  When the LUN is presented again to Data ONTAP after it is reconfigured, Data ONTAP is aware of the new LUN size or composition. Thereafter, in Data ONTAP you can assign the LUN to a gateway again.

  Note: You need to assign the LUN to a gateway again because all ownership information was removed from the LUN when you ran `disk remove -w`.

Changing array LUN size or composition

Reconfiguration of array LUN size or composition must be done on the storage array. If an array LUN has already been assigned to a storage system running Data ONTAP , you must change the state of the array LUN to unused, through Data ONTAP, before the storage array administrator can reconfigure the array LUN.

Before you begin

If the array LUN that the storage administrator wants to reconfigure is in an aggregate, you must take the aggregate to which the array LUN belongs offline and destroy the aggregate before starting this procedure. Taking the aggregate offline and destroying it changes the array LUN from a data array LUN to a spare array LUN.

About this task

Using the `disk remove -w` command on an array LUN removes the information that Data ONTAP wrote to the array LUN to identify which system running Data ONTAP is the assigned owner of the array LUN. After the ownership information is removed, the array LUN cannot be used by any system running Data ONTAP unless the array LUN is assigned again to a system.

Steps

1. On the system running Data ONTAP, enter the following command to remove ownership information:

   `disk remove -w` *`LUNfullname`*

2. On the storage array, complete the following steps:
   a. Unmap (unpresent) the array LUN from the systems in the gateway neighborhood so that they can no longer see the array LUN.
   b. Change the size or composition of the array LUN.
   c. Present the array LUN to the systems running Data ONTAP again.

      At this point, the array LUN is visible to the FC initiator ports to which the array LUN was presented, but it cannot be used by any systems running Data ONTAP yet.
3. On the system that you want to be the owner of the array LUN, use the `disk assign` command to assign the ownership of the array LUN to the storage system.

   You can leave the array LUN as a spare or add it to an aggregate. The array LUN cannot be used for storage until after it has been added to an aggregate.

Removing one array LUN from use by Data ONTAP

If you no longer want to use an array LUN for Data ONTAP, you must remove the information that Data ONTAP wrote to the LUN before you can reconfigure the LUN from the storage array for use by another host.

Before you begin

If the LUN that you no longer want Data ONTAP to use is in an aggregate, you must take the aggregate to which the LUN belongs offline and destroy the aggregate before starting this procedure. Taking an aggregate offline and destroying it changes the LUN from a data LUN to a spare LUN.

About this task

When Data ONTAP sees an array LUN, it writes information that it discovers about that LUN to that LUN. Additionally, Data ONTAP writes ownership information to a LUN when (through Data ONTAP) you assign a particular system to be the owner of the LUN. If you no longer want to use a LUN for Data ONTAP, you must use a Data ONTAP command to remove that information from the LUN before you reconfigure the LUN on the storage array. Otherwise that LUN is not available for other hosts.

Note: If you want a different gateway to own the LUN, use the `disk assign -s` or `disk assign -o` command to reassign the LUN to the other gateway.

Perform this procedure from the command line of your storage system running Data ONTAP.

Step

1. Enter the following command:

   `disk remove -w` *`LUNfullname`*

LUNfullname is the full name of the array LUN.

Removing a storage system using array LUNs from service

You must release the persistent reservations on all array LUNs assigned to the storage system running Data ONTAP before removing the system from service.

About this task

When you assign Data ONTAP ownership of an array LUN, Data ONTAP places persistent reservations (ownership locks) on that array LUN to identify which gateway owns the LUN. If you want the array LUNs to be available for use by other types of hosts, you must remove the persistent reservations that Data ONTAP put on those array LUNs. The reason is that some arrays do not allow you to destroy a reserved LUN if you do not remove the ownership and persistent reservations that Data ONTAP wrote to that LUN.

For example, the Hitachi USP storage array does not have a user command for removing persistent reservations from LUNs. If you do not remove persistent reservations through Data ONTAP before removing the gateway from service, you must call Hitachi technical support to remove the reservations.

Contact Technical Support for instructions about how to remove persistent reservations from LUNs before removing a gateway from service.

Note: If the system that you want to remove is part of an active/active configuration , you must remove the active/active software and interconnect cabling before you can remove the system from service. See the *Data ONTAP Active/Active Configuration Guide* for more information.

Commands to display information about your storage

Data ONTAP provides commands to display information about disks, array LUNs, disk space, and storage subsystems.

Next topics

Commands to display disk and array LUN information on page 39
Commands to display disk space information on page 41
Commands to display storage subsystem information on page 95

Commands to display disk and array LUN information

You can see information about your disks and array LUNs using several commands, including the `aggr`, `disk`, `fcstat`, `sasadmin`, `storage`, `sysconfig`, and `sysstat` commands.

Use this Data ONTAP command...	To display information about..
`aggr status -f`	Disks or array LUNs in your storage system that have failed, or that have been preemptively failed by Data ONTAP.
`aggr status -m`	Disks in your storage system that are currently in the maintenance center, that have been or are being sanitized, and that are being checked by Data ONTAP due to poor response time.
`aggr status -r`	All disks and array LUNs available in your storage system.
`aggr status -s`	Hot spare disks and spare array LUNs available in your storage system.
`disk maint status`	The status of disk maintenance tests that are in progress.
`disk sanitize status`	The status of the disk sanitization process, after the `disk sanitize start` command has been executed.
`disk shm_stats`	SMART (Self-Monitoring, Analysis, and Reporting Technology) data, disk error information, and log sense information for disks.
`disk show`	List of disks and array LUNs owned by a storage system, or unowned disks and array LUNs. This command is available only for systems using software-based disk ownership.

Use this Data ONTAP command...	**To display information about..**
`fcstat device_map`	A physical representation of where FC-AL attached disks reside in a loop and a mapping of the disks to the disk shelves.
`fcstat fcal_stats`	Error and exceptions conditions, and handler code paths executed.
`fcstat link_stats`	Link event counts.
`sasadmin devstats`	Statistics for SAS-connected disks: command completion counts, frame in and out counts, error and timeout counts.
`sasadmin shelf [short]`	Logical view of SAS shelf (long and short view).
`storage show acp`	The Alternate Control Path (ACP) module. Specifies whether the mode is enabled and displays connectivity and configuration information.
`storage show disk -a`	Detailed information about disks presented in a report form that is easily interpreted by scripts. This content also appears in the STORAGE section of an AutoSupport report.
`storage show disk -p`	Primary and secondary paths to all disks and array LUNs.
`storage show disk -T -x`	The disk type (FCAL, LUN, SATA, and so on) along with the disk and array LUN information.
`storage show disk -x`	The disk ID, shelf, bay, serial number, vendor, model, and revision level of all disks and array LUNs.
`sysconfig -d`	Disk name in the Device column, followed by the expansion slot number, shelf, bay, channel, and serial number.
`sysconfig -h`	Each disk, along with the size displayed in appropriate units (KB, GB, or TB) as calculated using the powers of two. (GB = 1024 × 1024 × 1024)
`sysstat`	The number of kilobytes per second (kB/s) of data being read and written.

Commands to display disk space information

You can see information about how disk space is being used in your aggregates and volumes and their Snapshot copies.

Use this Data ONTAP command...	To display information about...
`aggr show_space`	Disk space usage for aggregates
`df`	Disk space usage for volumes or aggregates
`snap delta`	The estimated rate of change of data between Snapshot copies in a volume
`snap reclaimable`	The estimated amount of space freed if you delete the specified Snapshot copies

For more information about the `snap` commands, see the *Data ONTAP Data Protection Online Backup and Recovery Guide*. For more information about the `df` and `aggr show_space` commands, see the appropriate man page.

Commands to display storage subsystem information

You can use the `acpadmin`, `environment`, `fcadmin`, `sasadmin`, `storage show`, and `sysconfig` commands to display information about your storage subsystems.

Note: For detailed information about these commands and their options, see the appropriate man pages.

Use this Data ONTAP command...	To display information about...
`acpadmin list_all`	Alternative Control Path (ACP) processors (EXN3000 and EXN3500 only).
`environment shelf`	Environmental information for each host adapter, including SES configuration and SES path.
`environment shelf_log`	Shelf-specific module log file information, for shelves that support this feature. Log information is sent to the `/etc/log/shelflog` directory and included as an attachment on AutoSupport reports.

Use this Data ONTAP command...	To display information about...
`fcadmin channels`	WWPN information.
`fcadmin device_map`	What disks are on each loop and shelf.
`fcadmin link_state`	How the ports are connected.
`sasadmin expander`	What disks are attached to expander PHYs.
`sasadmin expander_phy_state`	Expander PHY state, dongle state and event counters, PHY statistics.
`sasadmin shelf [short]`	The disks on each shelf (or a specific disk shelf), including a pictorial representation of disk placement (long or short view).
`storage show`	All disks and host adapters on the system.
`storage show acp`	Connectivity and status information for the Alternate Control Path (ACP) module (EXN3000 and EXN3500 only).
`storage show adapter`	FC host adapter attributes, including (as appropriate for the adapter type) a description, firmware revision level, Peripheral Component Interconnect (PCI) bus width, PCI clock speed, FC node name, cacheline size, FC packet size, link data rate, static random access memory (SRAM) parity, state, in use, redundant.
`storage show disk -p`	How many paths are available to each disk.
`storage show expander`	SAS expander attributes, including shelf name, channel, module, shelf ID, shelf UID, IOM state, and the following information for the disks attached to the expander: disk ID, port state, partial path timeout, link rate, invalid word count, running disparity count, PHY reset problem, CRC error count, and PHY change count.
`storage show hub`	Hub attributes: hub name, channel, loop, shelf ID, shelf user ID (UID), term switch, shelf state, ESH state, and hub activity for each disk ID: loop up count, invalid cyclic redundancy check (CRC) count, invalid word count, clock delta, insert count, stall count, util.

Use this Data ONTAP command...	**To display information about...**
`storage show mc`	All media changer devices that are installed in the system.
`storage show port`	Switch ports connected to the system.
`storage show switch`	Switches connected to the system.
`storage show tape`	All tape drive devices attached to the system.
`storage show tape supported [-v]`	All tape drives supported. With -v, information about density and compressions settings is also displayed.
`storage stats tape`	Statistics for all tape drives attached to the system.
`sysconfig -A`	All sysconfig reports, including configuration errors, disks, array LUNs, media changers, RAID details, tape devices, and aggregates.
`sysconfig -m`	Tape libraries.
`sysconfig -t`	Tape drives.

Enabling or disabling a host adapter

A host adapter can be enabled or disabled by using the `storage` command. You disable an adapter to replace hardware components or modules.

About this task

You might want to disable an adapter for the following reasons:

- You are replacing any of the hardware components connected to the adapter.
- You are replacing a malfunctioning I/O module.

You can disable an adapter only if all disks connected to it can be reached through another adapter. After an adapter connected to dual-connected disks has been disabled, the other adapter is not considered redundant; thus, the other adapter cannot be disabled.

Steps

1. Identify the name of the adapter whose state you want to change by entering the following command:

   `storage show adapter`

   The field that is labeled “Slot” lists the adapter name.

2. Enter the following command.

If you want to...	Then use this command
Enable the adapter	**`storage enable adapter`** ***`adapter_name`***
Disable the adapter	**`storage disable adapter`** ***`adapter_name`***

How Data ONTAP uses RAID to protect your data and data availability

RAID protects your data and data availability. Understanding how RAID provides this protection can help you administer your storage systems more effectively.

For native storage, Data ONTAP uses RAID-DP (double-parity) or RAID Level 4 (RAID4) protection to ensure data integrity within a group of disks even if one or two of those disks fail. Parity disks provide redundancy for the data stored in the data disks. If a disk fails (or, for RAID-DP, up to two disks), the RAID subsystem can use the parity disks to reconstruct the data in the drive that failed.

For third-party storage, Data ONTAP stripes across the array LUNs using RAID0. The storage arrays, not Data ONTAP, provide the RAID protection for the array LUNs that they make available to Data ONTAP.

Next topics

RAID protection levels for disks on page 101
RAID protection for third-party storage on page 103
Protection provided by RAID and SyncMirror on page 103
Understanding RAID disk types on page 36
How Data ONTAP RAID groups work on page 106
How Data ONTAP works with hot spare disks on page 109
How Data ONTAP handles a failed disk with a hot spare on page 111
How Data ONTAP handles a failed disk that has no available hot spare on page 113
How RAID-level disk scrubs verify data integrity on page 113

Related tasks

Controlling the performance impact of RAID-level scrubbing on page 120

RAID protection levels for disks

Data ONTAP supports two levels of RAID protection for disks in native disk shelves, RAID-DP and RAID4. RAID-DP can protect against double-disk failures or failures during reconstruction. RAID4 can protect against single-disk failures. You assign RAID level on a per-aggregate basis.

For more information about choosing RAID protection levels, see *Technical Report 3437: Storage Best Practices and Resiliency Guide*.

Note: This technical report contains information about NetApp products that IBM licenses and in some cases customizes. Technical reports might contain information about models and features that are not supported by IBM.

Next topics

What RAID-DP protection is on page 102
What RAID4 protection is on page 102

Related information

TR 3437: Storage Best Practices and Resiliency Guide

What RAID-DP protection is

If an aggregate is configured for RAID-DP protection, Data ONTAP reconstructs the data from one or two failed disks within a RAID group and transfers that reconstructed data to one or two spare disks as necessary.

RAID-DP provides double-parity disk protection when the following conditions occur:

- There is a single-disk or double-disk failure within a RAID group.
- There are media errors on a block when Data ONTAP is attempting to reconstruct a failed disk.

The minimum number of disks in a RAID-DP group is three: at least one data disk, one regular parity disk, and one double-parity (or dParity) disk.

If there is a data-disk or parity-disk failure in a RAID-DP group, Data ONTAP replaces the failed disk in the RAID group with a spare disk and uses the parity data to reconstruct the data of the failed disk on the replacement disk. If there is a double-disk failure, Data ONTAP replaces the failed disks in the RAID group with two spare disks and uses the double-parity data to reconstruct the data of the failed disks on the replacement disks.

RAID-DP is the default RAID type for all aggregates.

What RAID4 protection is

RAID4 provides single-parity disk protection against single-disk failure within a RAID group. If an aggregate is configured for RAID4 protection, Data ONTAP reconstructs the data from a single failed disk within a RAID group and transfers that reconstructed data to a spare disk.

The minimum number of disks in a RAID4 group is two: at least one data disk and one parity disk.

If there is a single data or parity disk failure in a RAID4 group, Data ONTAP replaces the failed disk in the RAID group with a spare disk and uses the parity data to reconstruct the failed disk's data on the replacement disk. If no spare disks are available, Data ONTAP goes into degraded mode and alerts you of this condition.

Attention: With RAID4, if there is a second disk failure before data can be reconstructed from the data on the first failed disk, there will be data loss. To avoid data loss when two disks fail, you can

select RAID-DP. This provides two parity disks to protect you from data loss when two disk failures occur in the same RAID group before the first failed disk can be reconstructed.

Note: Non-disruptive upgrade is not supported for aggregates configured for RAID4. For more information about non-disruptive upgrade, see the *Data ONTAP Upgrade Guide*.

Related concepts

How Data ONTAP handles a failed disk with a hot spare on page 111
How Data ONTAP handles a failed disk that has no available hot spare on page 113
About degraded mode on page 110

RAID protection for third-party storage

Third-party storage arrays provide the RAID protection for the array LUNs they make available to systems running Data ONTAP.

Data ONTAP supports a variety of RAID types used by storage arrays, but imposes restrictions on storage arrays using RAID0 for the LUNs that they make available to Data ONTAP. Data ONTAP uses RAID0 to stripe across the array LUNs, which splits data evenly across two or more array LUNs. Performance is maximized because more disk spindles are used.

RAID0 provides no data protection. Therefore, when creating "RAID groups" on storage arrays, follow the best practices of the storage array vendor to ensure that there is an adequate level of protection on the storage array so that disk failure does not result in loss of data or loss of access to data.

Note: A "RAID group" on a storage array is the arrangement of disks that together form the defined RAID level. Each RAID group supports only one RAID type. The number of disks that you select for a RAID group determines the RAID type that a particular RAID group supports. Different storage array vendors use different terms to describe this entity—RAID groups, parity groups, disk groups, Parity RAID groups, and other terms.

Starting in Data ONTAP 7.3, gateways support native disk shelves as well as third-party storage. Data ONTAP supports RAID4 and RAID-DP on the native disk shelves connected to a gateway but does not support RAID4 and RAID-DP with array LUNs.

See the gateway implementation guide for your vendor to determine whether there are specific requirements or limitations about RAID types in configurations with storage systems running Data ONTAP.

Protection provided by RAID and SyncMirror

Combining RAID and SyncMirror provides protection against more types of disk failures than using RAID alone.

RAID can be used in combination with the Data ONTAP SyncMirror feature, which also offers protection against data loss due to disk or other hardware component failure. SyncMirror protects

against data loss by maintaining two copies of the data contained in the aggregate, one in each plex. Any data loss due to disk failure in one plex is repaired by the undamaged data in the other plex.

Note: SyncMirror can be used to provide mirroring of data in array LUNs on third-party storage arrays. However, Data ONTAP provides only RAID0 for data in array LUNs, which does not provide RAID protection. The RAID protection for array LUNs is provided by the third-party storage array.

For more information about SyncMirror, see the *Data Protection Online Backup and Recovery Guide*.

The following tables outline the differences between using RAID alone and using RAID with SyncMirror.

Table 1: RAID-DP and SyncMirror

Criteria	RAID-DP alone	RAID-DP with SyncMirror
Failures protected against	Single-disk failure Double-disk failure within a single RAID group Multiple-disk failures, as long as no more than two disks within a single RAID group fail	All failures protected against by RAID-DP alone Any combination of failures protected against by RAID-DP alone in one plex, concurrent with an unlimited number of failures in the other plex Storage subsystem failures (HBA, cables, shelf), as long as only one plex is affected
Failures *not* protected against	Three or more concurrent disk failures within a single RAID group Storage subsystem failures (HBA, cables, shelf) that lead to three or more concurrent disk failures within a single RAID group	Three or more concurrent disk failures in a single RAID group on both plexes
Required disk resources per RAID group	*n* data disks + 2 parity disks	2 x (*n* data disks + 2 parity disks)
Performance cost	Almost none	Low mirroring overhead; can improve performance
Additional cost and complexity	None	SyncMirror license and configuration

Table 2: RAID4 and SyncMirror

Criteria	RAID4 alone	RAID4 with SyncMirror
Failures protected against	Single-disk failure Multiple-disk failures, as long as no more than one disk within a single RAID group fails	All failures protected against by RAID4 alone Any combination of failures protected against by RAID4 alone in one plex, concurrent with an unlimited number of failures in the other plex Storage subsystem failures (HBA, cables, shelf), as long as only one plex is affected
Failures *not* protected against	Two or more concurrent disk failures within a single RAID group Storage subsystem failures (HBA, cables, shelf) that lead to two or more concurrent disk failures within a single RAID group	Two or more concurrent disk failures in a single RAID group on both plexes
Required disk resources per RAID group	*n* data disks + 1 parity disk	2 x (*n* data disks + 1 parity disk)
Performance cost	None	Low mirroring overhead; can improve performance
Additional cost and complexity	None	SyncMirror license and configuration

Table 3: RAID0 and SyncMirror

Criteria	RAID0 alone	RAID0 with SyncMirror
Failures protected against	RAID0 does not provide protection against any failures. RAID protection is provided by the RAID implemented on the third-party storage array.	Any combination of array LUN, connectivity, or hardware failures, as long as only one plex is affected

Criteria	RAID0 alone	RAID0 with SyncMirror
Failures *not* protected against	RAID0 does not provide protection against any failures. RAID protection is provided by the RAID implemented on the storage array.	Any concurrent failures that affect both plexes.
Required array LUN resources per RAID group	No extra array LUNs required other than *n* data array LUNs	2 x *n* data array LUNs
Performance cost	None	Low mirroring overhead; can improve performance
Additional cost and complexity	None	SyncMirror license and configuration

Understanding RAID disk types

Data ONTAP classifies disks as one of four types for RAID: data, hot spare, parity, or dParity. The RAID disk type is determined by how RAID is using a disk.

Data disk Holds data stored on behalf of clients within RAID groups (and any data generated about the state of the storage system as a result of a malfunction).

Spare disk Does not hold usable data, but is available to be added to a RAID group in an aggregate. Any functioning disk that is not assigned to an aggregate but is assigned to a system functions as a hot spare disk.

Parity disk Stores data reconstruction information within RAID groups.

dParity disk Stores double-parity information within RAID groups, if RAID-DP is enabled.

How Data ONTAP RAID groups work

A RAID group consists of one or more data disks or array LUNs, across which client data is striped and stored, and up to two parity disks, depending on the RAID level of the aggregate that contains the RAID group.

RAID-DP uses two parity disks to ensure data recoverability even if two disks within the RAID group fail.

RAID4 uses one parity disk to ensure data recoverability if one disk within the RAID group fails.

RAID0 does not use any parity disks; it does not provide data recoverability if any disks within the RAID group fail.

For native storage, Data ONTAP uses RAID-DP or RAID4 groups to provide parity protection. For third-party storage, Data ONTAP uses RAID0 groups to optimize performance and storage utilization. The storage arrays provide the parity protection for third-party storage.

Next topics

How RAID groups are named on page 107
About RAID group size on page 107
Considerations for sizing RAID groups for disks on page 107
Considerations for Data ONTAP RAID groups for array LUNs on page 108

How RAID groups are named

Within each aggregate, RAID groups are named rg0, rg1, rg2, and so on in order of their creation. You cannot specify the names of RAID groups.

About RAID group size

A RAID group has a maximum number of disks or array LUNs that it can contain. This is called its maximum size, or its size. A RAID group can be left partially full, with fewer than its maximum number of disks or array LUNs, but storage system performance is optimized when all RAID groups are full.

Related references

Storage limits on page 345

Considerations for sizing RAID groups for disks

Configuring an optimum RAID group size for an aggregate made up of disks requires a trade-off of factors. You must decide which factor—speed of recovery, assurance against data loss, or maximizing data storage space—is most important for the aggregate that you are configuring.

In most cases, the default RAID group size is the best size for your RAID groups. However, you can change the maximum size of your RAID groups.

Note: You change the size of RAID groups on a per-aggregate basis. You cannot change the size of an individual RAID group.

Configuring an optimum RAID group size for an aggregate requires a trade-off of factors. Adding more data disks to a RAID group increases the striping of data across those disks, which typically improves I/O performance. Additionally, a smaller percentage of disks is used for parity rather than data. However, with more disks in a RAID group, there is a greater risk that one of the disks might fail.

Note: With RAID-DP, you can use larger RAID groups because they offer more protection. A RAID-DP group is more reliable than a RAID4 group that is half its size, even though a RAID-DP group has twice as many disks. Thus, the RAID-DP group provides better reliability with the same parity overhead.

Large RAID group configurations offer the following advantages:

- More data drives available. An aggregate configured into a few large RAID groups requires fewer drives reserved for parity than that same aggregate configured into many small RAID groups.
- Small improvement in storage system performance. Write operations are generally faster with larger RAID groups than with smaller RAID groups.

Small RAID group configurations offer the following advantages:

- Shorter disk reconstruction times. In case of disk failure within a small RAID group, data reconstruction time is usually shorter than it would be within a large RAID group.
- Decreased risk of data loss due to multiple disk failures. The probability of data loss through double-disk failure within a RAID4 group or through triple-disk failure within a RAID-DP group is lower within a small RAID group than within a large RAID group.

Considerations for Data ONTAP RAID groups for array LUNs

Setting up Data ONTAP RAID groups for array LUNs requires planning and coordination with the storage array administrator so that the administrator makes the number and size of array LUNs you need available to Data ONTAP.

For array LUNs, Data ONTAP uses RAID0 RAID groups to determine where to allocate data to the LUNs on the storage array. The RAID0 RAID groups are not used for RAID data protection. The storage arrays provide the RAID data protection.

Note: Data ONTAP RAID groups are similar in concept to what storage array vendors call RAID groups, parity groups, disk groups, Parity RAID groups, and other terms.

Follow these steps when planning your Data ONTAP RAID groups for array LUNs:

1. Plan the size of the aggregate that best meets your data needs.
2. Plan the number and size of RAID groups that you need for the size of the aggregate.

   Follow these guidelines:

   - RAID groups in the same aggregate should be the same size with the same number of LUNs in each RAID group. For example, you should create four RAID groups of 8 LUNs each, not three RAID groups of 8 LUNs and one RAID group of 6 LUNs.
   - Use the default RAID group size for array LUNs, if possible. The default RAID group size is adequate for most organizations.

     Note: The default RAID group size is different for array LUNs and disks.

3. Plan the size of the LUNs that you need in your RAID groups.

   - To avoid a performance penalty, all array LUNs in a particular RAID group should be the same size.
   - The LUNs should be the same size in all RAID groups in the aggregate.

4. Ask the storage array administrator to create the number of LUNs of the size you need for the aggregate.

The LUNs should be optimized for performance, according to the instructions in the storage array vendor documentation.

5. Create all the RAID groups in the aggregate at the same time.

   Note: Do not mix array LUNs from storage arrays with different characteristics in the same Data ONTAP RAID group.

   Note: If you create a new RAID group for an existing aggregate, be sure that the new RAID group is the same size as the other RAID groups in the aggregate, and that the array LUNs are the same size as the LUNs in the other RAID groups in the aggregate.

How Data ONTAP works with hot spare disks

A hot spare disk is a disk that is assigned to a storage system but is not in use by a RAID group. It does not yet hold data but is ready for use. If a disk failure occurs within a RAID group, Data ONTAP automatically assigns hot spare disks to RAID groups to replace the failed disks.

Next topics

How many hot spares you should have

At a minimum, you should have at least one matching or appropriate hot spare available for each kind of disk installed in your storage system. However, having two available hot spares for all disks provides the best protection against disk failure.

Having at least two available hot spares for all disks provides the following benefits:

- At least two hot spares must be available in order to put a disk into the maintenance center.
- Having two hot spares means that when a disk fails, you still have a spare available if another disk fails before you replace the first failed disk.

Note: One disk can be the hot spare for multiple disks.

What disks can be used as hot spares

A disk must conform to certain criteria to be used as a hot spare for a particular data disk.

For a disk to be used as a hot spare for another disk, it must conform to the following criteria:

- It must be either an exact match for the disk it is replacing or an appropriate alternative.

- If SyncMirror is in use, the spare must be in the same pool as the disk it is replacing.
- The spare must be owned by the same system as the disk it is replacing.

What a matching spare is

A matching hot spare exactly matches a data disk for several characteristics.

A matching spare is a disk that exactly matches a data disk for all of the following criteria:

- Type (FC, SAS, ATA, BSAS, or SATA)

  Note: On systems with the `raid.disktype.enable` option set to `off`, FC and SAS disks are considered to be the same type and SATA, ATA, and BSAS disks are considered to be the same type.

- Size
- Speed (RPM)

What an appropriate hot spare is

If a disk fails and no hot spare disk that exactly matches the failed disk is available, Data ONTAP uses the best available spare.

Data ONTAP picks a non-matching hot spare based on the following criteria:

- If the available hot spares are not the correct size, Data ONTAP uses one that is the next size up if possible.

  Note: The replacement disk is downsized to match the size of the disk it is replacing; the extra capacity is not available.

- If the hot spares are not the correct speed, Data ONTAP uses one that is a different speed.

  Note: Using drives with different speeds within the same aggregate is not optimal. Replacing a disk with a slower disk can cause performance degradation, and replacing with a faster disk is not a cost-effective solution.

- If SyncMirror is in use and the hot spares are not in the correct pool, Data ONTAP uses a spare from the other pool.

  Note: Using drives from the wrong pool is not optimal because you no longer have fault isolation for your SyncMirror configuration. Warning messages go to the logs and console to alert you to this issue.

- The hot spare must be of the same disk type (FC, SAS, and so on) as the failed disk, or of a type that is considered to be equivalent.

Related concepts

Disk formats supported by Data ONTAP on page 35

About degraded mode

When a disk fails, Data ONTAP can continue to serve data, but it must reconstruct the data from the failed disk using RAID parity. When this happens, the affected RAID group is said to be in *degraded*

mode. The performance of a storage system with one or more RAID groups in degraded mode is decreased.

A RAID group goes into degraded mode in the following scenarios:

- A single disk fails in a RAID4 group.
  After the failed disk is reconstructed to a spare, the RAID group returns to normal mode.
- One or two disks fail in a RAID-DP group.
  If two disks have failed in a RAID-DP group, the RAID group goes into *double-degraded mode*.
- A disk in a RAID4 group is taken offline by Data ONTAP.
  After the offline disk is brought back online, the RAID group returns to normal mode.

Note: If another disk fails in a RAID-DP group in double-degraded mode or a RAID4 group in degraded mode, data loss could occur (unless the data is mirrored). For this reason, always minimize the amount of time a RAID group is in degraded mode by ensuring that appropriate hot spares are available.

About low spare warnings

By default, Data ONTAP issues warnings to the console and logs if you have fewer than one hot spare disk that matches the attributes of each disk in your storage system. You can change the threshold value for these warning messages by using the `raid.min_spare_count` option.

To make sure that you always have two hot spares for every disk (a best practice), you can set the `raid.min_spare_count` option to 2.

Setting the `raid.min_spare_count` option to 0 disables low spare warnings. You might want to do this if you do not have enough disks to provide hot spares (for example if your storage system does not support external disk shelves). You can disable the warnings only if the following requirements are met:

- Your system has 16 or fewer disks.
- You have no RAID groups that use RAID4.

  Note: You cannot create aggregates that use RAID4 protection while the `raid.min_spare_count` option is set to 0. If either of these requirements is no longer met after this option has been set to 0, the option is automatically set back to 1.

How Data ONTAP handles a failed disk with a hot spare

Using an available matching hot spare, Data ONTAP can use RAID to reconstruct the missing data from the failed disk onto the hot spare disk with no data service interruption.

If a disk fails and a matching or appropriate spare is available, Data ONTAP performs the following tasks:

- Replaces the failed disk with a hot spare disk.

If RAID-DP is enabled and double-disk failure occurs in the RAID group, Data ONTAP replaces each failed disk with a separate spare disk.

- In the background, reconstructs the missing data onto the hot spare disk or disks.

  Note: During reconstruction, the system is in degraded mode, and file service might slow down.

- Logs the activity in the `/etc/messages` file.
- Sends an AutoSupport message.

Attention: After Data ONTAP is finished reconstructing data, replace the failed disk or disks with new hot spare disks as soon as possible, so that hot spare disks are always available in the storage system.

Note: If the available spare disks are not the correct size, Data ONTAP chooses a disk of the next larger size and restricts its capacity to match the size of the disk it is replacing.

Example: A larger disk is used for reconstructing a failed disk

Suppose you have an aggr, `aggr1`, which contains only 68-GB disks.

```
sys1> aggr status -r aggr1
Aggregate aggr1 (online, raid4) (block checksums)
Plex /aggr1/plex0 (online, normal, active)
RAID group /aggr1/plex0/rg0 (normal)
RAID Disk Device HA SHELF BAY CHAN Pool Type  RPM  Used (MB/blks)  Phys (MB/blks)
--------- ------ -- ----- --- ---- ---- ---- ----- --------------  --------------
parity    0a.19  0a   1   3   FC:A   -  FCAL 10000 68000/139264000 69536/142410400
data      0a.21  0a   1   5   FC:A   -  FCAL 10000 68000/139264000 69536/142410400
```

The only spare available is a 136-GB disk.

```
sys1> aggr status -s
Spare disks
RAID Disk Device HA SHELF BAY CHAN Pool Type  RPM  Used (MB/blks)  Phys (MB/blks)
--------- ------ -- ----- --- ---- ---- ---- ----- --------------  --------------
Spare disks for block or zoned checksum traditional volumes or aggregates
spare     0c.48  0c   3   0   FC:A   -  FCAL 10000 136000/280790184 137104/280790184
```

Disk 0a.21, a 68-GB disk, fails. Disk 0c.48, a 136-GB drive, is the only available spare. Disk 0c.48 is used for reconstruction. Its Used size is restricted to 68 GB, even though its Physical size remains at 136 GB.

```
sys1> aggr status -r aggr1
Aggregate aggr1 (online, raid4, reconstruct) (block checksums)
Plex /aggr1/plex0 (online, normal, active)
RAID group /aggr1/plex0/rg0 (reconstruction 1% completed)

RAID Disk Device HA SHELF BAY CHAN Pool Type  RPM  Used (MB/blks)  Phys (MB/blks)
--------- ------ -- ----- --- ---- ---- ---- ----- --------------  --------------
parity    0a.19  0a   1   3   FC:A   -  FCAL 10000 68000/139264000  69536/142410400
data      0c.48  0c   3   1   FC:A   -  FCAL 10000 68000/139264000 137104/280790184
```

Later, you add a 68-GB disk to the system. You can now replace the 136-GB disk with the new 68-GB disk using the `disk replace` command.

```
sys1> disk replace start 0c.48 0a.22
*** You are about to copy and replace the following file system disk ***
Disk /aggr1/plex0/rg0/0c.48
RAID Disk Device HA SHELF BAY CHAN Pool Type  RPM  Used (MB/blks)  Phys (MB/blks)
--------- ------ -- ----- --- ---- ---- ---- ----- --------------  --------------
```

```
data      0c.48  0c   3   1   FC:A  -   FCAL 10000 68000/139264000 137104/280790184
Really replace disk 0c.48 with 0a.22? y
disk replace: Disk 0c.48 was marked for replacing.

sys1> aggr status -r aggr1
Aggregate aggr1 (online, raid4) (block checksums)
Plex /aggr1/plex0 (online, normal, active)
RAID group /aggr1/plex0/rg0 (normal)

RAID Disk Device HA SHELF BAY CHAN Pool Type  RPM  Used (MB/blks)  Phys (MB/blks)
--------- ------ -- ----- --- ---- ---- ---- ----- --------------  --------------
parity    0a.19  0a   1   3   FC:A   -  FCAL 10000 68000/139264000  69536/142410400
data      0c.48  0c   3   1   FC:A   -  FCAL 10000 68000/139264000 137104/280790184
(replacing, copy in progress)
-> copy   0a.22  0a   1   6   FC:A   -  FCAL 10000 68000/139264000  69536/142410400
(copy 1% completed)
```

Related concepts

How Data ONTAP handles a failed disk that has no available hot spare on page 113

Related tasks

Removing a failed disk on page 76
Adding disks to a storage system on page 73

How Data ONTAP handles a failed disk that has no available hot spare

When a failed disk has no appropriate hot spare available, Data ONTAP puts the affected RAID group into degraded mode indefinitely and the storage system automatically shuts down within a specified time period.

If the maximum number of disks have failed in a RAID group (two for RAID-DP, one for RAID4), the storage system automatically shuts down in the period of time specified by the `raid.timeout` option. The default timeout value is 24 hours.

To ensure that you are aware of the situation, Data ONTAP sends an AutoSupport message whenever a disk fails. In addition, it logs a warning message in the `/etc/message` file once per hour after a disk fails.

Attention: If a disk fails and no hot spare disk is available, contact technical support.

Related concepts

About degraded mode on page 110
How Data ONTAP handles a failed disk with a hot spare on page 111

How RAID-level disk scrubs verify data integrity

RAID-level scrubbing means checking the disk blocks of all disks in use in aggregates (or in a particular aggregate, plex, or RAID group) for media errors and parity consistency. If Data ONTAP

finds media errors or inconsistencies, it uses RAID to reconstruct the data from other disks and rewrites the data.

RAID-level scrubs help improve data availability by uncovering and fixing media and checksum errors while the RAID group is in a normal state (for RAID-DP, RAID-level scrubs can also be performed when the RAID group has a single-disk failure).

RAID-level scrubs can be scheduled or run manually.

Next topics

How you schedule automatic RAID-level scrubs on page 114
How you run a manual RAID-level scrub on page 115

How you schedule automatic RAID-level scrubs

By default, Data ONTAP performs a weekly RAID-level scrub starting on Sunday at 1:00 a.m. for a duration of six hours. You can change the start time and duration of the weekly scrub, add more automatic scrubs, or disable the automatic scrub.

To schedule an automatic RAID-level scrub, you use the `raid.scrub.schedule` option.

To change the duration of automatic RAID-level scrubbing without changing the start time, you use the `raid.scrub.duration` option, specifying the number of minutes you want automatic RAID-level scrubs to run. If you set this option to `-1`, all automatic RAID-level scrubs run to completion.

Note: If you specify a duration using the `raid.scrub.schedule` option, that value overrides the value you specify with this option.

To enable or disable automatic RAID-level scrubbing, you use the `raid.scrub.enable` option.

For more information about these options, see the na_options(1) man page.

Scheduling example

The following command schedules two weekly RAID scrubs. The first scrub is for 240 minutes (four hours) every Tuesday starting at 2 a.m. The second scrub is for eight hours every Saturday starting at 10 p.m.

```
options raid.scrub.schedule 240m@tue@2,8h@sat@22
```

Verification example

The following command displays your current RAID-level automatic scrub schedule. If you are using the default schedule, nothing is displayed.

```
options raid.scrub.schedule
```

Reverting to the default schedule example

The following command reverts your automatic RAID-level scrub schedule to the default (Sunday at 1:00 am, for six hours):

```
options raid.scrub.schedule " "
```

Related tasks

Controlling the performance impact of RAID-level scrubbing on page 120

How you run a manual RAID-level scrub

You can manually run a RAID-level scrub on individual RAID groups, plexes, aggregates, or all aggregates using the `aggr scrub` command. You can also stop, suspend, and resume manual RAID-level scrubs.

If you try to run a RAID-level scrub on a RAID group that is not in a normal state (for example, a group that is reconstructing or degraded), the scrub returns errors and does not check that RAID group. You can run a RAID-level scrub on a RAID-DP group with one failed disk.

Scrubbing all aggregates

The following command starts a RAID-level scrub on all of the aggregates in the storage system:

```
aggr scrub start
```

Scrubbing a particular RAID group

The following command starts a RAID-level scrub on rg0 in plex1 of aggregate aggr2:

```
aggr scrub start aggr2/plex1/rg0
```

Stopping a manual RAID-level scrub

The following command stops a manual RAID-level scrub currently running on plex1 or aggr0:

```
aggr scrub stop aggr0/plex1
```

If you do not specify a name of an aggregate, plex, or RAID group, Data ONTAP stops all manual RAID-level scrubs. After you stop a scrub, it cannot be resumed.

Suspending a manual RAID-level scrub

The following command suspends a manual RAID-level scrub currently running on aggregate aggr3:

```
aggr scrub suspend aggr3
```

You can resume this scrub later by using the `aggr scrub resume` command.

Viewing RAID-level scrub status

The following command displays the status of all currently running RAID-level scrubs, along with the date and time when the last full scrub completed:

```
aggr scrub status -v
```

Customizing the size of your RAID groups

You can customize the size of your RAID groups based on your requirements for data availability, performance, and disk utilization.

About this task

You change the size of RAID groups on a per-aggregate basis, by setting the `raidsize` aggregate option. You cannot change the size of individual RAID groups.

The following list outlines some facts about changing the `raidsize` aggregate option:

- If you increase the `raidsize` option, more disks or array LUNs will be added to the most recently created RAID group until it reaches the new size.
- All other existing RAID groups in that aggregate remain the same size, unless you explicitly add disks to them.
- You cannot decrease the size of already created RAID groups.
- The new size applies to all subsequently created RAID groups in that aggregate.

Step

1. Enter the following command:

   `aggr options *aggr_name* raidsize *size*`

   Example
   The following command changes the raidsize setting of the aggregate aggr3 to 16 disks or array LUNs:

   `aggr options aggr3 raidsize 16`

Related concepts

How Data ONTAP uses RAID to protect your data and data availability on page 101
Considerations for sizing RAID groups for disks on page 107
Considerations for Data ONTAP RAID groups for array LUNs on page 108
How Data ONTAP RAID groups work on page 106

Related tasks

Increasing the size of an aggregate on page 140

Related references

Storage limits on page 345

Controlling the impact of RAID operations on system performance

You can reduce the impact of RAID operations on system performance by decreasing the speed of RAID operations.

About this task

You can control the speed of the following RAID operations with RAID options:

- RAID data reconstruction
- Disk scrubbing
- Plex resynchronization
- Synchronous mirror verification

The speed that you select for each of these operations might affect the overall performance of the storage system. However, if the operation is already running at the maximum speed possible and it is fully utilizing one of the three system resources (the CPU, disks, or the disk-to-controller connection bandwidth), changing the speed of the operation has no effect on the performance of the operation or the storage system.

If the operation is not yet running, you can set a speed that minimally slows storage system network operations or a speed that severely slows storage system network operations. For each operation, use the following guidelines:

- If you want to reduce the performance impact on client access to the storage system, change the specific RAID option from medium to low. Doing so also causes the operation to slow down.
- If you want to speed up the operation, change the RAID option from medium to high. Doing so might decrease the performance of the storage system in response to client access.

Next topics

Controlling the performance impact of RAID data reconstruction on page 120
Controlling the performance impact of RAID-level scrubbing on page 120
Controlling the performance impact of plex resynchronization on page 121
Controlling the performance impact of mirror verification on page 122

Controlling the performance impact of RAID data reconstruction

Because RAID data reconstruction consumes CPU resources, increasing the speed of data reconstruction sometimes slows storage system network and disk operations. You can control the speed of data reconstruction with the `raid.reconstruc.perf_impact` option.

About this task

When RAID data reconstruction and plex resynchronization are running at the same time, Data ONTAP limits the combined resource utilization to the greater impact set by either operation. For example, if `raid.resync.perf_impact` is set to `medium` and `raid.reconstruct.perf_impact` is set to `low`, the resource utilization of both operations has a medium impact.

Step

1. Enter the following command:

   `options raid.reconstruct.perf_impact` ***`impact`***

   `impact` can be `high`, `medium`, or `low`.

   `high` means that the storage system uses most of the system resources—CPU time, disks, and disk-to-controller bandwidth—available for RAID data reconstruction; this setting can heavily affect storage system performance. However, reconstruction finishes faster, reducing the time that the storage system is running in degraded mode.

   `low` means that the storage system uses very little of the system resources; this setting lightly affects storage system performance. However, reconstruction takes more time to complete, increasing the time that the storage system is running in degraded mode.

   The default speed is `medium`.

   Note: The setting for this option also controls the speed of Rapid RAID recovery.

Controlling the performance impact of RAID-level scrubbing

When Data ONTAP performs a RAID-level scrub, it checks the disk blocks of all disks on the storage system for media errors and parity consistency. You can control the impact this operation has on system performance with the `raid.verify.perf_impact` option.

About this task

When RAID-level scrubbing and mirror verification are running at the same time, Data ONTAP limits the combined resource utilization to the greater impact set by either operation. For example, if

`raid.verify.perf_impact` is set to `medium` and `raid.scrub.perf_impact` is set to `low`, the resource utilization by both operations has a medium impact.

Note: If there are times during the day where the load on your storage system is decreased, you can also limit the performance impact of the automatic RAID-level scrub by changing the start time or duration of the automatic scrub.

Step

1. Enter the following command:

   `options raid.scrub.perf_impact` ***`impact`***

   `impact` can be `high`, `medium`, or `low`.

   `high` means that the storage system uses most of the system resources—CPU time, disks, and disk-to-controller bandwidth—available for scrubbing; this setting can heavily affect storage system performance, but the scrub will complete in less time.

   `low` means that the storage system uses very little of the system resources; this setting lightly affects storage system performance, and the scrub will take longer to complete.

   The default value for *`impact`* is `low`.

Related concepts

How you schedule automatic RAID-level scrubs on page 114
How Data ONTAP uses RAID to protect your data and data availability on page 101

Controlling the performance impact of plex resynchronization

You can control the performance impact of plex resynchronization by using the `raid.reconstruct.perf_impact` option.

About this task

Plex resynchronization is a process that ensures two plexes of a mirrored aggregate have exactly the same data. When plexes are unsynchronized, one plex contains data that is more up to date than that of the other plex. Plex resynchronization updates the out-of-date plex so that both plexes are identical.

Data ONTAP resynchronizes the two plexes of a mirrored aggregate if one of the following situations occurs:

- One of the plexes was taken offline and then brought online later.
- You add a plex to an unmirrored aggregate.

When plex resynchronization and RAID data reconstruction are running at the same time, Data ONTAP limits the combined resource utilization to the greatest impact set by either operation. For

example, if `raid.resync.perf_impact` is set to `medium` and `raid.reconstruct.perf_impact` is set to `low`, the resource utilization by both operations has a medium impact.

Step

1. Enter the following command:

   `options raid.resync.perf_impact` ***`impact`***

   `impact` can be `high`, `medium`, or `low`.

   `high` means that the storage system uses most of the system resources available for plex resynchronization; this setting can heavily affect storage system performance, but the resynchronization finishes sooner.

   `low` means that the storage system uses very little of the system resources; this setting lightly affects storage system performance, but the resynchronization will take longer to finish.

   The default impact is `medium`.

Controlling the performance impact of mirror verification

You use mirror verification to ensure that the two plexes of a synchronous mirrored aggregate are identical. You can control the speed of mirror verification, and its effect on system resources, by using the `raid.verify.perf_impact` option.

About this task

When mirror verification and RAID-level scrubbing are running at the same time, Data ONTAP limits the combined resource utilization to the greatest impact set by either operation. For example, if `raid.verify.perf_impact` is set to `medium` and `raid.scrub.perf_impact` is set to `low`, the resource utilization of both operations has a medium impact.

For more information about synchronous mirroring, see the *Data Protection Online Backup and Recovery Guide*.

Step

1. Enter the following command:

   `options raid.verify.perf_impact` ***`impact`***

   `impact` can be `high`, `medium`, or `low`.

   `high` means that the storage system uses most of the system resources available for mirror verification; this setting can heavily affect storage system performance, but the mirror verification finishes faster.

`low` means that the storage system uses very little of the system resources; this setting lightly affects storage system performance, but the mirror verification finishes more slowly.

The default speed is `low`.

How you use aggregates to provide storage to your volumes

To support the differing security, backup, performance, and data sharing needs of your users, you group the physical data storage resources on your storage system into one or more aggregates. These aggregates provide storage to the volume or volumes that they contain.

Each aggregate has its own RAID configuration, plex structure, and set of assigned disks or array LUNs. When you create an aggregate without an associated traditional volume, you can use it to hold one or more FlexVol volumes—the logical file systems that share the physical storage resources, RAID configuration, and plex structure of that common containing aggregate. When you create an aggregate with its tightly-bound traditional volume, then it can contain only that volume.

Aggregates can be mirrored or unmirrored. An unmirrored aggregate has only one plex; a mirrored aggregates have two plexes.

For information about best practices for working with aggregates, see *Technical Report 3437: Storage Best Practices and Resiliency Guide*.

Note: This technical report contains information about NetApp products that IBM licenses and in some cases customizes. Technical reports might contain information about models and features that are not supported by IBM.

Next topics

How unmirrored aggregates work on page 126
How mirrored aggregates work on page 127
Aggregate states and status on page 128
How you can use disks with mixed speeds in the same aggregate on page 130
How to control disk selection from heterogeneous storage on page 131
Rules for mixing disk types in aggregates on page 132
Rules for mixing array LUNs in an aggregate on page 133
Checksum rules for adding storage to an aggregate on page 134
What happens when you add larger disks to an aggregate on page 134

Related concepts

Disk speeds supported by Data ONTAP on page 34

Related references

Storage limits on page 345

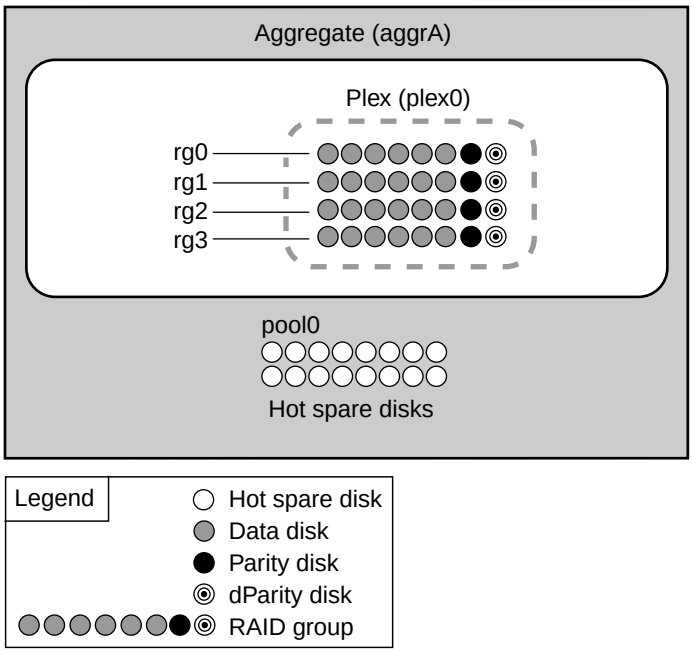
Related information

TR 3437: Storage Best Practices and Resiliency Guide

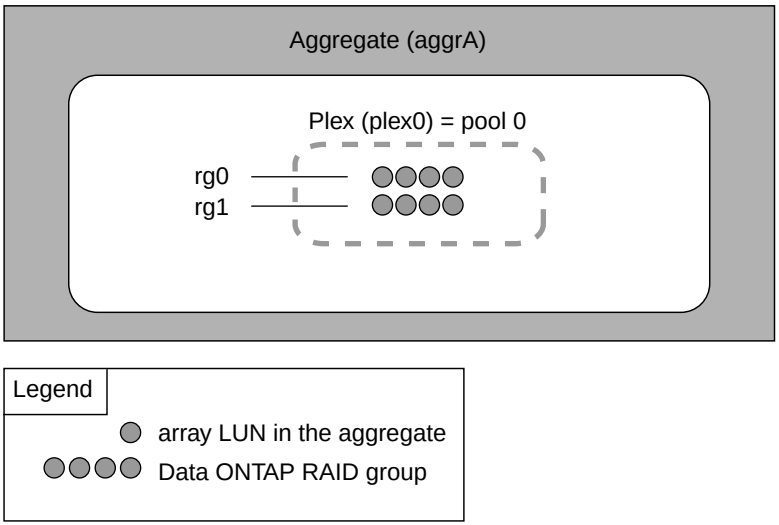
How unmirrored aggregates work

Unless you are using SyncMirror, all of your aggregates are unmirrored. Unmirrored aggregates have only one *plex* (copy of their data), which contains all of the RAID groups belonging to that aggregate.

The following diagram shows an unmirrored aggregate with disks, with its one plex.



The following diagram shows an unmirrored aggregate with array LUNs, with its one plex.



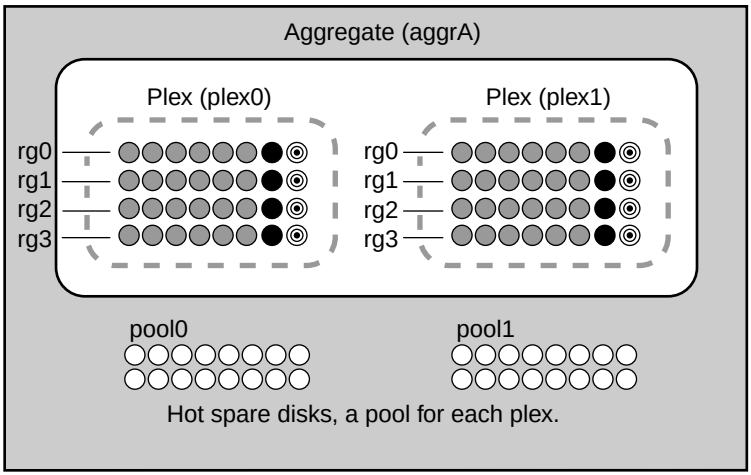
How mirrored aggregates work

Mirrored aggregates have two *plexes* (copies of their data), which use the SyncMirror functionality to duplicate the data to provide redundancy.

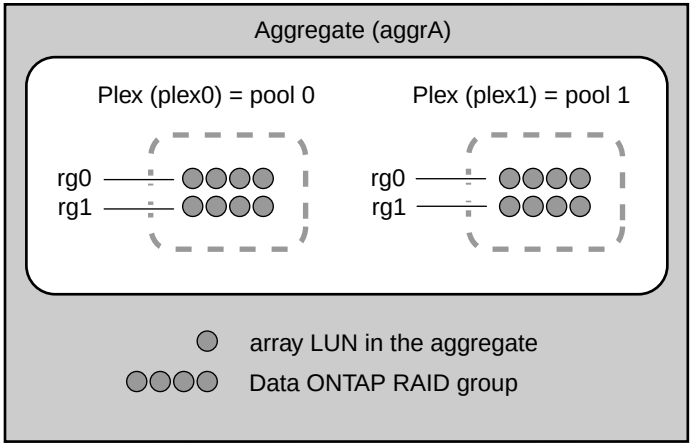
When SyncMirror is enabled, all the disks or array LUNs are divided into two pools, and a copy of the plex is created. The plexes are physically separated (each plex has its own RAID groups and its own pool), and the plexes are updated simultaneously. This provides added protection against data loss if more disks fail than the RAID level of the aggregate protects against or there is a loss of connectivity, because the unaffected plex continues to serve data while you fix the cause of the failure. After the plex that had a problem is fixed, you can resynchronize the two plexes and reestablish the mirror relationship.

Note: Before an aggregate can be enabled for mirroring, the storage system must have the syncmirror_local license installed and enabled, and the storage configuration must support RAID-level mirroring.

In the following diagram of a storage system using disks, SyncMirror is enabled and implemented, so Data ONTAP copies plex0 and automatically names the copy plex1. Plex0 and plex1 contain copies of one or more file systems. In this diagram, 32 disks were available prior to the SyncMirror relationship being initiated. After initiating SyncMirror, the spare disks are allocated to pool0 or pool1.



The following diagram shows a storage system using array LUNs with SyncMirror enabled and implemented.



Aggregate states and status

Aggregates can be in one of three states—online, offline, or restricted. In addition, they can show one or more status values, depending on how they are configured and the health of their disks. You can determine an aggregate's current state and status by using the `aggr status` command.

The following table displays the possible states for aggregates.

State	Description
Online	Read and write access to volumes hosted on this aggregate is allowed.
Restricted	Some operations, such as parity reconstruction, are allowed, but data access is not allowed.
Offline	No access to the aggregate is allowed.

The following table displays the possible status values for aggregates.

Status	Description
`aggr`	This aggregate is capable of containing FlexVol volumes.
`copying`	The aggregate is currently the target aggregate of an active copy operation.
`degraded`	The aggregate contains at least one RAID group with single disk failure that is not being reconstructed..
`double degraded`	The aggregate contains at least one RAID group with double disk failure that is not being reconstructed (RAID-DP aggregates only).
`foreign`	Disks that the aggregate contains were moved to the current storage system from another storage system.
`growing`	Disks are in the process of being added to the aggregate.
`initializing`	The aggregate is in the process of being initialized.
`invalid`	The aggregate contains no volumes and none can be added. Typically this happens only after an aborted `aggr copy` operation.
`ironing`	A WAFL consistency check is being performed on the aggregate.
`mirror degraded`	The aggregate is mirrored and one of its plexes is offline or resynchronizing.
`mirrored`	The aggregate is mirrored.
`needs check`	WAFL consistency check needs to be performed on the aggregate.
`normal`	The aggregate is unmirrored and all of its RAID groups are functional.
`out-of-date`	The aggregate is mirrored and needs to be resynchronized.

Status	Description
`partial`	At least one disk was found for the aggregate, but two or more disks are missing.
`raid0`	The aggregate consists of RAID0 (no parity) RAID groups (gateways only).
`raid4`	The aggregate consists of RAID4 RAID groups.
`raid_dp`	The aggregate consists of RAID-DP RAID groups.
`reconstruct`	At least one RAID group in the aggregate is being reconstructed.
`redirect`	Aggregate reallocation or file reallocation with the `-p` option has been started on the aggregate. Read performance on volumes in the aggregate might be degraded.
`resyncing`	One of the mirrored aggregate's plexes is being resynchronized.
`snapmirrored`	The aggregate is a SnapMirror replica of another aggregate (traditional volumes only).
`trad`	The aggregate is a traditional volume and cannot contain FlexVol volumes.
`verifying`	A mirror verification operation is currently running on the aggregate.
`wafl inconsistent`	The aggregate has been marked corrupted; contact technical support.

How you can use disks with mixed speeds in the same aggregate

Whenever possible, you should use disks of the same speed in an aggregate. However, if needed, you can configure Data ONTAP to allow mixed speed aggregates based on the disk type.

To configure Data ONTAP to allow mixed speed aggregates, you use the following options:

- `raid.rpm.fcal.enable`
- `raid.rpm.ata.enable`

When these options are set to `off`, Data ONTAP allows mixing speeds for the designated disk type.

By default, `raid.rpm.fcal.enable` is set to `on`, and `raid.rpm.ata.enable` is set to `off`.

Note: Even if Data ONTAP is not configured to allow mixing speeds, you can still add disks with different speeds to an aggregate using the `-f` option of the `aggr create` or `aggr add` commands.

Related concepts

Disk speeds supported by Data ONTAP on page 34

How to control disk selection from heterogeneous storage

When disks with different characteristics coexist on the same storage system, the system is said to have heterogeneous storage. When you create an aggregate from heterogeneous storage, you can explicitly select disks with the correct characteristics to ensure that Data ONTAP uses the disks you expect.

When you create a new aggregate using heterogeneous storage, you should use one of the following methods to ensure that the correct disks or disk types are selected:

- Specify the disk attributes you want to use:
  - You can specify disk size by using the `@size` option to the number of disks. For example, **`6@300G`** tells Data ONTAP to use six 300-GB disks.
  - You can specify disk speed by using the `-R` option.
  - You can specify disk type by using the `-T` option.

    Note: The `-R` and `-T` options are not available when you are adding disks to an existing aggregate; they are available only for creating a new aggregate.
- Use an explicit disk list.

  You can list the names of specific disks you want to use.
- Use disk selection preview.

  You can use the `-n` option to identify which disks Data ONTAP will select automatically. If you are happy with the disks selected, you can proceed with automatic disk selection. Otherwise, you can use one of the previous methods to ensure that the correct disks are selected.

Note: For unplanned events such as disk failures, which cause Data ONTAP to add another disk to a RAID group automatically, the best way to ensure that Data ONTAP will choose the best disk for any RAID group on your system is to always have at least one spare (and preferably two) available to match all disk types and sizes in use in your system.

Rules for mixing disk types in aggregates

You can mix disks from different loops or stacks within the same aggregate. Depending on the value of the `raid.disktype.enable` option, you might be able to mix certain types of disks within the same aggregate.

The following table shows what types of disks can be mixed within an aggregate when the `raid.disktype.enable` option is set to `off`:

	SAS disks in EXN3000/3500 disk shelves	SATA disks in EXN3000 disk shelves	FC disks in EXN2000 or EXN4000 disk shelves	ATA disks in EXN2000 or EXN4000 disk shelves
Internal SAS disks	Y	N	Y	N
Internal SATA disks	N	Y	N	Y
SAS disks in EXN3000/3500 disk shelves	Y	N	Y*	N
SATA disks in EXN3000 disk shelves	N	Y	N	Y*
FC disks in EXN2000 or EXN4000 disk shelves	Y*	N	Y	N
ATA disks in EXN2000 or EXN4000 disk shelves	N	Y*	N	Y

*Data ONTAP does not prevent these combinations. However, due to the large difference in performance between the two disk types, you should avoid these combinations.

BSAS disks are considered to be the same as SATA disks in this table.

SAS and SATA disks are *not* allowed in the same aggregate.

If the `raid.disktype.enable` option is set to `on`, all aggregates must contain disks of a single type.

Note: If you set the `raid.disktype.enable` option to `on` for a system that already contains aggregates with disks of mixed type, those mixed aggregates continue to function normally and

accept both types of disks. However, no other aggregates will accept mixed disk types as long as the `raid.disktype.enable` option is set to `on`.

For information about best practices for working with different types of disks, see *Technical Report 3437: Storage Best Practices and Resiliency Guide*.

Note: This technical report contains information about NetApp products that IBM licenses and in some cases customizes. Technical reports might contain information about models and features that are not supported by IBM.

Related concepts

What disk types Data ONTAP supports on page 31

Related information

TR 3437: Storage Best Practices and Resiliency Guide

Rules for mixing array LUNs in an aggregate

Data ONTAP does not support mixing different types of storage in the same aggregate because it causes performance degradation.

There are restrictions on the types of array LUNs that you can mix in the same aggregate, which you must observe when you add array LUNs to an aggregate. Data ONTAP does not *prevent* you from mixing different types of array LUNs.

Note: Data ONTAP prevents you from mixing native disks and array LUNs in the same aggregate.

For aggregates for third-party storage, you cannot mix the following storage types in the same aggregate:

- Array LUNs from storage arrays from different vendors
- Array LUNs from storage arrays from the same vendor but from different storage array families

  Note: Storage arrays in the same family share the same characteristics---for example, the same performance characteristics. See the gateway implementation guide for your vendor for information about how Data ONTAP defines family members for the vendor.
- Array LUNs from storage arrays with 4-Gb HBAs and array LUNs from storage arrays with 2-Gb HBAs
- Array LUNs from Fibre Channel and SATA drives

  You can deploy Fibre Channel and SATA drives behind the same gateway. However, you cannot mix array LUNs from SATA disks and Fibre Channel disks in the same aggregate, even if they are from the same series and the same vendor. Before setting up this type of configuration, consult your authorized reseller to plan the best implementation for your environment.

Checksum rules for adding storage to an aggregate

If you have disks or array LUNs of both checksum types (blocks and zoned) in your storage system, you must follow the checksum type rules when you add storage to an aggregate.

Data ONTAP enforces the following rules when creating aggregates or adding storage to existing aggregates:

- An aggregate can have only one checksum type, and it applies to the entire aggregate.
- To use block checksum storage when you create a new aggregate, you must have at least the number of block checksum spare disks or array LUNs available that you specified in the `aggr create` command.
- When you add storage to an existing aggregate, the following rules apply:
  - You can add block checksum storage to either a block checksum aggregate or a zoned checksum aggregate.
  - You cannot add zoned checksum storage to a block checksum aggregate.

The following table shows the types of storage that you can add to an existing aggregate of each type.

Storage checksum type	Block checksum aggregate	Zoned checksum aggregate
Block checksum	OK	OK
Zoned checksum	Not allowed	OK

What happens when you add larger disks to an aggregate

What Data ONTAP does when you add disks to an aggregate that are larger than the existing disks depends on the RAID level (RAID4 or RAID-DP) of the aggregate.

- When an aggregate configured for RAID4 protection is created, Data ONTAP assigns the role of parity disk to the largest disk in each RAID group.

  When an existing RAID4 group is assigned an additional disk that is larger than the group's existing parity disk, then Data ONTAP reassigns the new disk as parity disk for that RAID group.
- When an aggregate configured for RAID-DP protection is created, Data ONTAP assigns the role of dParity disk and regular parity disk to the largest and second largest disk in the RAID group.

  When an existing RAID-DP group is assigned an additional disk that is larger than the group's existing dParity disk, then Data ONTAP reassigns the new disk as the regular parity disk for that RAID group and restricts its capacity to be the same size as the existing dParity disk. Note that Data ONTAP does *not* replace the existing dParity disk, even if the new disk is larger than the dParity disk.

Note: Because the smallest parity disk limits the effective size of disks added to a RAID-DP group, you can maximize available disk space by ensuring that the regular parity disk is as large as the dParity disk.

Note: If needed, you can replace a capacity-restricted disk with a more suitable (smaller) disk later, to avoid wasting disk space. However, replacing a disk already in use in an aggregate with a larger disk does not result in any additional usable disk space; the new disk is capacity-restricted to be the same size as the smaller disk it replaced.

Example: adding a larger disk to a mixed-size RAID-DP aggregate

In this example, aggr2 has two 136-GB disks and one 68-GB disk. The 136-GB disks were used as the parity disks.

```
sys1> aggr status -r aggr2
Aggregate aggr2 (online, raid_dp) (block checksums)
  Plex /aggr2/plex0 (online, normal, active)
    RAID group /aggr2/plex0/rg0 (normal)

    RAID Disk Device HA SHELF BAY CHAN  Type  RPM   Used (MB/blks)      Phys (MB/blks)
    --------- ------ -- ----- --- ----- ----- ----- ------------------ ----------------
    dparity   0c.48  0c 3     0   FC:A  FCAL  10000 136000/278528000   137104/280790184
    parity    0c.50  0c 3     2   FC:A  FCAL  10000 136000/278528000   137104/280790184
    data      0a.28  0a 1     12  FC:A  FCAL  10000  68000/139264000    69536/142410400
```

When another 136-GB disk is added to the aggregate, the disk is added as a data disk and is not restricted in size.

```
sys1> aggr add aggr2 -d 0c.49

sys1> aggr status -r aggr2
Aggregate aggr2 (online, raid_dp) (block checksums)
  Plex /aggr2/plex0 (online, normal, active)
    RAID group /aggr2/plex0/rg0 (normal)

    RAID Disk Device HA SHELF BAY CHAN Type RPM   Used (MB/blks)     Phys (MB/blks)
    --------- ------ -- ----- --- ---- ---- ----- ------------------ ----------------
    dparity   0c.48  0c 3     0   FC:A FCAL 10000 136000/278528000   137104/280790184
    parity    0c.50  0c 3     2   FC:A FCAL 10000 136000/278528000   137104/280790184
    data      0a.28  0a 1     12  FC:A FCAL 10000  68000/139264000    69536/142410400
    data      0c.49  0c 3     1   FC:A FCAL 10000 136000/278528000   137104/280790184
```

Note: If the parity disk had been a 68-GB disk, then the newly added disk would have been restricted to 68 GB.

Related tasks

Replacing disks that are currently being used in an aggregate on page 74

Managing aggregates

You manage aggregates by creating them, increasing their size, setting their RAID level, and managing their state. In addition, you can destroy, undestroy and move aggregates.

About this task

Note: You cannot reduce the number of disks in an aggregate by removing data disks. The only way to reduce the number of data disks in an aggregate is to copy the data and transfer it to a new aggregate that has fewer data disks.

Next topics

Creating an aggregate

You create an aggregate to provide storage to one or more FlexVol volumes (or one traditional volume).

Before you begin

Determine the name of the aggregate. Aggregate names must conform to the following requirements:

- Begin with either a letter or an underscore (_)
- Contain only letters, digits, and underscores
- Contain no more than 250 characters

Note: You can change the name of an aggregate later by using the `aggr rename` command.

Determine what disks or array LUNs will be used in the new aggregate. You can specify disks by listing their IDs, or by specifying a disk characteristic such as speed or type. You can display a list of the available spares on your storage system by using the `aggr status -s` command.

Note: If your storage system is attached to more than one type of disk, or to both disks and array LUNs, and you do not use the `-T` option, Data ONTAP creates the aggregate using the disk type (including array LUNs) with the highest number of available disks. To ensure that Data ONTAP uses the disk type that you expect, always use the `-T` option when creating aggregates from heterogeneous storage.

Step

1. Enter the following command:

   `aggr create` *`aggr_name`* `[-f] [-m] [-n] [-t {raid0 | raid4 | raid_dp}] [-r` *`raidsize`*`] [-T` *`disk-type`*`] -R` *`rpm`*`] [-L]` *`disk-list`*

   `aggr_name` is the name for the new aggregate.

   `-f` overrides the default behavior that does not permit disks in a plex to belong to different disk pools. This option also allows you to mix disks with different RPM speeds even if the appropriate `raid.rpm` option is not off.

   `-m` specifies the optional creation of a SyncMirror-replicated volume if you want to supplement RAID protection with SyncMirror protection. A SyncMirror license is required for this feature.

   `-n` displays the results of the command but does not execute it. This is useful for displaying the disks that would be automatically selected prior to executing the command.

   `-t {raid0 | raid4 | raid_dp}` specifies the level of RAID protection you want to provide for this aggregate. If no RAID level is specified for an aggregate composed of disks, the default value (`raid_dp`) is applied. `raid0` is used only for array LUNs.

   `-r` *`raidsize`* is the maximum size of the RAID groups for this aggregate. If no size is specified, the default is used.

   `-T` *`disk-type`* specifies one of the following types of disk to be used: `ATA`, `SATA`, `SAS`, `BSAS`, `FCAL`, or `LUN`. This option is only needed when creating aggregates on systems that have mixed disk types or both disks and array LUNs. Use `SATA` for SAS-attached ATA disks, `SAS` for SAS-attached SAS disks, `FCAL` for FC disks, `ATA` for ATA disks connected through FC-AL, and `LUN` for array LUNs.

   Note: If the `raid.disktype.enable` option is set to `off` (its default value), FCAL and SAS disks are considered to be the same type for the purposes of creating an aggregate and may be combined even if the `-T` option is used. Similarly, ATA, BSAS, and SATA disks are considered to be the same type and may be combined, even when the `-T` option is used.

   `-R` *`rpm`* specifies the type of disk to use based on its speed. Valid values for *`rpm`* include `5400`, `7200`, `10000`, and `15000`.

`-L` creates a SnapLock aggregate. For more information, see the na_aggr(1) man page or the *Data ONTAP Archive and Compliance Management Guide*.

`disk-list` is one of the following values:

- *`ndisks`*[`@`*`disk-size`*]
  `ndisks` is the number of disks to use. It must be at least 3 for RAID-DP aggregates, 2 for RAID-4 aggregates, or 1 for RAID0 aggregates.
  `disk-size` is the disk size to use, in gigabytes.
- `-d` *`disk_name1 disk_name2... disk_nameN`*
  `disk_name1`, *`disk_name2`*, and *`disk_nameN`* are disk IDs of available disks; use a space to separate disk IDs.

Examples

The following command creates an aggregate called newaggr, with a RAID group size of 8, consisting of the disks with disk IDs 8a.16, 8a.17, 8a.18, and 8a.19:

```
aggr create newaggr -r 8 -d 8a.16 8a.17 8a.18 8a.19
```

The following command creates an aggregate called newfastaggr, with 20 disks, the default RAID group size, and all disks with 15K RPM:

```
aggr create newfastaggr -R 15000 20
```

The following command creates an aggregate called newFCALaggr. Note that if SAS disks are present, they might be used, because FC and SAS disks are considered to be the same type.

```
aggr create newFCALaggr -T FCAL 15
```

After you finish

You can use the `aggr status -r` command to verify the RAID groups and disks used in the aggregate you just created.

Related concepts

Considerations for sizing RAID groups for disks on page 107
Considerations for Data ONTAP RAID groups for array LUNs on page 108
Protection provided by RAID and SyncMirror on page 103
How you use aggregates to provide storage to your volumes on page 125
How Data ONTAP uses RAID to protect your data and data availability on page 101

Related references

Storage limits on page 345

Increasing the size of an aggregate

You can add disks or array LUNs to an aggregate to increase its size, so it can provide more storage space to its contained volumes. You might also want to increase the size of a specific RAID group.

Before you begin

Make sure you understand the following concepts:

- The requirement to add disks or array LUNs owned by the same system and pool
- How the type of your aggregate affects its maximum size.
  You cannot change the type of an aggregate by adding storage to it.
- The maximum size of your aggregate
  You cannot add storage to an aggregate that would cause the aggregate to exceed its maximum size.
- For aggregates composed of disks:
  - Benefits of keeping your RAID groups homogenous for disk size and speed
  - What types of disks can be used together
  - How to ensure that the correct disks are added to the aggregate (the `aggr add` command cannot be undone)
  - How to add disks to aggregates from heterogenous storage
  - The minimum number of disks to add for best performance
    For best performance, you should add a complete RAID group to prevent the new disks from becoming a performance bottleneck.
  - How many hot spares you need to provide for maximum protection against disk failures

About this task

You can specify a RAID group to add disks or array LUNs to. If you do not specify a RAID group, the disks or array LUNs are added to the most recently created RAID group if there is room in that RAID group. Otherwise, a new RAID group is created.

To see the number and types of disks or array LUNs in each RAID group, you can use the `aggr status -r` command.

Steps

1. Verify that appropriate spare disks or array LUNs are available for you to add by entering the following command:

   `aggr status -s`

   For disks, make sure that enough of the spares listed are of the correct type, size, speed, and checksum type for the target RAID group in the aggregate to which you are adding the disks.

2. Add the disks or array LUNs by entering the following command:

   aggr add *aggr_name* [-f] [-n] [-g *raid_group_name*] *disk_list*

   `-f` enables you to add disks or array LUNs from a different pool or, for disks, of a different speed.

   `-n` displays the results of the command but does not execute it. This is useful for displaying the disks or array LUNs that Data ONTAP would automatically select. You can then decide whether to accept the selection provided by Data ONTAP or to add different disks or array LUNs.

   If you specify the `-g` option, the storage is added to the RAID group you specify. *raid_group_name* is the name that Data ONTAP gave to the group—for example, rg0.

   To add the storage to a new RAID group, use the `new` keyword instead of the group name.

   disk_list is one of the following parameters:

   - *ndisks*[*disk_size*]
   - `-d` *disk1* [*disk2*...]

   The *disk_size* parameter is the approximate size of the disk in GBs. Disks that are within approximately 20 percent of the specified size are selected.

Examples

The following command adds four 300-GB disks to the aggr1 aggregate:

```
aggr add aggr1 4@300
```

The following command adds the disks 5a.17, 5a.19, 5a.20, and 5a.26 to the rg1 RAID group of the aggr2 aggregate:

```
aggr add aggr2 -g rg1 -d 5a.17 5a.19 5a.20 5a.26
```

The following command adds four disks to each plex of a mirrored aggregate aggr_mir:

```
aggr add aggr_mir -d 5a.18 5a.19 5a.20 5a.21 -d 8b.14 8b.15 8b.16 8b.
17
```

After you finish

After you add storage to an aggregate, run a full reallocation job on each FlexVol volume contained in that aggregate. For information about reallocation, see the *System Administration Guide*.

Next topics

What happens when you add storage to an aggregate on page 142
Forcibly adding disks to aggregates on page 142

Related concepts

How to control disk selection from heterogeneous storage on page 131
How many hot spares you should have on page 109
How you use aggregates to provide storage to your volumes on page 125

Related references

Storage limits on page 345

What happens when you add storage to an aggregate

By default, Data ONTAP adds new disks or array LUNs to the most recently created RAID group until it reaches its maximum size. Then Data ONTAP creates a new RAID group. Alternatively, you can specify a RAID group you want to add storage to.

When you create an aggregate or add storage to an aggregate, Data ONTAP creates new RAID groups as each RAID group is filled with its maximum number of disks or array LUNs. The last RAID group formed might contain fewer disks or array LUNs than the maximum RAID group size for the aggregate. In that case, any storage added to the aggregate is added to the last RAID group until the specified RAID group size is reached.

If you increase the RAID group size for an aggregate, new disks or array LUNs are added only to the most recently created RAID group; the previously created RAID groups remain at their current size unless you explicitly add storage to them using the `-g` option of the `aggr add` command.

Note: You are advised to keep your RAID groups homogeneous when possible. If needed, you can replace a mismatched disk with a more suitable disk later.

Related tasks

Replacing disks that are currently being used in an aggregate on page 74

Forcibly adding disks to aggregates

You might want to override some of the restrictions on what disks can be added to an aggregate if you do not have disks of the right speed or enough disks in the correct pool. You can do so by using the `aggr add -f` command.

About this task

Forcibly adding disks can be useful in the following situations:

- You need to add disks from two different spare disk pools to a mirrored aggregate.

  Note: Using disks from the wrong pool in a mirrored aggregate removes an important fault isolation property of the SyncMirror functionality. You should do so only when absolutely necessary, and you should return to a supported configuration as soon as possible.

- You need to add disks of a different speed than that of existing disks in the aggregate.

Step

1. Add the disks by entering the following command:

   aggr add *aggr_name* -f [-n] [-g *raid_group_name*] *disk_list*

Related concepts

How to control disk selection from heterogeneous storage on page 131
How you can use disks with mixed speeds in the same aggregate on page 130
How mirrored aggregates work on page 127

Taking an aggregate offline

You use the `aggr offline` command to take an aggregate offline to perform maintenance on the aggregate, move it, or destroy it.

Steps

1. If the aggregate you want to take offline contains FlexVol volumes, boot into maintenance mode.

   Note: This step is not necessary for traditional volumes.

2. Enter the following command:

   aggr offline *aggr_name*

3. If you previously booted into maintenance mode, return to normal mode.

Result

The aggregate is now offline. You cannot access any data in the aggregate's volumes.

Related tasks

Taking a volume offline on page 174

Bringing an aggregate online

After you restrict an aggregate or take it offline, you can use the `aggr online` command to make it available to the storage system again by bringing it back online.

Step

1. Enter the following command:

   aggr online *aggr_name*

If the aggregate is inconsistent, the command prompts you for confirmation.

Attention: If you bring an inconsistent aggregate online, it might suffer further file system corruption. If you have an inconsistent aggregate, contact technical support.

Result

The aggregate is online and available for use.

Related tasks

Bringing a volume online on page 175

Putting an aggregate into restricted state

You use the `aggr restrict` command to put the aggregate into a restricted state if you want the aggregate to be the target of an aggregate copy or SnapMirror replication operation.

About this task

For information about aggregate copy and SnapMirror replication, see the *Data Protection Online Backup and Recovery Guide*.

Steps

1. If the aggregate you want to restrict contains FlexVol volumes, boot into maintenance mode.

   Note: This step is not necessary for traditional volumes.

2. Enter the following command:

   `aggr restrict` *`aggr_name`*

3. If you previously booted into maintenance mode, return to normal mode.

Result

The aggregate is now restricted. Data in the aggregate's volumes is unavailable to clients.

Related tasks

Putting a volume into restricted state on page 174

Changing the RAID level of an aggregate

When you change an aggregate's RAID level (from RAID4 to RAID-DP, for example), Data ONTAP reconfigures existing RAID groups to the new level and applies the new level to subsequently created RAID groups.

About this task

Note: You cannot change the Data ONTAP RAID level of aggregates containing array LUNs. Aggregates that contain array LUNs must have a Data ONTAP RAID level of RAID0. RAID protection for aggregates that contain array LUNs is provided by the storage array.

Next topics

Changing an aggregate's RAID level from RAID4 to RAID-DP on page 145
Changing an aggregate's RAID level from RAID-DP to RAID4 on page 146

Changing an aggregate's RAID level from RAID4 to RAID-DP

You can change an existing aggregate's RAID level from RAID4 to RAID-DP if you want the increased protection that RAID-DP provides.

Steps

1. Determine the number of RAID groups and the size of their parity disks in the aggregate in question by entering the following command:

   ```
   aggr status aggr_name -r
   ```

2. List the available hot spares on your system by entering the following command:

   ```
   aggr status -s
   ```

3. Make sure that at least one, and preferably two hot spare disks exist for each RAID group listed. If necessary, add additional hot spare disks.

4. Enter the following command:

   ```
   aggr options aggr_name raidtype raid_dp
   ```

Result

When you change the RAID level of an aggregate from RAID4 to RAID-DP, Data ONTAP makes the following changes:

- Adds an additional disk to each existing RAID group from the storage system's hot spare disks; assigns the new disk the dParity disk function for the RAID-DP group. A reconstruction begins for each RAID group to populate the dParity disk.

- Changes the `raidsize` option for the aggregate to the appropriate RAID-DP default value.

  Note: You can change the `raidsize` option after the RAID level change is complete.

After you finish

You can verify the new RAID level by using the `aggr options` command.

Related concepts

How you use aggregates to provide storage to your volumes on page 125
How Data ONTAP works with hot spare disks on page 109

Related tasks

Customizing the size of your RAID groups on page 117

Related references

Storage limits on page 345

Changing an aggregate's RAID level from RAID-DP to RAID4

When you change an aggregate's RAID level from RAID-DP to RAID4, the extra parity disks are converted to spares. In addition, the `raidsize` option is changed.

Step

1. Enter the following command:

   `aggr options` *`aggr_name`* `raidtype raid4`

Result

When you change the RAID level of an aggregate from RAID4 to RAID-DP, Data ONTAP makes the following changes:

- In each of the aggregate's existing RAID groups, the RAID-DP second parity disk (dParity) is removed and designated as a hot spare, thus reducing each RAID group's size by one parity disk.
- Data ONTAP changes the setting for the aggregate's `raidsize` option to the size of the largest RAID group in the aggregate, except in the following situations:
  - If the aggregate's largest RAID group is larger than the maximum RAID4 group size, then the aggregate's `raidsize` option is set to the maximum.
  - If the aggregate's largest RAID group is smaller than the default RAID4 group size, then the aggregate's `raidsize` option is set to the default group size.
  - If the aggregate's `raidsize` option is already below the default value for RAID4, it is reduced by 1.

After you finish

You can verify the new RAID level by using the `aggr options` command.

Related concepts

How you use aggregates to provide storage to your volumes on page 125

Related tasks

Customizing the size of your RAID groups on page 117

Related references

Storage limits on page 345

Determining how the space in an aggregate is being used

Not all of the disk space you add to an aggregate is available for user data. You use the `aggr show_space` command to display how the disk space in an aggregate is being used.

About this task

If you specify the name of an aggregate, the command only displays information about that aggregate. Otherwise, the command displays information about all of the aggregates in the storage system.

For more information about the values returned by this command, see the na_aggr(1) man page.

Example

`aggr show_space aggr1`

```
Aggregate 'aggr1'

Total space WAFL reserve    Snap reserve    Usable space   BSR NVLOG
33GB          3397MB          1529MB            28GB             0KB

Space allocated to volumes in the aggregate

Volume          Allocated            Used       Guarantee
newvol             2344KB          2344KB       (offline)
vol1               1024MB          1328KB          volume
dest1               868KB           868KB          volume

Aggregate       Allocated            Used           Avail
Total space        1027MB          4540KB            27GB
```

```
Snap reserve       1529MB          6640KB          1522MB
WAFL reserve       3397MB          1280KB          3396MB
```

Destroying an aggregate

You destroy an aggregate when you no longer need the data in that aggregate or when you have copied the content of the aggregate to another location.

Before you begin

Before you can destroy an aggregate, you must destroy all of the FlexVol volumes contained by that aggregate.

About this task

When you destroy an aggregate, Data ONTAP converts its parity disks and its data disks back into hot spares. You can then use the spares in other aggregates and other storage systems.

Attention: If you destroy an aggregate, the data in the aggregate is no longer accessible by normal access methods, unless you undestroy it before any of its disks are zeroed or reused in another aggregate.

Note: If you want to make the data in the aggregate inaccessible by any means, you can sanitize its disks.

Note: You cannot destroy a SnapLock Compliance aggregate until the retention periods for all data contained in it have expired. For more information about the SnapLock functionality, see the *Data ONTAP Archive and Compliance Management Guide*.

Steps

1. Take the aggregate offline by entering the following command:

   aggr offline *aggr_name*

2. Destroy the aggregate by entering the following command:

   aggr destroy *aggr_name*

   The following message is displayed:

   ```
   Are you sure you want to destroy this aggregate ?
   ```

3. Enter the following command to confirm that you want to destroy the aggregate:

   y

   The following message is displayed:

   ```
   Aggregate 'aggr_name' destroyed.
   ```

Related concepts

How disk sanitization works on page 37

Undestroying an aggregate

If you previously destroyed an aggregate and have changed your mind, you can undestroy the aggregate if the data is still intact and the aggregate was not SnapLock-compliant.

Before you begin

You must know the name of the aggregate you want to undestroy, because there is no Data ONTAP command available to display destroyed aggregates, nor do they appear in FilerView.

Step

1. Undestroy the aggregate by entering the following command:

   aggr undestroy *aggr_name*

   Example

   aggr undestroy aggr1

   The following message is displayed:

   ```
   To proceed with aggr undestroy, select one of the following options [1]
   abandon the command [2] undestroy aggregate aggr1 ID:
   0xf8737c0-11d9c001-a000d5a3-bb320198 Selection (1-2)?
   ```

   If you select 2, a message with a date and time stamp appears for each disk that is restored to the aggregate. The message concludes with:

   ```
   Aggregate 'aggr1' undestroyed. Run wafliron to bring the aggregate
   online.
   ```

After you finish

After undestroying an aggregate, you must run the `wafliron` program with the privilege level set to `advanced`. If you need assistance, contact technical support.

Physically moving an aggregate composed of disks

To move an aggregate composed of disks from one storage system (the source) to another (the target), you need to physically move disks, disk shelves, or entire loops or stacks. You might move

an aggregate to move data to a new storage system model or remove data from an impaired storage system.

Before you begin

Ensure that the *target* storage system meets the following requirements:

- It must be running a version of Data ONTAP that is the same or later than the version running on the source system.
- It must support the shelf, module, and disk types being moved.
- It must support the size of the aggregate being moved.

About this task

The procedure described here applies to both aggregates with FlexVol volumes and to traditional volumes.

The procedure described here does *not* apply to aggregates composed of array LUNs.

Steps

1. Enter the following command at the source storage system to locate the disks that contain the aggregate:

   aggr status *aggr_name* -r

   The locations of the data, parity, and dParity disks in the aggregate appear under the HA, SHELF, and BAY columns (dParity disks appear for RAID-DP aggregates only).

2. Complete the appropriate steps, depending on whether you are moving an aggregate or a traditional volume.

If you are moving...	Then...
A traditional volume	Take the volume offline by entering the following command: **aggr offline *vol_name***
An aggregate	**a.** Boot the source storage system into maintenance mode. **b.** Take the aggregate offline by entering the following command: **aggr offline *aggr_name*** **c.** Reboot into normal mode.

3. If the storage system is using software-based disk ownership, remove the software ownership information from the disk by entering the following commands in the specified order for each disk:

   priv set advanced

   disk remove_ownership *disk_name*

priv set

4. Follow the instructions in the disk shelf hardware guide to remove the disks or shelves you identified previously from the source storage system.
5. Install the disks or disk shelves in the target storage system.

   When the target storage system sees the new disks, it sees the new aggregate as a *foreign aggregate*. Data ONTAP takes the foreign aggregate offline. If the foreign aggregate has the same name as an existing aggregate on the target storage system, Data ONTAP renames it *aggr_name*(1), where *aggr_name* is the original name of the aggregate.
6. If the target storage system uses software-based disk ownership, assign the disks that you moved to the target storage system by entering the following command for each moved disk:

   disk assign *disk_name*
7. Confirm that the foreign aggregate is complete by entering the following command:

   aggr status *aggr_name*

   Attention: If the foreign aggregate is incomplete (if it has a status of partial), add all missing disks before proceeding. Do not try to add missing disks after the aggregate comes online—doing so causes them to become hot spare disks. You can identify the disks currently used by the aggregate using the `aggr status -r` command.
8. If the storage system renamed the foreign aggregate because of a name conflict, enter the following command to rename the aggregate:

   aggr rename *aggr_name new_name*
9. Enter the following command to bring the aggregate online in the destination storage system:

   aggr online *aggr_name*
10. Enter the following command to confirm that the added aggregate came online:

    aggr status *aggr_name*
11. Boot the source storage system out of maintenance mode.

    For more information about maintenance mode, see the *Data ONTAP System Administration Guide*.

Moving an aggregate composed of array LUNs

You might want to move an aggregate composed of array LUNs to a less loaded system in the gateway neighborhood to balance the load processing over the systems.

Before you begin

- You should plan the number and size of your aggregates ahead of time so that you have flexibility in the amount of the workload that you can shift from one system in the gateway neighborhood to another.
- You should ensure that the *target* system meets the following requirements:
  - The target system must be running a version of Data ONTAP that is the same as or later than the version running on the source system.
  - The target system must support the size of the aggregate being moved.

About this task

To move the aggregate composed of array LUNs from one storage system (the source) to another (the target), you need to change the ownership of each array LUN in the aggregate from the source system to the target system. You can move both aggregates and traditional volumes using this procedure.

Note: If there are vFiler units in the aggregate you want to move, you might prefer to use SnapMover to move the aggregate. When SnapMover is used to move a vFiler unit, all aggregates in the vFiler unit are moved with the vFiler unit. To use vFiler units, you must have MultiStore software and SnapMover. See the *Data ONTAP MultiStore Management Guide* for more information.

Steps

1. Enter the following commands on the target system:
   a. Obtain the system ID of the target system by entering either of the following commands:

      `disk show`
      or

      `sysconfig`

      You need to provide the target system's ID on the source system when you assign each of the array LUNs to the target system.
2. Enter the following commands on the source system:
   a. Enter the following command to display the array LUNs that the aggregate contains:

      `aggr status` ***`aggr_name`*** **`-r`**

The array LUNs that are displayed are the LUNs that you need to reassign to the target system to be able to move the aggregate.

b. Write down the names of the array LUNs in the aggregate that you want to move.

c. Enter the following command to shut down the source system:

halt

d. At the boot environment prompt, enter the following command to boot the source system:

bye

e. Interrupt the boot process by pressing Ctrl-C when you see the following message on the console:
`Press Ctrl-C for Boot menu`

f. Enter Maintenance mode.

g. When prompted whether you want to continue with booting, enter the following:

y

h. Enter the following command to take the aggregate offline:

aggr offline *aggr_name*

aggr_name is the name of the traditional volume or aggregate.

i. Enter the following and confirm that the aggregate is offline:

aggr status

j. In Maintenance mode, enter the following command *separately* for each array LUN in the aggregate that you are moving to the target system:

disk assign -s *system_id_target* *disk_id* -f

system_id_target is the system ID of the target system (the system to which you want to move the array LUN.)

disk_id is the ID of the array LUN you want to move.

Note: Entering this command automatically removes ownership of the array LUN from the source system and assigns it to the target system.

3. Enter the following commands on the target system.

a. Enter the following command to start a scan so that the target system can recognize the LUNs you moved to it as its own:

disk show

b. Enter the following command:

aggr status

The display shows the *foreign* aggregate as offline. (The aggregate you are moving is a foreign aggregate to the target system.) If the foreign aggregate has the same name as an

existing aggregate on the system, Data ONTAP renames it *aggr_name(1)*, where *aggr_name* is the original name of the aggregate.

Attention: If the foreign aggregate is incomplete, that is, if you have not moved all the array LUNs in the aggregate, go back to the source system to add the missing array LUNs to the aggregate you moved to the target system. (Enter the following on the source system:

disk assign -s *system_id_target disk_id* -f

c. If Data ONTAP renamed the foreign aggregate because of a name conflict and you want to change the name, enter the following command to rename the aggregate :

aggr rename *aggr_name new_name*

aggr_name is the name of the aggregate you want to rename.

new_name is the new name of the aggregate.

Example
The following command renames the users(1) aggregate as newusers:

```
aggr rename users(1) newusers
```

d. Enter the following command to confirm that the aggregate you moved came online:

aggr status *aggr_name*

aggr_name is the name of the aggregate.

4. On the source system, reboot the system out of Maintenance mode.

How volumes work

Volumes contain file systems that hold user data that is accessible using one or more of the access protocols supported by Data ONTAP, including NFS, CIFS, HTTP, FTP, FC, and iSCSI.

Each volume depends on its containing aggregate for all its physical storage, that is, for all storage in the aggregate's disks and RAID groups.

Next topics

Related references

How FlexVol volumes work

A FlexVol volume is a volume that is loosely coupled to its containing aggregate. A FlexVol volume can share its containing aggregate with other FlexVol volumes. Thus, a single aggregate can be the shared source of all the storage used by all the FlexVol volumes contained by that aggregate.

Because a FlexVol volume is managed separately from the aggregate, you can create small FlexVol volumes (20 MB or larger), and you can increase or decrease the size of FlexVol volumes in increments as small as 4 KB.

When a FlexVol volume is created, it reserves a small amount of extra space (approximately 0.5 percent of its nominal size) from the free space of its containing aggregate. This space is used to store the volume's metadata. Therefore, upon creation, a FlexVol volume with a space guarantee of `volume` uses free space from the aggregate equal to its size × 1.005. A newly-created FlexVol volume with a space guarantee of `none` or `file` uses free space equal to .005 × its nominal size.

Note:

FlexVol volumes and traditional volumes have different best practices, optimal configurations, and performance characteristics. Make sure you understand these differences and deploy the configuration that is optimal for your environment.

Related references

Storage limits on page 345

How traditional volumes work

A traditional volume is a volume that is contained by a single, dedicated, aggregate. It is tightly coupled with its containing aggregate. No other volumes can get their storage from this containing aggregate.

The only way to increase the size of a traditional volume is to add entire disks to its containing aggregate. You cannot decrease the size of a traditional volume. The smallest possible traditional volume uses all the space on two disks (for RAID4) or three disks (for RAID-DP).

Attributes you can set for volumes

Volumes have a set of attributes that determine how they can be used.

You assign the following attributes to every volume, whether it is a traditional or FlexVol volume, except where noted:

- The name of the volume
- The size of the volume (assigned only for FlexVol volumes; the size of traditional volumes is determined by the size and number of their disks or array LUNs)
- A security style, which determines whether a volume can contain files that use UNIX security, files that use NT file system (NTFS) file security, or both types of files
- Whether the volume uses CIFS oplocks (opportunistic locks)
- The language of the volume
- The level of space guarantees (for FlexVol volumes only)
- Disk space and file limits (quotas, optional)
- A Snapshot copy schedule (optional)
- Whether the volume is designated as a SnapLock volume
- Whether the volume is a *root* volume

How the volume language attribute affects data visibility and availability

Every volume has a language. The language of the volume determines the character set Data ONTAP uses to display file names and data for that volume.

Attention: You are strongly advised to set all volumes to have the same language as the root volume, and to set the volume language at volume creation time. Changing the language of an existing volume can cause some files to become inaccessible.

The language of the root volume has special significance, because it affects or determines the following items:

- Default language for all volumes
- System name
- Domain name
- Console commands and command output
- NFS user and group names
- CIFS share names
- CIFS user account names
- Access from CIFS clients that don't support Unicode
- How configuration files in /etc are read
- How the home directory definition file is read

Note: Regardless of the language you specify for the root volume, names of the following objects must be in ASCII characters:

- Qtrees
- Snapshot copies
- Volumes
- Aggregates

For more information about the root volume, see the *System Administration Guide*.

How file access protocols affect what language to use for your volumes

Your choice of file access protocol (CIFS and NFS) affects the languages you should choose for your volumes.

Protocols in use	**Volume language**
NFS Classic (v2 or v3) only	Language setting does not matter
NFS Classic (v2 or v3) and CIFS	Language of the clients

Protocols in use	Volume language
NFS v4, with or without CIFS	*cl_lang*.UTF-8, where *cl_lang* is the language of the clients. **Note:** If you use NFS v4, all NFS Classic clients must be configured to present file names using UTF-8.

How you manage duplicate volume names

Data ONTAP does not support having more than one volume with the same name on a storage system. Data ONTAP renames such volumes, but the name it uses can cause problems, so you need to take corrective action.

When Data ONTAP detects a potential duplicate volume name, it appends the string "(*d*)" to the end of the name of the new volume, where *d* is a digit that makes the name unique.

For example, if you have a volume named vol1, and you copy a volume named vol1 from another storage system, Data ONTAP renames the newly copied volume to vol1(1).

You must rename any volume with an appended digit as soon as possible, for the following reasons:

- The name containing the appended digit is not guaranteed to persist across reboots. Renaming the volume prevents the name of the volume from changing unexpectedly later on.
- The parentheses characters, "(" and ")", are not legal characters for NFS. Any volume whose name contains those characters cannot be exported to NFS clients.
- The parentheses characters could cause problems for client scripts.

Volume states and status

Volumes can be in one of three states—online, offline, or restricted. In addition, they can show one or more status values, depending on how they are configured and the health of their disks.

You can determine a volume's current state and status by using the `vol status` command.

The following table displays the possible states for volumes.

State	Description
Online	Read and write access to this volume is allowed.
Restricted	Some operations, such as parity reconstruction, are allowed, but data access is not allowed.
Offline	No access to the volume is allowed.

State	Description
Quiesced	The volume is in the final stages of a move. Data access is not allowed, and many volume, qtree and quota management operations are temporarily unavailable.

The following table displays the possible status values for volumes.

Note: Although FlexVol volumes do not directly involve RAID, the state of a FlexVol volume includes the state of its containing aggregate. Thus, the states pertaining to RAID apply to FlexVol volumes as well as traditional volumes.

Status	Description
`access denied`	The origin system is not allowing access. (FlexCache volumes only.)
`active redirect`	The volume's containing aggregate is undergoing reallocation (with the `-p` option specified). Read performance may be reduced while the volume is in this state.
`connecting`	The caching system is trying to connect to the origin system. (FlexCache volumes only.)
`copying`	The volume is currently the target of an active `vol copy` or `snapmirror` operation.
`degraded`	The volume's containing aggregate contains at least one degraded RAID group that is not being reconstructed after single disk failure.
`double degraded`	The volume's containing aggregate contains at least one degraded RAID-DP group that is not being reconstructed after double disk failure.
`flex`	The volume is a FlexVol volume.
`flexcache`	The volume is a FlexCache volume.
`foreign`	Disks used by the volume's containing aggregate were moved to the current storage system from another storage system.
`growing`	Disks are being added to the volume's containing aggregate.
`initializing`	The volume's containing aggregate is being initialized.
`invalid`	The volume does not contain a valid file system.

Status	Description
`ironing`	A WAFL consistency check is being performed on the volume's containing aggregate.
`lang mismatch`	The language setting of the origin volume was changed since the caching volume was created. (FlexCache volumes only.)
`mirror degraded`	The volume's containing aggregate is mirrored and one of its plexes is offline or resynchronizing.
`mirrored`	The volume's containing aggregate is mirrored.
`needs check`	A WAFL consistency check needs to be performed on the volume's containing aggregate.
`out-of-date`	The volume's containing aggregate is mirrored and needs to be resynchronized.
`partial`	At least one disk was found for the volume's containing aggregate, but two or more disks are missing.
`raid0`	The volume's containing aggregate consists of RAID0 (no parity) groups (array LUNs only).
`raid4`	The volume's containing aggregate consists of RAID4 groups.
`raid_dp`	The volume's containing aggregate consists of RAID-DP groups.
`reconstruct`	At least one RAID group in the volume's containing aggregate is being reconstructed.
`redirect`	The volume's containing aggregate is undergoing aggregate reallocation or file reallocation with the `-p` option. Read performance to volumes in the aggregate might be degraded.
`rem vol changed`	The origin volume was deleted and re-created with the same name. Re-create the FlexCache volume to reenable the FlexCache relationship. (FlexCache volumes only.)
`rem vol unavail`	The origin volume is offline or has been deleted. (FlexCache volumes only.)
`remote nvram err`	The origin system is experiencing problems with its NVRAM. (FlexCache volumes only.)

Status	Description
`resyncing`	One of the plexes of the volume's containing mirrored aggregate is being resynchronized.
`snapmirrored`	The volume is in a SnapMirror relationship with another volume.
`trad`	The volume is a traditional volume.
`unrecoverable`	The volume is a FlexVol volume that has been marked unrecoverable; contact technical support.
`unsup remote vol`	The origin system is running a version of Data ONTAP the does not support FlexCache volumes or is not compatible with the version running on the caching system. (FlexCache volumes only.)
`verifying`	RAID mirror verification is running on the volume's containing aggregate.
`wafl inconsistent`	The volume or its containing aggregate has been marked corrupted; contact technical support .

Related concepts

About FlexCache volumes on page 191

About the CIFS oplocks setting

Usually, you should leave CIFS oplocks on for all volumes and qtrees. This is the default setting. However, you might turn CIFS oplocks off under certain circumstances.

CIFS oplocks (opportunistic locks) enable the redirector on a CIFS client in certain file-sharing scenarios to perform client-side caching of read-ahead, write-behind, and lock information. A client can then work with a file (read or write it) without regularly reminding the server that it needs access to the file. This improves performance by reducing network traffic.

You might turn CIFS oplocks off on a volume or a qtree under either of the following circumstances:

- You are using a database application whose documentation recommends that CIFS oplocks be turned off.
- You are handling critical data and cannot afford even the slightest data loss.

Otherwise, you can leave CIFS oplocks on.

For more information about CIFS oplocks, see the CIFS section of the *Data ONTAP File Access and Protocols Management Guide*.

Related tasks

Enabling or disabling CIFS oplocks for the entire storage system on page 296
Enabling CIFS oplocks for a specific volume or qtree on page 296
Disabling CIFS oplocks for a specific volume or qtree on page 296

How security styles affect access to your data

Every qtree and volume has a security style setting—NTFS, UNIX, or mixed. The setting determines whether files use Windows NT or UNIX (NFS) security. How you set up security styles depends on what protocols are licensed on your storage system.

Although security styles can be applied to volumes, they are not shown as a volume attribute, and are managed for both volumes and qtrees using the `qtree` command. The security style for a volume applies only to files and directories in that volume that are not contained in any qtree. The volume security style does not affect the security style for any qtrees in that volume.

The following table describes the three security styles and the effects of changing them.

Security Style	Description	Effect of changing to this style
NTFS	For CIFS clients, security is handled using Windows NTFS ACLs. For NFS clients, the NFS UID (user id) is mapped to a Windows SID (security identifier) and its associated groups. Those mapped credentials are used to determine file access, based on the NFTS ACL. **Note:** To use NTFS security, the storage system must be licensed for CIFS. You cannot use an NFS client to change file or directory permissions on qtrees with the NTFS security style.	If the change is from a mixed qtree, Windows NT permissions determine file access for a file that had Windows NT permissions. Otherwise, UNIX-style (NFS) permission bits determine file access for files created before the change. **Note:** If the change is from a CIFS storage system to a multiprotocol storage system, and the /etc directory is a qtree, its security style changes to NTFS.
UNIX	Files and directories have UNIX permissions.	The storage system disregards any Windows NT permissions established previously and uses the UNIX permissions exclusively.

Security Style	Description	Effect of changing to this style
Mixed	Both NTFS and UNIX security are allowed: A file or directory can have either Windows NT permissions or UNIX permissions. The default security style of a file is the style most recently used to set permissions on that file.	If NTFS permissions on a file are changed, the storage system recomputes UNIX permissions on that file. If UNIX permissions or ownership on a file are changed, the storage system deletes any NTFS permissions on that file.

Note: When you create an NTFS qtree or change a qtree to NTFS, every Windows user is given full access to the qtree, by default. You must change the permissions if you want to restrict access to the qtree for some users. If you do not set NTFS file security on a file, UNIX permissions are enforced.

For more information about file access and permissions, see the *Data ONTAP File Access and Protocols Management Guide*.

Next topics

How UNIX permissions are affected when files are edited using Windows applications on page 163
What the default security style is for new volumes and qtrees on page 164

How UNIX permissions are affected when files are edited using Windows applications

Many Windows applications incorrectly interpret the ACLs when reading files that have UNIX security. When the application saves the file, the original UNIX permissions are lost. Using the `cifs.preserve_unix_security` option avoids this problem.

You should set the `cifs.preserve_unix_security` option to `on` if you serve files under the following conditions:

- The files have UNIX permissions (that is, mode bits are set using the `chmod` or `umask` commands).
- NFS v4 Access Control Lists (ACLs) are not applied to the files.
- The files are in a qtree with UNIX or mixed security.
- The files are edited using Windows applications.

Note: When this option is enabled, a UNIX-style qtree appears as an NTFS volume instead of a FAT volume when viewed from a Windows client.

When the `cifs.preserve_unix_security` option is set, you can view and edit UNIX permissions using the Security tab in the Windows Properties dialog box . However, you cannot modify permissions from a Windows client if the operation is not permitted by the UNIX system. For example, you cannot change the ownership of a file you do not own, because the UNIX system

does not permit this operation. This restriction prevents Windows clients from bypassing UNIX permissions set on the storage system.

For more information about the `cifs.preserve_unix_security` option, see the options(1) man page.

What the default security style is for new volumes and qtrees

The default security style for new volumes and qtrees depends on whether your storage system is licensed for CIFS, NFS, or both.

License	Default security style
CIFS only	NTFS
NFS only	UNIX
CIFS and NFS	UNIX

How Data ONTAP can automatically provide more free space for full volumes

Data ONTAP can automatically make more free space available for a FlexVol volume when that volume is nearly full. You can choose to make the space available by first allowing the volume size to increase, or by first deleting Snapshot copies.

You enable this capability for a FlexVol volume by using the `vol options` command with the `try_first` option.

Data ONTAP can automatically provide more free space for the volume by using one of the following methods:

- Increase the size of the volume when it is nearly full.
  This method is useful if the volume's containing aggregate has enough space to support a larger volume. You can increase the size in increments and set a maximum size for the volume.

  Note: The autosize capability is disabled by default, so you must enable and configure it by using the `vol autosize` command. You can use the `vol status -v` command to view the current autosize settings for a volume.
- Delete Snapshot copies when the volume is nearly full.
  For example, you can automatically delete Snapshot copies that are not linked to Snapshot copies in cloned volumes or LUNs, or you can define which Snapshot copies you want to delete first—your oldest or newest Snapshot copies. You can also determine when to begin deleting Snapshot copies—for example, when the volume is nearly full or when the volume’s Snapshot reserve is nearly full.

You use the `snap autodelete` command to configure automatic Snapshot copy deletion. For more information about deleting Snapshot copies automatically, see the *Data ONTAP Data Protection Online Backup and Recovery Guide*.

You can choose which method (increasing the size of the volume or deleting Snapshot copies) you want Data ONTAP to try first. If the first method does not provide sufficient extra free space to the volume, Data ONTAP will try the other method next.

Related tasks

Configuring a FlexVol volume to grow automatically on page 184
Configuring automatic free space preservation for a FlexVol volume on page 184

About the maximum number of files allowed on a volume

Volumes have a maximum number of files that they can contain. You can increase the maximum number of files for a volume, but before doing so you should understand how this change affects the volume.

The storage system automatically sets the maximum number of files for a newly-created volume based on the amount of disk space in the volume. The storage system increases the maximum number of files when you increase the size of a volume, up to a 1 TB volume size. For volumes larger than 1 TB or volumes that contain an unusually large number of small files, you can use the `maxfiles` command to increase the maximum number of files if needed. The requirement for manually increasing the limit prevents a storage system with terabytes of storage from creating a larger than necessary inode file, which wastes storage.

When you increase the number of files a volume can contain, you are increasing that volume's number of inodes. An inode is a data structure that contains information about files. Increasing a volume's number of inodes increases the amount of space that volume uses for metadata. Once you increase the number of inodes for a volume, you cannot decrease it.

How to manage the root volume

The storage system's root volume contains special directories and configuration files that help you administer your storage system.

The root volume is installed at the factory on filers and on gateways ordered with disk shelves.

Note: For a gateway system that does not have a disk shelf, you need to install the root volume on the third-party storage. If you use a FlexVol volume for the root volume, you must ensure that it has a space guarantee of `volume`. For more information, see the *Data ONTAP Software Setup Guide*.

The factory-installed root volume is a FlexVol volume.

Unless the installer selected a unique volume name during setup, the default root volume name, `/vol/vol0`, is used.

Next topics

Recommendations regarding the root volume on page 166
Size requirement for root FlexVol volumes on page 167

Recommendations regarding the root volume

There are recommendations and considerations to keep in mind when choosing what kind of volume to use for the root volume.

The following are the general recommendations regarding the root volume:

- Root volumes can use either FlexVol or traditional volumes.
- For small storage systems where cost concerns outweigh resiliency, a FlexVol based root volume on a regular aggregate might be more appropriate.
- Avoid storing user data in the root volume, regardless of the type of volume used for the root volume.
- For a gateway system with a disk shelf, the root volume can reside on the disk shelf (recommended) or on the third-party storage. For a gateway system that does not have a disk shelf, the root volume resides on the third-party storage. You can install only one root volume per gateway system, regardless of the number of storage arrays or disk shelves that the gateway system uses for storage.

The following are additional facts and considerations if the root volume is on a disk shelf:

- Data ONTAP supports two levels of RAID protection, RAID4 and RAID-DP. RAID4 requires a minimum of two disks and can protect against single-disk failures. RAID-DP requires a minimum of three disks and can protect against double-disk failures. The root volume can exist as the traditional stand-alone two-disk volume (RAID4) or three-disk volume (RAID-DP). Alternatively, the root volume can exist as a FlexVol volume that is part of a larger hosting aggregate.
- Smaller stand-alone root volumes offer fault isolation from general application storage. On the other hand, FlexVol volumes have less impact on overall storage utilization, because they do not require two or three disks to be dedicated to the root volume and its small storage requirements.
- If a FlexVol volume is used for the root volume, file system consistency checks and recovery operations could take longer to finish than with the two- or three-disk traditional root volume. FlexVol recovery commands work at the aggregate level, so all of the aggregate's disks are targeted by the operation. One way to mitigate this effect is to use a smaller aggregate with only a few disks to house the FlexVol volume containing the root volume.
- In practice, having the root volume on a FlexVol volume makes a bigger difference with smaller capacity storage systems than with very large ones, in which dedicating two disks for the root volume has little impact.
- For higher resiliency, use a separate two-disk root volume.

  Note: You should convert a two-disk root volume to a RAID-DP volume when performing a disk firmware update, because RAID-DP is required for disk firmware updates to be

nondisruptive. When all disk firmware and Data ONTAP updates have been completed, you can convert the root volume back to RAID4.

For Data ONTAP 7.3 and later, the default RAID type for traditional root volume is RAID-DP. If you want to use RAID4 as the raid type for your traditional root volume to minimize the number of disks required, you can change the RAID type from RAID-DP to RAID4 by using `vol options vol0 raidtype raid4`.

The following requirement applies if the root volume is on a storage array:

- For storage systems whose root volume is on a storage array, only one array LUN is required for the root volume regardless of whether the root volume is a traditional volume or a FlexVol volume.

Size requirement for root FlexVol volumes

The root volume must have enough space to contain system files, log files, and core files. If a system problem occurs, these files are needed to provide technical support.

It is possible to create a FlexVol volume that is too small to be used as the root volume. Data ONTAP prevents you from setting the root option on a FlexVol volume that is smaller than the minimum root volume size for your storage system model. Data ONTAP also prevents you from resizing the root volume below the minimum allowed size or changing the space guarantee for the root volume.

The minimum size for a root FlexVol volume depends on your storage system model. The following table lists the required minimum size for root volumes. Check to ensure that the FlexVol volume to be used as the root volume meets the minimum size requirement. If you are using third-party storage, ensure that the array LUN you are using for the root volume is large enough to meet the minimum size requirement for the root volume.

Storage system model	**Minimum root FlexVol volume size**
N3700	10 GB
N3300	10 GB
N3400	16 GB
N3600	12 GB
N5200	12 GB
N5300	16 GB
N5500	16 GB
N5600	23 GB
N6040	16 GB
N6060	23 GB

Storage system model	Minimum root FlexVol volume size
N6070	37 GB
N6210	17 GB
N6240	22 GB
N6270	30 GB
N7600	37 GB
N7700	37 GB
N7800	69 GB
N7900	69 GB

Note: You cannot increase the root volume to more than 95 percent of the available aggregate size. The output of `df -A` displays the space used by the aggregates in the system.

The minimum array LUN size shown in the *Gateway Interoperability Matrix* does not apply to the root volume.

General volume operations

General volume operations are operations you can perform on either a FlexVol volume or a traditional volume. They include managing a volume's language, viewing or changing its state, renaming or destroying it, increasing the number of files it can contain, and running a reallocation operation on it.

Next topics

Migrating from traditional volumes to FlexVol volumes

You cannot convert directly from a traditional volume to a FlexVol volume. You must create a new FlexVol volume and then move the data to the new volume.

Before you begin

FlexVol volumes have best practices, optimal configurations, and performance characteristics different from those of traditional volumes. Make sure you understand these differences by referring to the available documentation on FlexVol volumes. Deploy the configuration that is optimal for your environment.

In addition, if your target volume is on the same storage system as the source volume, ensure that your system has enough free space to contain both copies of the volume during the migration.

About this task

If you are using this procedure to migrate your root volume, observe the notes specific to root volume migration.

If you want to migrate from a FlexVol volume to a traditional volume, you follow the same basic procedure, with the volume types reversed.

Note: Snapshot copies that currently exist on the source volume are not affected by this procedure. However, they are not replicated to the new target FlexVol volume as part of the migration.

Next topics

Preparing your destination volume on page 170
Migrating your data on page 172
Completing your migration on page 172

Related concepts

How volumes work on page 155
How FlexVol volumes work on page 155

Preparing your destination volume

Before migrating, you need to create and name a destination volume of the correct size and number of inodes.

About this task

If the new FlexVol volume will be the root volume, it must meet the minimum size requirements for root volumes, which are based on your storage system. Data ONTAP prevents you from designating as root a volume that does not meet the minimum size requirement. For more information, see the *System Administration Guide*.

Steps

1. Enter the following command to determine the amount of space your traditional volume uses:

   `df -Ah vol_name`

   Example

   ```
   sys1> df -Ah vol0
   Aggregate        total       used     avail    capacity
   vol0              24GB     1434MB      22GB          7%
   vol0/.snapshot  6220MB     4864MB    6215MB          0%
   ```

   The total space used by the traditional volume is displayed as `used` for the volume name.

2. Enter the following command to determine the number of inodes your traditional volume uses:

   `df -I vol_name`

Example

```
sys1> df -I vol0
Filesystem           iused      ifree  %iused  Mounted on
vol0               1010214   27921855     3%   /vol/vol0
```

The number of inodes your traditional volume uses is displayed as `iused`.

3. Identify or create an aggregate to contain the new FlexVol volume.

   Note: To determine if an existing aggregate is large enough to contain the new FlexVol volume, you can use the `df -Ah` command. The space listed under `avail` should be large enough to contain the new FlexVol volume.

4. If you want the destination (FlexVol) volume to have the same name as the source (traditional) volume, and they are on the same storage system, you must rename the source volume before creating the destination volume. Do this by entering the following command:

   `aggr rename *vol_name new_vol_name*`

   Example

   `aggr rename vol0 vol0trad`

5. Create the destination volume in the containing aggregate.

   Example

   `vol create vol0 aggrA 90g`

   Note: For root volumes, you must use the (default) volume space guarantee, because it ensures that writes to the volume do not fail due to a lack of available space in the containing aggregate.

6. Confirm that the size of the destination volume is at least as large as the source volume by entering the following command on the target volume:

   `df -h *vol_name*`

7. Confirm that the destination volume has at least as many inodes as the source volume by entering the following command on the destination volume:

   `df -I *vol_name*`

   Note: If you need to increase the number of inodes in the destination volume, use the `maxfiles` command.

Result

You have created a destination volume with sufficient resources to accept the data from the source volume.

Related tasks

Creating an aggregate on page 137

Creating a FlexVol volume on page 181

Migrating your data

You use the ndmpcopy command from the Data ONTAP prompt to migrate your data to the target volume.

Steps

1. Ensure that NDMP is configured correctly by entering the following commands:

   options ndmpd.enable on

   options ndmpd.authtype challenge

   Note: If you are migrating your volume between storage systems, make sure that these options are set correctly on both systems.

2. Disable data access to the source volume.
3. Migrate the data by entering the following command at the storage system prompt:

   ndmpcopy *src_vol_name dest_vol_name*

   Example

   ndmpcopy /vol/vol0trad /vol/vol0

   Attention: Make sure that you use the storage system command-line interface to run the ndmpcopy command. If you run this command from a client, your data will not migrate successfully.

   For more information about the ndmpcopy command, see the *Data ONTAP Data Protection Online Backup and Recovery Guide*.

4. Verify that the ndmpcopy operation completed successfully by validating the copied data.

Result

The target volume now contains the data from the source volume.

Completing your migration

After you copy your data, you need to perform some additional tasks before the migration is complete.

Steps

1. If you are migrating your root volume, complete the following steps:
   a. Make the new FlexVol volume the root volume by entering the following command:

      vol options *vol_name* root

Example

```
vol options vol0 root
```

b. Reboot the storage system.

2. Update the clients to point to the new FlexVol volume.

   - In a CIFS environment, complete these steps:

   a. Point CIFS shares to the new FlexVol volume.

   b. Update the CIFS maps on the client machines so that they point to the new FlexVol volume.

   - In an NFS environment, complete these steps:

   a. Point NFS exports to the new FlexVol volume.

   b. Update the NFS mounts on the client machines so that they point to the new FlexVol volume.

3. Make sure that all clients can see the new FlexVol volume and read and write data:

   a. Using a CIFS or NFS client, create a new folder or directory.

   b. Using the client, copy some scratch data into the new folder or directory and confirm that you can access that data from the client.

   c. Delete the new folder.

4. If you are migrating the root volume, and you changed the name of the root volume, update the `httpd.rootdir` option to point to the new root volume.

5. If quotas were used with the traditional volume, configure the quotas on the new FlexVol volume.

6. Take a Snapshot copy of the target volume and create a new Snapshot schedule as needed.

   For more information, see the *Data Protection Online Backup and Recovery Guide*.

7. Start using the migrated volume for the data source for your applications.

8. When you are confident the volume migration was successful, you can take the original volume offline or destroy it.

   Note: You should preserve the original volume and its Snapshot copies until the new FlexVol volume has been stable for some time.

Putting a volume into restricted state

You use the `vol restrict` command to put a volume into restricted state, which makes it unavailable for read or write access by clients. You might want to do this if you want the volume to be the target of a volume copy or SnapMirror replication operation.

About this task

When you restrict a FlexVol volume, it relinquishes any unused space that has been allocated for it in its containing aggregate. If this space is allocated for another volume and then you bring the volume back online, this can result in an overcommitted aggregate.

Related concepts

How volumes work on page 155
Considerations for bringing a volume online in an overcommited aggregate on page 283

Related tasks

Putting an aggregate into restricted state on page 144

Taking a volume offline

You use the `vol offline` command to take a volume offline to perform maintenance on the volume, move it, or destroy it. When a volume is offline, it is unavailable for read or write access by clients.

About this task

When you take a FlexVol volume offline, it relinquishes any unused space that has been allocated for it in its containing aggregate. If this space is allocated for another volume and then you bring the volume back online, this can result in an overcommitted aggregate.

Note: You cannot take the root volume offline.

Note: If you attempt to take a volume offline while any files contained by that volume are open, the `volume offline` command fails and displays the names (or inodes, if i2p is disabled) of the files that are open, along with the processes that opened them.

Related concepts

How volumes work on page 155
Considerations for bringing a volume online in an overcommited aggregate on page 283

Related tasks

Taking an aggregate offline on page 143

Bringing a volume online

After you restrict a volume or take it offline, you can make it available to the storage system again by bringing it online using the `vol online` command.

About this task

If you bring a FlexVol volume online into an aggregate that does not have sufficient free space to fulfill the space guarantee for that volume, this command fails.

Attention: If the volume you are bringing online is inconsistent, the `vol online` command prompts you for confirmation. If you bring an inconsistent volume online, it might suffer further file system corruption.

Related concepts

How volumes work on page 155

Considerations for bringing a volume online in an overcommited aggregate on page 283

Related tasks

Bringing an aggregate online on page 143

Renaming a volume

You use the `vol rename` command to rename a volume. You can rename volumes without interrupting data service.

Step

1. Enter the following command:

   `vol rename *vol_name new-name*`

Result

The following events occur:

- The volume is renamed.
- If NFS is in use and the `nfs.export.auto-update` option is On, the `/etc/exports` file is updated to reflect the new volume name.
- If CIFS is running, shares that refer to the volume are updated to reflect the new volume name.

- The in-memory information about active exports gets updated automatically, and clients continue to access the exports without problems.

After you finish

If you access the storage system using NFS, add the appropriate mount point information to the `/etc/fstab` or `/etc/vfstab` file on clients that mount volumes from the storage system.

Destroying a volume

If you no longer need a volume and the data it contains, you can destroy the volume to free up its space for other data.

About this task

When you destroy a FlexVol volume, all the disks included in its containing aggregate remain assigned to that containing aggregate, although the space associated with the volume is returned as free space to the containing aggregate.

When you destroy a traditional volume, however, you also destroy the traditional volume's dedicated containing aggregate. This converts its parity disk and all its data disks back into hot spares. After the disks have been zeroed, you can use them in other aggregates, traditional volumes, or storage systems.

Attention: If you destroy a volume, the data in the volume is no longer accessible.

Steps

1. Take the volume offline by entering the following command:

   `vol offline` *`vol_name`*

2. Enter the following command to destroy the volume:

   `vol destroy` *`vol_name`*

Result

The following events occur:

- The volume is destroyed.
- If NFS is in use and the `nfs.exports.auto-update` option is `on`, entries in the `/etc/exports` file that refer to the destroyed volume are removed.
- If CIFS is running, any shares that refer to the destroyed volume are deleted.
- If the destroyed volume was a FlexVol volume, its allocated space is freed, becoming available for allocation to other FlexVol volumes contained by the same aggregate.
- If the destroyed volume was a traditional volume, the disks it used become hot spare disks.

After you finish

If you access your storage system using NFS, update the appropriate mount point information in the `/etc/fstab` or `/etc/vfstab` file on clients that mount volumes from the storage system.

Changing the maximum number of files allowed in a volume

Volumes have a limit on the number of files they can contain. You can increase this limit using the `maxfiles` command, which increases the number of inodes allocated for the volume. However, you cannot decrease the limit after you have increased it.

About this task

You should use caution when increasing the maximum number of files, because after you increase this number, you can never decrease it. As new files are created, the file system consumes the additional disk space required to hold the inodes for the additional files; there is no way for the storage system to release that disk space.

Steps

1. Enter the following command:

   `maxfiles` ***`vol_name max_num_files`***

   Note: Inodes are added in blocks. If the requested increase in the number of files is too small to require a new inode block to be added, the maxfiles value is not increased. If this happens, repeat the command with a larger value for ***`max_num_files`***.

2. You can confirm the new maximum number of files, as well as the number of files currently present in the volume, by entering the following command:

   `maxfiles` ***`vol_name`***

   Note: The value returned reflects only the number of files that can be created by users; the inodes reserved for internal use are not included in this number.

Changing the language for a volume

You should use caution when changing the language for an existing volume, because doing so could affect the system's ability to display your data. In addition, a system reboot is necessary before the language change is complete.

Before you begin

Before changing the language that a volume uses, be sure you understand how volumes use the language attribute and how this change could affect access to your data.

Steps

1. Determine the correct language code for your volume.

   You can view the possible language codes by using the `vol lang` command.

2. Enter the following command to change the volume language:

   `vol lang` ***`vol_name language`***

   Note: If you are changing the NFS character set, you are asked to confirm your choice, and also whether you want to halt the system so that WAFL_check can be run to check for any files that will no longer be accessible using NFS. The default answer for this question is **`yes`**. If you do not want to halt the system, you must enter **`n`**.

3. Reboot the storage system.

   Note: Although the language change is effective for the target volume immediately, the full effect of the change is not complete until after the reboot.

After you finish

You can verify the new language by using the `vol status -l` command.

Related concepts

How the volume language attribute affects data visibility and availability on page 157
How volumes work on page 155

Changing the root volume

Every storage system must have a root volume. Therefore, you must always have one volume designated as the root volume. However, you can change which volume on your storage system is used as the root volume.

Before you begin

Before designating a volume to be the new root volume, ensure that the volume meets the minimum size requirement. The required minimum size for the root volume varies, depending on the storage system model. If the volume is too small to become the new root volume, Data ONTAP prevents you from setting the root option.

If you use a FlexVol volume for the root volume, ensure that it has a space guarantee of `volume`.

About this task

You might want to change the storage system's root volume, for example, when you migrate your root volume from a traditional volume to a FlexVol volume.

Steps

1. Identify an existing volume to use as the new root volume, or create the new root volume using the `vol create` command.
2. Using `ndmpcopy`, copy the `/etc` directory and all of its subdirectories from the current root volume to the new root volume. For more information about `ndmpcopy`, see the *Data ONTAP Data Protection Tape Backup and Recovery Guide*.
3. Enter the following command:

   `vol options` *`vol_name`* `root`

   `vol_name` is the name of the new root volume.

   After a volume is designated to become the root volume, it cannot be brought offline or restricted.

   Note: Besides the volume `root` option that you use to determine which volume will be the root volume after the next storage system reboot, there is also an aggregate `root` option. The aggregate `root` option is used only when, for some reason, the storage system cannot determine which volume to use as the root volume.

   If you move the root volume outside the current root aggregate, you must also change the value of the aggregate root option (using `aggr options` *`aggr_name`* `root`) so that the aggregate containing the root volume becomes the root aggregate.

   For more information about the aggregate `root` option, see the na_aggr(1) man page.

4. Enter the following command to reboot the storage system:

   `reboot`

   When the storage system finishes rebooting, the root volume is changed to the specified volume.

5. Update the `httpd.rootdir` option to point to the new root volume.

FlexVol volume operations

You can create FlexVol volumes, clone them, determine the amount of space they use, resize them, and display their containing aggregate, among other tasks.

Next topics

Creating a FlexVol volume on page 181
Resizing a FlexVol volume on page 183
Configuring a FlexVol volume to grow automatically on page 184
Configuring automatic free space preservation for a FlexVol volume on page 184
Displaying a FlexVol volume's containing aggregate on page 185

Related concepts

How FlexVol volumes work on page 155

Creating a FlexVol volume

You create FlexVol volumes to provide resizeable, flexible file systems that can be mounted and accessed using all data access protocols supported by Data ONTAP.

Before you begin

Before creating a FlexVol volume, you must first determine the following items:

- The name of the volume
  The volume name must conform to the following requirements:
  - Begin with either a letter or an underscore (_)
  - Contain only letters, digits, and underscores
  - Contain no more than 250 characters
  - Be different from all other volume names on the storage system
- The size of the volume
  The volume must be at least 20 MB in size. Its maximum size depends on whether it is in a 32-bit or a 64-bit aggregate and the model of the storage system that hosts the volume.
- The language used for the volume (optional)
  The default language is the language of the root volume.
- The space guarantee setting for the new volume (optional)
  The default space guarantee is `volume`.
- The CIFS oplocks setting for the new volume.
- The security style setting for the new volume.

Steps

1. If you have not already done so, create the aggregate that will contain the FlexVol volume that you want to create.
2. Enter the following command:

   vol create *vol_name* [-l *language_code*] [-s {volume|file|none}] *aggr_name* *size*{k|m|g|t}

   vol_name is the name for the new FlexVol volume (without the /vol/ prefix)

   language_code specifies a language other than that of the root volume.

   -s {volume|file|none} specifies the space guarantee setting that is enabled for the specified FlexVol volume. If no value is specified, the default value is volume

   aggr_name is the name of the containing aggregate for the new FlexVol volume.

   size{k | m | g | t} specifies the volume size in kilobytes, megabytes, gigabytes, or terabytes. For example, you would enter **20m** to indicate twenty megabytes. If you do not specify a unit, size is taken as bytes and rounded up to the nearest multiple of 4 KB.

   Example
   The following command creates a 200-MB volume called newvol, in the aggregate called aggr1, using the French character set:

   vol create newvol -l fr aggr1 200M

   The new volume is created and, if NFS is in use, an entry is added to the /etc/export file for the new volume. The default automatic snapshot schedule is applied to the new volume.
3. If you access the storage system using CIFS, update the share information for the new volume.
4. If you access the storage system using NFS, complete the following steps:
   a. Verify that the line added to the /etc/exports file for the new volume is correct for your security model.
   b. Add the appropriate mount point information to the /etc/fstab or /etc/vfstab file on clients that mount volumes from the storage system.

After you finish

Verify that the CIFS oplocks and security style settings are correct, and modify them as needed.

Note: You should set these values as soon as possible after creating the volume. If you change these values after files are in the volume, the files might become inaccessible to users because of conflicts between the old and new values. For example, UNIX files available under mixed security might not be available after you change to NTFS security.

If the default automatic snapshot schedule does not match your data protection strategies, update the snapshot schedule for the newly created volume with a more appropriate schedule. For more information, see the *Data ONTAP Data Protection Online Backup and Recovery Guide*.

Related concepts

How the volume language attribute affects data visibility and availability on page 157
What space guarantees are on page 279
About the CIFS oplocks setting on page 161
How security styles affect access to your data on page 162
How volumes work on page 155

Related tasks

Creating an aggregate on page 137

Related references

Storage limits on page 345

Resizing a FlexVol volume

You can increase or decrease the amount of space that an existing FlexVol volume is allowed to occupy in its containing aggregate. A FlexVol volume can grow to the size you specify as long as the containing aggregate has enough free space to accommodate that growth.

Steps

1. Check the available space of the containing aggregate by entering the following command:

   df -A *aggr_name*

2. If you want to determine the current size of the volume, enter one of the following commands:

   vol size *vol_name*

   df *vol_name*

3. Enter the following command to resize the volume:

   vol size *vol_name* [+|-] *n*{k|m|g|t}

   If you include the + or -, *n*{k|m|g|t} specifies how many kilobytes, megabytes, gigabytes or terabytes to increase or decrease the volume size. If you do not specify a unit, size is taken as bytes and rounded up to the nearest multiple of 4 KB.

   If you omit the + or -, the size of the volume is set to the size you specify, in kilobytes, megabytes, gigabytes, or terabytes. If you do not specify a unit, size is taken as bytes and rounded up to the nearest multiple of 4 KB.

   Note:

   If you attempt to decrease the size of a FlexVol volume to less than the amount of space that it is currently using, the command fails.

Decreasing the size of a FlexVol volume does not decrease the space reserved for metadata for the volume (it remains .5 percent of the original nominal size of the volume).

4. You can verify the success of the resize operation by entering the following command:

   vol size *vol_name*

Related references

Storage limits on page 345

Configuring a FlexVol volume to grow automatically

You configure FlexVol volumes to grow automatically to ensure that space in your aggregates is used efficiently, and to reduce the likelihood that your volumes will run out of space.

Step

1. Enter the following command:

   vol autosize *vol_name* [-m *size*] [-I *size*] on

   `-m` *size* is the maximum size to which the volume will grow. Specify a size in `k` (KB), `m` (MB), `g` (GB) or `t` (TB).

   `-I` *size* is the increment by which the volume's size increases. Specify a size in `k` (KB), `m` (MB), `g` (GB) or `t` (TB).

Result

If the specified FlexVol volume is about to run out of free space and is smaller than its maximum size, and if there is space available in its containing aggregate, its size will increase by the specified increment.

Related concepts

How Data ONTAP can automatically provide more free space for full volumes on page 164

Configuring automatic free space preservation for a FlexVol volume

When you configure a FlexVol volume for automatic free space preservation, the FlexVol volume attempts to provide more free space when it becomes nearly full. It can provide more free space by

increasing its size or by deleting Snapshot copies, depending on how you have configured the volume.

Step

1. Enter the following command:

   `vol options *vol-name* try_first [volume_grow|snap_delete]`

   If you specify `volume_grow`, Data ONTAP attempts to increase the volume's size before deleting any Snapshot copies. Data ONTAP increases the volume size based on specifications you provided using the `vol autosize` command.

   If you specify `snap_delete`, Data ONTAP attempts to create more free space by deleting Snapshot copies, before increasing the size of the volume. Data ONTAP deletes Snapshot copies based on the specifications you provided using the `snap autodelete` command.

Related concepts

How Data ONTAP can automatically provide more free space for full volumes on page 164

Displaying a FlexVol volume's containing aggregate

You display a FlexVol volume's containing aggregate by using the `vol container` command.

Traditional volume operations

Operations that apply exclusively to traditional volumes generally involve management of the aggregate to which that volume is closely coupled.

About this task

Additional traditional volume operations described in other chapters or other guides include:

- Configuring and managing SyncMirror replication of volume data
  See the *Data ONTAP Data Protection Online Backup and Recovery Guide*.
- Configuring and managing SnapLock volumes
  See the *Data ONTAP Archive and Compliance Management Guide*.

Related concepts

How Data ONTAP uses RAID to protect your data and data availability on page 101

Related tasks

Increasing the size of an aggregate on page 140
Changing the RAID level of an aggregate on page 145
Physically moving an aggregate composed of disks on page 149

Creating a traditional volume

Traditional volumes don't provide the flexibility that FlexVol volumes do, because they are tightly coupled with their containing aggregate. However, if you want a single-volume aggregate, you can create a traditional volume.

Before you begin

Determine the name of the volume. Volume names must conform to the following requirements:

- Begin with either a letter or an underscore (_)
- Contain only letters, digits, and underscores
- Contain no more than 250 characters

Note: You can change the name of an traditional volume later by using the `aggr rename` command.

Determine what disks will be used in the new volume. You can specify disks by listing their IDs, or by specifying a disk characteristic such as speed or type. You can display a list of the available spares on your storage system by using the `aggr status -s` command.

Determine the CIFS oplocks setting for the new volume.

Determine the security setting for the new volume.

Steps

1. Enter the following command:

   aggr create *vol_name* -v [-l *language_code*] [-f] [-m] [-n] [-v] [-t {raid4|raid_dp}] [-r *raidsize*] [-T *disk-type*] -R *rpm*] [-L] *disk-list*

   vol_name is the name for the new volume (without the /vol/ prefix).

   language_code specifies the language for the new volume. The default is the language of the root volume.

   Note: For a description of the RAID-related parameters, see the na_aggr(1) man page or the information about creating aggregates.

   The new volume is created and, if NFS is in use, an entry for the new volume is added to the `/etc/exports` file. The default automatic snapshot schedule is applied to the new volume.

2. You can enter the following command to verify that the volume exists as you specified:

   aggr status *vol_name* -r

   The system displays the RAID groups and disks of the specified volume on your storage system.

3. If you access the storage system using CIFS, update your CIFS shares as necessary.
4. If you access the storage system using NFS, complete the following steps:
   a. Verify that the line added to the `/etc/exports` file for the new volume is correct for your security model.
   b. Add the appropriate mount point information to the `/etc/fstab` or `/etc/vfstab` file on clients that mount volumes from the storage system.

After you finish

Verify that the CIFS oplocks and security style settings are correct, and modify them as needed.

Note: You should update these values as soon as possible after creating the volume. If you change the values after files are in the volume, the files might become inaccessible to users because of conflicts between the old and new values. For example, UNIX files available under mixed security might not be available after you change to NTFS security.

If the default automatic snapshot schedule does not match your data protection strategies, update the snapshot schedule for the newly created volume with a more appropriate schedule. For more information, see the *Data ONTAP Data Protection Online Backup and Recovery Guide*.

Related concepts

How the volume language attribute affects data visibility and availability on page 157
How to control disk selection from heterogeneous storage on page 131
About the CIFS oplocks setting on page 161
How security styles affect access to your data on page 162
How volumes work on page 155
How you use aggregates to provide storage to your volumes on page 125

Related tasks

Creating an aggregate on page 137

Related references

Storage limits on page 345

About FlexCache volumes

A FlexCache volume is a sparsely-populated volume on a local storage system that is backed by a volume on a different, possibly remote, storage system. A sparsely-populated volume, sometimes called a sparse volume, provides access to data in the remote volume without requiring that all the data be in the sparse volume.

You use FlexCache volumes to speed up access to remote data, or to offload traffic from heavily accessed volumes. Because the cached data must be ejected when the data is changed, FlexCache volumes work best for data that does not change often.

When a client requests data from the FlexCache volume, the data is read from the origin system and cached on the FlexCache volume. Subsequent requests for that data are then served directly from the FlexCache volume. This improves performance when the same data is accessed repeatedly, because after the first request, the data no longer has to travel across the network, or be served from an overloaded system.

Next topics

Related tasks

FlexCache hardware and software requirements

Before you can create FlexCache volumes and use them to access data in their origin volumes, you must ensure that both your origin and caching systems meet the hardware and software requirements for the FlexCache functionality.

The requirements for the caching system and the origin system are different.

For the caching system, the following requirements must be met:

- The caching system must have one of the following versions of Data ONTAP:
  - Data ONTAP 7.2.1 or later in the 7.2 release family
  - Any version in the Data ONTAP 7.3 release family

  Note: The caching and origin systems do not need to have the same version of Data ONTAP.
- A valid FlexCache license
- A valid NFS license, with NFS enabled

For the origin system, the following requirements must be met:

- The system must have one of the following versions of Data ONTAP:
  - Any version in the Data ONTAP 7.*x* release families
- A valid NFS license, with NFS enabled
- The `flexcache.access` option set to allow access to FlexCache volumes

  Note: For more information about this option, see the na_protocolaccess(8) man page.

  If the origin volume is in a vFiler unit, set this option for the vFiler context.
- The `flexcache.enable` option set to `on`

  Note: If the origin volume is in a vFiler unit, set this option for the vFiler context.

Limitations of FlexCache volumes

You can have a maximum of 100 FlexCache volumes on a storage system. In addition, there are certain features of Data ONTAP that are not available on FlexCache volumes, and others that are not available on volumes that are backing FlexCache volumes.

You cannot use the following Data ONTAP capabilities on FlexCache volumes (these limitations do not apply to the origin volumes):

- Client access using any protocol other than NFSv2 or NFSv3
- Client access using IPv6
- Snapshot copy creation

- SnapRestore
- SnapMirror (qtree or volume)
- SnapVault
- FlexClone volume creation
- The `ndmp` command
- Quotas
- Qtrees
- Volume copy
- Deduplication
- Creation of FlexCache volumes in any vFiler unit other than vFiler0
- Creation of FlexCache volumes in the same aggregate as their origin volume
- Mounting the FlexCache volume as a read-only volume

You cannot use the following Data ONTAP capabilities on FlexCache origin volumes or storage systems without rendering all of the FlexCache volumes backed by that volume or storage system unusable:

Note: If you want to perform these operations on an origin system, you can destroy the affected FlexCache volumes, perform the operation, and re-create the FlexCache volumes. However, the FlexCache volumes will need to be repopulated.

- You cannot move an origin volume between vFiler units or to vFiler0 using any of the following commands:
  - `vfiler move`
  - `vfiler add`
  - `vfiler remove`
  - `vfiler destroy`

    Note: You can use SnapMover (`vfiler migrate`) to migrate an origin volume without having to re-create FlexCache volumes backed by that volume.

    Origin volumes can be owned by any vFiler unit.
- You cannot use a FlexCache origin volume as the destination of a `snapmirror migrate` command.
- You cannot change the language of the origin volume if the change causes the underlying character set to change, or if the new language is not available on the caching system.

  For example, you can change the language of the origin volume from English to US English. However, if you want to change the language from English to a language that uses a different character set, such as Japanese, then you need to destroy and re-create all of the FlexCache volumes backed by the origin volume.
- Qtrees contained by the origin volume that belong to a vFiler unit other than the vFiler unit that owns the origin volume are not accessible to a FlexCache volume.

For example, suppose that volume `vol1` is owned by vFiler0 but `qtree1`, which is contained by `vol1`, is owned by another vFiler unit. FlexCache volumes created with `vol1` as the backing volume will not be able to access the data contained in `qtree1`.

Types of volumes you can use for FlexCache

A FlexCache volume must be a FlexVol volume. The origin volume can be a FlexVol or a traditional volume; it can also be a SnapLock volume. There are some restrictions on what can be used as an origin volume.

You cannot use the following storage containers as a FlexCache origin volume:

- A FlexCache volume
- A volume that contains SnapVault destinations
- A qtree

How the FlexCache Autogrow capability works

For best caching performance, you should allow Data ONTAP to control the size of your FlexCache volumes, by using the FlexCache Autogrow capability.

Making your FlexCache volume too small can negatively impact your caching performance. When the FlexCache volume begins to fill up, it flushes randomly chosen, previously cached files to make room for newly requested data. When data from the flushed files is requested again, it must be retrieved again from the origin volume.

Therefore it is best to use the Autogrow capability and allow Data ONTAP to increase the size of your FlexCache volumes as the size of the working set increases. This method has the following advantages:

- If the size of the FlexCache volume's working set increases, as long as there is space in the containing aggregate, the FlexCache volume automatically increases its size rather than ejecting data from the cache, which could affect data access performance.
- These size increases happen without operator intervention.
- If you have several FlexCache volumes sharing the same aggregate, the volumes that are getting the most data accesses will also receive the most space.
- If you increase the size of an aggregate, the FlexCache volumes contained by that aggregate will automatically take advantage of the extra space if needed.

The Autogrow capability is enabled by default in new FlexCache volumes created without specifying a size using Data ONTAP 7.3 and later. You can enable the Autogrow capability on existing FlexCache volumes by using the `vol options` command with the `flexcache_autogrow` option.

Note: Before the Autogrow capability was available, the preferred sizing strategy for FlexCache volumes was to size the FlexCache volume to the same size as its containing aggregate. If this

approach is providing you with the performance and space utilization you need, you do not need to reconfigure those existing FlexCache volumes to use the Autogrow capability.

How FlexCache volumes use space management

FlexCache volumes do not use space management in the same manner as regular FlexVol volumes. The amount of disk space reserved for a FlexCache volume is determined by the value of the `flexcache_min_reserved` volume option, rather than the nominal size of the FlexCache volume.

The default value for the `flexcache_min_reserved` volume option is 100 MB. In general, you should not change the value of this option.

Attention: FlexCache volumes' space guarantees must be honored. When you take a FlexCache volume offline, the space allocated for the FlexCache becomes available for use by other volumes in the aggregate (as with all FlexVol volumes). However, unlike regular FlexVol volumes, FlexCache volumes cannot be brought online if there is insufficient space in the aggregate to honor their space guarantee.

Related concepts

What space guarantees are on page 279
How volumes work on page 155

How FlexCache volumes share space with other volumes

You can have multiple FlexCache volumes in the same aggregate; you can also have regular FlexVol volumes in the same aggregate as your FlexCache volumes. To set up your system most efficiently, you should understand the way these volumes share space.

When you put multiple FlexCache volumes in the same aggregate, each FlexCache volume reserves only a small amount of space (as specified by the `flexcache_min_reserved` volume option—by default, 100 MB). The rest of the space is allocated as needed. This means that a "hot" FlexCache volume (one that is being accessed heavily) is permitted to take up more space, while a FlexCache volume that is not being accessed as often will gradually be reduced in size.

Note: When an aggregate containing FlexCache volumes runs out of free space, Data ONTAP randomly selects a FlexCache volume in that aggregate to be truncated. Truncation means that files are ejected from the FlexCache volume until the size of the volume is decreased to a predetermined percentage of its former size.

If you have regular FlexVol volumes in the same aggregate as your FlexCache volumes, and you start to fill up the aggregate, the FlexCache volumes can lose some of their unreserved space (if they are not currently using it). In this case, when the FlexCache volume needs to fetch a new data block and it does not have enough free space to accommodate it, a data block is ejected from one of the FlexCache volumes to make room for the new data block.

If ejected data is causing too many cache misses (as shown by the `flexcache stats` command), you can add more space to your aggregate or move some of your data to another aggregate.

How you display FlexCache statistics

Data ONTAP provides statistics about FlexCache volumes to help you understand the access patterns and administer the FlexCache volumes effectively.

You can display statistics for your FlexCache volumes using the following methods:

- The `flexcache stats` command (client and server statistics)
- The `nfsstat` command (client statistics only)
- The `perfstat` utility
- The `stats` command

For more information about the commands, see the na_flexcache(1), na_stats(1), and nfsstat(1) man pages.

Related tasks

Displaying FlexCache client statistics on page 207
Displaying FlexCache server statistics on page 208

What happens when connectivity to the origin system is lost

You can control how the FlexCache volume functions when connectivity between the caching and origin systems is lost by using the `disconnected_mode` and `acdisconnected` volume options.

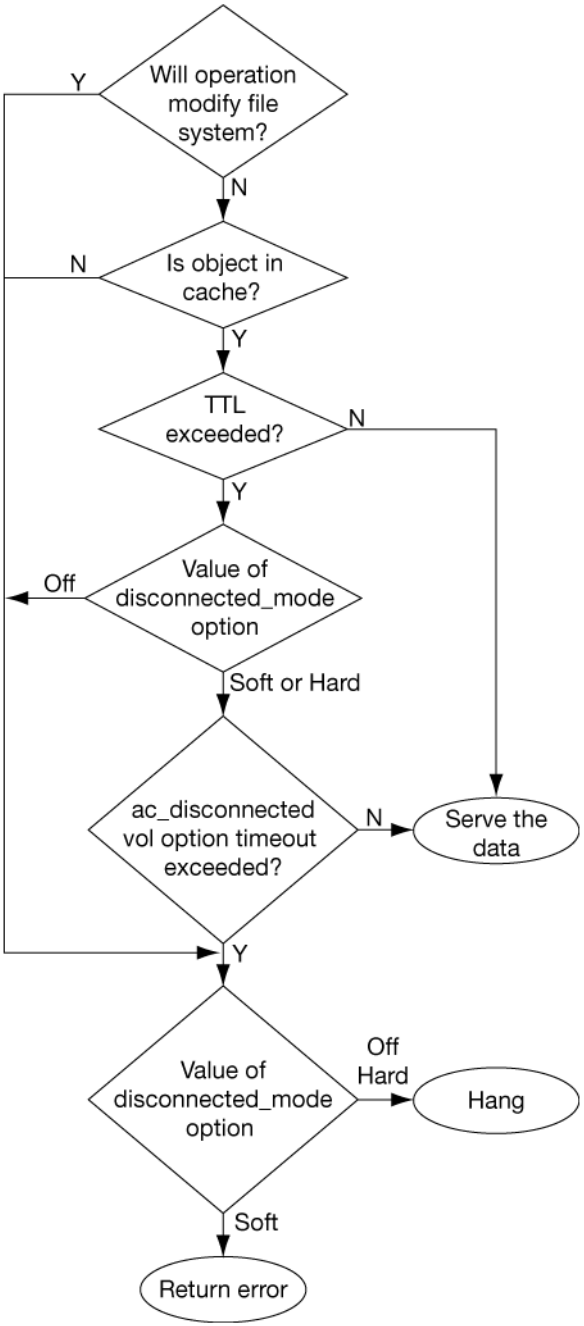
The `disconnected_mode` volume option and the `acdisconnected` timeout, combined with the regular TTL timeouts (`acregmax`, `acdirmax`, `acsymmax`, and `actimeo`), enable you to control the behavior of the FlexCache volume when contact with the origin volume is lost.

When you configure the FlexCache disconnected options, you should consider the following questions:

- Would your applications or file access protocols react better if an I/O request returned an error or if it did not return at all?
- How long can you safely serve stale data when connectivity is lost?

The following flowchart shows the multi-step decision process used by Data ONTAP to determine what happens when a FlexCache volume is disconnected from its origin volume. The possible outcomes of this process are:

- The data is served.
- An error is returned.
- The operation hangs.


Will operation modify file system?
Y
N
Is object in cache?
N
Y
TTL exceeded?
N
Y
Value of disconnected_mode option
Off
Soft or Hard
ac_disconnected vol option timeout exceeded?
N
Serve the data
Y
Value of disconnected_mode option
Off
Hard
Hang
Soft
Return error

How the NFS export status of the origin volume affects FlexCache access

A volume does not need to be exported to serve as an origin volume for a FlexCache volume. If you want to prevent a volume from being an origin volume, set the `flexcache.access` option to `none`.

How FlexCache caching works

Understanding how FlexCache determines the validity of cached data will help you determine whether your data set is a good candidate for a FlexCache.

Next topics

What it means for a file to be cached

When a data block from a specific file is requested from a FlexCache volume, then the attributes of that file are cached, and that file is considered to be cached, even if not all of its data blocks are present.

If the requested data is cached and valid, a read request for that data is fulfilled without access to the origin volume.

How data changes affect FlexCache volumes

How data changes affect FlexCache volumes depends on where the change is made: on the FlexCache volume, the origin volume, or another FlexCache volume.

Writes to a file on the origin volume

When a change is made to a file on the origin system, Data ONTAP revokes the delegation for that file and invalidates the entire file for all FlexCache volumes backed by that origin volume.

Note: The FlexCache copy of the file is not invalidated until an access to that file is made on the FlexCache volume.

The cache is not affected when only the access time of a file is updated.

Writes to a file on the FlexCache volume

When a write is made to a file on the FlexCache volume, the write request is relayed to the origin volume. When the origin volume acknowledges the request, the blocks that were changed are invalidated on the FlexCache volume, but the rest of the file remains valid.

Changes to a directory

When any change to a directory is made on either the FlexCache volume or the origin volume, that directory object is invalidated on all FlexCache volumes backed by that origin volume.

How cache consistency is achieved

Cache consistency for FlexCache volumes is achieved using three primary techniques: *delegations*, *attribute cache timeouts*, and *write operation proxy*.

Next topics

Delegations on page 199
Attribute cache timeouts on page 200
Write operation proxy on page 201

Delegations

You can think of a delegation as a contract between the origin system and the caching volume; as long as the caching volume has the delegation, the file has not changed. Delegations are used only in certain situations.

When data from a file is retrieved from the origin volume, the origin system can give a delegation for that file to the caching volume. Before that file is modified on the origin volume, whether due to a request from another caching volume or due to direct client access, the origin system revokes the delegation for that file from all caching volumes that have that delegation.

Delegations are not always used. The following list outlines situations when delegations cannot be used to guarantee that an object has not changed:

- Objects other than regular files
  Directories, symbolic links, and other objects that are not regular files have no delegations.
- Origin volumes that are SnapMirror destinations
  If the origin volume is a SnapMirror destination, delegations are not used.
- When connectivity is lost
  If connectivity is lost between the caching and origin systems, then delegations cannot be honored and must be considered to be revoked.
- When the maximum number of delegations has been reached
  If the origin volume cannot store all of its delegations, it might revoke an existing delegation to make room for a new one.

Note: Delegations can cause a small performance decrease for writes to the origin volume, depending on the number of caching volumes holding delegations for the file being modified.

If a FlexCache volume is taken offline, all its delegations are destroyed.

Attribute cache timeouts

When data is retrieved from the origin volume, the file that contains that data is considered valid in the FlexCache volume as long as a delegation exists for that file. If no delegation exists, the file is considered valid for a certain length of time, specified by the attribute cache timeout.

If a client requests data from a file for which there are no delegations, and the attribute cache timeout has been exceeded, the FlexCache volume compares the file attributes of the cached file with the attributes of the file on the origin system. Then one of the following actions is taken:

- If the two sets of file attributes match, the requested data is directly returned to the client (if it was already in the FlexCache volume) or retrieved from the origin system and then returned to the client.
- If the two sets of file attributes do not match, the file is marked as invalid in the cache. Then the requested data blocks are read from the origin system and stored in the FlexCache volume, as if it were the first time that file had been accessed from that FlexCache volume.

With attribute cache timeouts, clients can get stale data when all of the following conditions are true:

- There are no delegations for the file on the caching volume.
- The file's attribute cache timeout has not been reached.
- The file has changed on the origin volume since it was last accessed by the caching volume.

Note: Clients can get stale data when a file on the origin volume is added to or removed from a directory that is already stored on the FlexCache volume. The file addition or deletion does not become visible on the FlexCache until the length of time specified in the directory attribute cache timeout (`acdirmax`) has passed since the last time the directory was updated on the FlexCache volume.

To prevent clients from ever getting stale data, you can set the attribute cache timeout to `0`. However, this negatively affects your caching performance, because every data request for which there is no delegation causes an access to the origin system.

The attribute cache timeouts are determined by using volume options. The option names and default values are outlined in the following table.

Volume option	Description	Default value (seconds)
`acdirmax`	Attribute cache timeout for directories	`15s`
`acregmax`	Attribute cache timeout for regular files	`15s`

Volume option	Description	Default value (seconds)
`acsymmax`	Attribute cache timeout for symbolic links	`15s`
`actimeo`	Attribute cache timeout for all objects	`15s`

For more information about modifying these options, see the na_vol(1) man page.

Write operation proxy

If a client modifies a file that is cached, that operation is passed back, or proxied through, to the origin system, and the file is ejected from the cache.

When the write is proxied, the attributes of the file on the origin volume are changed. This means that when another client requests data from that file, any other FlexCache volume that has that data cached will re-request the data after the attribute cache timeout is reached.

Cache hits and misses

There are several types of cache hits and misses. Factors include whether data is present in the cache, whether the attribute cache timeout has been exceeded, and whether the file's attributes have changed.

When a client makes a read request, if the relevant block is cached in the FlexCache volume, the data is read directly from the FlexCache volume. This is called a *cache hit.* Cache hits are the result of a previous request.

A cache hit can be one of the following types:

- Hit
  The requested data is cached and no verification is required; the request is fulfilled locally and no access to the origin system is made.
- Hit-Verify
  The requested data is cached but the attribute cache timeout has been exceeded, so the file attributes are verified against the origin system. No data is requested from the origin system.

If data is requested that is not currently on the FlexCache volume, or if requested data has changed since it was cached, the caching system loads the data from the origin system and then returns it to the requesting client. This is called a *cache miss*.

A cache miss can be one of the following types:

- Miss
  The requested data is not in the cache; it is read from the origin system and cached.
- Miss-Verify
  The requested data is cached, but the file attributes have changed since the file was cached; the file is ejected from the cache and the requested data is read from the origin system and cached.

Typical FlexCache deployments

FlexCache is typically used in WAN deployments (which decrease average access time for remote clients) and LAN deployments (which reduce the workload of an overloaded storage system).

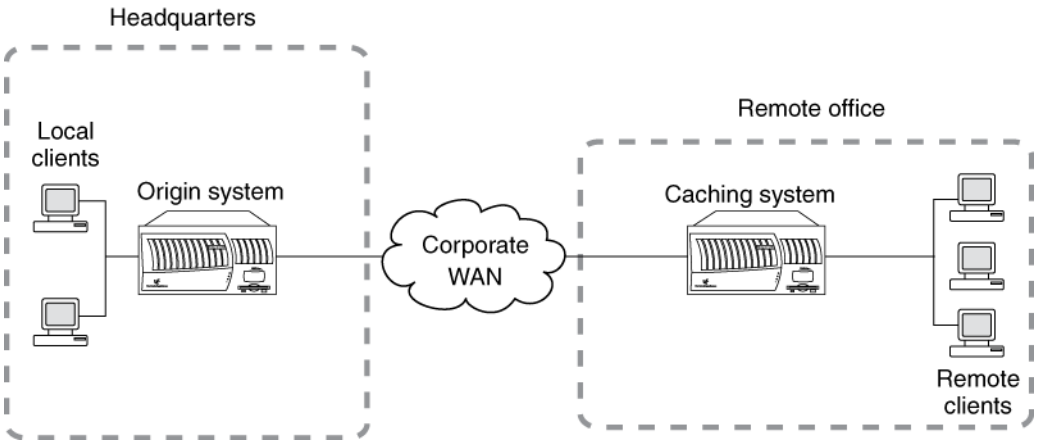
Next topics

WAN deployment on page 202
LAN deployment on page 202

WAN deployment

In a WAN deployment, the FlexCache volume is remote from the data center. As clients request data, the FlexCache volume caches popular data, giving the end user faster access to information.

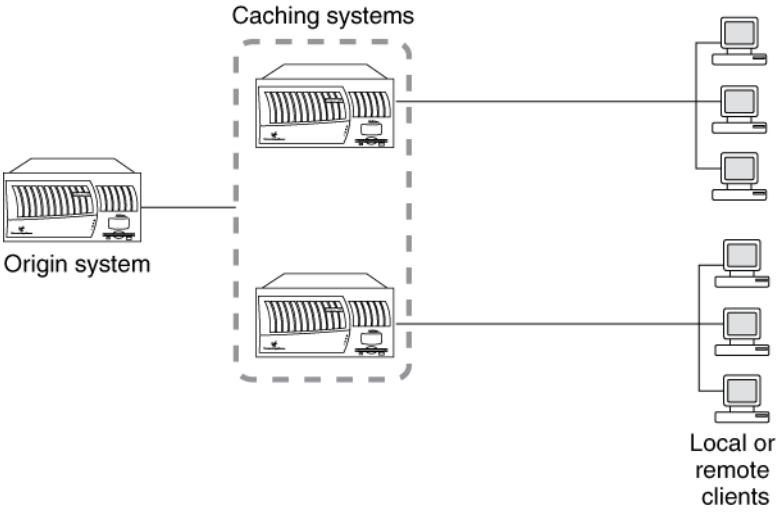
The FlexCache volume is placed as close as possible to the remote office. Client requests are then explicitly directed to the FlexCache volume. If valid data exists in the cache, that data is served directly to the client. If the data does not exist in the cache, it is retrieved across the WAN from the origin system, cached in the FlexCache volume, and returned to the client. A WAN deployment is shown in the following diagram.



LAN deployment

In a LAN deployment, or accelerator mode, the FlexCache volume is local to the administrative data center, and is used to offload work from busy file servers and free system resources.

Frequently accessed data, or "hot objects," are replicated and cached by the FlexCache volumes. This reduces network collisions and latency because the data access load is shared amongst all of the caching systems. A LAN deployment is shown in the following diagram.



About using LUNs in FlexCache volumes

You cannot use SAN access protocols to access FlexCache volumes. You can cache a volume that contains LUNs, but this configuration can change system behavior.

When you attempt to access, in a FlexCache volume, a directory that contains a LUN, the command sometimes returns "stale NFS file handle" for the LUN. If you get that error message, you should repeat the command.

If you use the `fstat` command on a LUN, fstat always indicates that the LUN is not cached. This is expected behavior.

Note: LUNs in this context refer to the LUNs that Data ONTAP serves to clients, not to the array LUNs used for storage on a storage array.

What FlexCache status messages mean

When you enter the `vol status` command for a FlexCache volume, and the status of the FlexCache volume is not normal, you get a FlexCache status message.

The following table lists the status messages you might see for a FlexCache volume and what they mean.

FlexCache status	Description
`access denied`	The origin system is not allowing FlexCache access. Check the setting of the `flexcache.access` option on the origin system.
`connecting`	The caching system is trying to connect to the origin system.

FlexCache status	Description
`lang mismatch`	The language setting of the origin volume was changed since the FlexCache volume was created.
`rem vol changed`	The origin volume was deleted and re-created with the same name. Re-create the FlexCache volume to reenable the FlexCache relationship.
`rem vol unavail`	The origin volume is offline or has been deleted.
`remote nvram err`	The origin system is experiencing problems with its NVRAM.
`unsup remote vol`	The origin system is running a version of Data ONTAP that either does not support FlexCache volumes or is not compatible with the version running on the caching system.

How FlexCache volumes connect to their origin volume

FlexCache volumes use a proprietary protocol to connect to their origin volume. The protocol uses port 2050.

FlexCache volume operations

Operations you can perform with FlexCache volumes include creating them, displaying their status and free space, configuring the Autogrow capability, and flushing files that they are caching.

Next topics

Related concepts

About FlexCache volumes on page 191

Creating FlexCache volumes

You use FlexCache volumes to speed up access to remote data, or to offload traffic from heavily accessed volumes.

Before you begin

Ensure that you have configured and enabled the FlexCache feature correctly on both the origin and caching systems.

Step

1. Enter the following command:

   vol create *cache_vol aggr* [*size*{k|m|g|t}] -S *origin*:*source_vol*

   origin is the name or IP address of the origin system. If you use the name, then changing the IP address of the origin system does not affect the FlexCache volume.

   cache_vol is the name of the new FlexCache volume you want to create.

   aggr is the name of the containing aggregate for the new FlexCache volume.

   size{ k | m | g | t } specifies the FlexCache volume size in kilobytes, megabytes, gigabytes, or terabytes. If you do not specify a unit, size is taken as bytes and rounded up to the nearest multiple of 4 KB.

Note: For best performance, do not specify a size when you create a FlexCache volume. Specifying a size disables the FlexCache Autogrow capability.

source_vol is the name of the volume you want to use as the origin volume on the origin system.

Result

The new FlexCache volume is created and an entry is added to the `/etc/export` file for the new volume.

Example

The following command creates a FlexCache volume called `newcachevol`, with the Autogrow capability enabled, in the aggregate called `aggr1`, with a source volume `vol1` on storage system corp_toaster:

```
vol create newcachevol aggr1 -S corp_toaster:vol1
```

Related concepts

FlexCache hardware and software requirements on page 192
How the FlexCache Autogrow capability works on page 194
About FlexCache volumes on page 191
How volumes work on page 155

Displaying free space for FlexCache volumes

When you use the `df` command on the caching storage system, you display the disk free space for the *origin* volume, rather than the local caching volume. You can display the disk free space for the local caching volume by using the `-L` option for the `df` command.

Configuring the FlexCache Autogrow capability

With the Autogrow capability enabled, Data ONTAP increases the size of a FlexCache volume when the volume starts to fill up. The Autogrow capability is enabled and disabled per FlexCache volume, and is enabled by default on new FlexCache volumes.

Step

1. Enter the command below, depending on the operation you want to perform:

If you want to..	Then enter...
Enable the Autogrow capability	`vol options vol_name flexcache_autogrow on`
Disable the Autogrow capability	`vol options vol_name flexcache_autogrow off`

Example

To enable the FlexCache Autogrow capability on the FlexCache volume fc1, enter the following command:

```
vol options fc1 flexcache_autogrow on
```

Related concepts

How the FlexCache Autogrow capability works on page 194

Flushing files from FlexCache volumes

If you know that a specific file has changed on the origin volume and you want to flush it from your FlexCache volume before it is accessed, you can use the `flexcache eject` command. For more information about this command, see the na_flexcache(1) man page.

Displaying FlexCache client statistics

You can use client statistics to see how many operations are being served by the FlexCache volume rather than the origin system. A large number of cache misses might indicate that the FlexCache volume is too small and data is being discarded and fetched again later.

Before you begin

Give the cache time to become populated before tracking cache misses.

Step

1. Depending on what statistics you want to see, enter the appropriate command.

If you want to...	Use this command:
Display FlexCache statistics	`flexcache stats -C`
Display NFS statistics for the FlexCache volume	`nfsstat -C`

Related concepts

How you display FlexCache statistics on page 196

Displaying FlexCache server statistics

If you are using the LAN deployment to offload an overloaded volume, you can use server statistics to get information about the origin system and ensure that the load is evenly distributed among the caching volumes.

Step

1. Depending on what statistics you want to see, enter the appropriate command.

If you want to...	Use this command:
Display overall server statistics	`flexcache stats -S`
Display server statistics per client	`flexcache stats -S -c` **Note:** To get per-client statistics, the `flexcache.per_client_stats` option must be set to `on`.

Related concepts

How you display FlexCache statistics on page 196

Displaying FlexCache status

You display the status for a FlexCache volume using the `vol status` command. If your FlexCache volume has a problem, a FlexCache status is displayed as the last line of the volume status output. If the status of the FlexCache is normal, no FlexCache status is displayed.

Related concepts

About FlexCache volumes on page 191

Related references

What FlexCache status messages mean on page 203

About FlexClone volumes

FlexClone volumes are writable, point-in-time copies of a parent FlexVol volume. Often, you can manage them as you would a regular FlexVol volume, but they also have some extra capabilities and restrictions.

FlexClone volumes are created when you clone a parent volume by using the `vol clone create` command. With the FlexClone license, you can also clone files and LUNs by using the `clone start` command.

Next topics

How FlexClone volumes work on page 209
Operations not supported on FlexClone volumes or their parents on page 210
FlexClone volumes and space guarantees on page 211
FlexClone volumes and shared Snapshot copies on page 212
How you can identify shared Snapshot copies in FlexClone volumes on page 212
How you use volume SnapMirror replication with FlexClone volumes on page 212
How splitting a FlexClone volume from its parent works on page 213
FlexClone volumes and LUNs on page 214

Related tasks

FlexClone volume operations on page 215

How FlexClone volumes work

FlexClone volumes can be managed similarly to regular FlexVol volumes, with a few key differences.

The following list outlines some key facts about FlexClone volumes:

- A FlexClone volume is a point-in-time, writable copy of the parent volume. Changes made to the parent volume after the FlexClone volume is created are not reflected in the FlexClone volume.
- You must install the license for the FlexClone feature before you can create FlexClone volumes.
- FlexClone volumes are fully functional volumes; you manage them using the `vol` command, just as you do the parent volume.
- FlexClone volumes always exist in the same aggregate as their parent volumes.
- Traditional volumes cannot be used as parent volumes for FlexClone volumes. To create a copy of a traditional volume, you must use the `vol copy` command, which creates a distinct copy that uses additional storage space equivalent to the amount of storage space used by the volume you copied.

- FlexClone volumes can themselves be cloned to create another FlexClone volume.
- FlexClone volumes and their parent volumes share the same disk space for any common data. This means that creating a FlexClone volume is instantaneous and requires no additional disk space (until changes are made to the FlexClone volume or its parent).
- A FlexClone volume is created with the same space guarantee as its parent. The space guarantee setting is enforced for the new FlexClone volume only if there is enough space in the containing aggregate.
- A FlexClone volume is created with the same space reservation and fractional reserve settings as its parent.
- While a FlexClone volume exists, some operations on its parent are not allowed.
- You can sever the connection between the parent volume and the FlexClone volume. This is called *splitting* the FlexClone volume. Splitting removes all restrictions on the parent volume and causes the FlexClone to use its own additional disk space rather than sharing space with its parent.

  Attention: Splitting a FlexClone volume from its parent volume deletes all existing Snapshot copies of the FlexClone volume, and disables the creation of new Snapshot copies while the splitting operation is in progress.
- Quotas applied to the parent volume are *not* automatically applied to the FlexClone volume.
- When a FlexClone volume is created, any LUNs present in the parent volume are present in the FlexClone volume but are unmapped and offline.

Related concepts

Operations not supported on FlexClone volumes or their parents on page 210
How splitting a FlexClone volume from its parent works on page 213
What space guarantees are on page 279

Related tasks

FlexClone volume operations on page 215

Related references

Storage limits on page 345

Operations not supported on FlexClone volumes or their parents

Not all Data ONTAP capabilities are available on FlexClone volumes.

The following restrictions apply to parent volumes or their clones:

- You cannot delete the base Snapshot copy in a parent volume while a FlexClone volume using that Snapshot copy exists. The base Snapshot copy is the Snapshot copy that was used to create the FlexClone volume, and is marked `busy, vclone` in the parent volume.
- You cannot perform a volume SnapRestore operation on the parent volume using a Snapshot copy that was taken before the base Snapshot copy was taken.
- You cannot destroy a parent volume if any clone of that volume exists.
- You cannot create a FlexClone volume from a parent volume that has been taken offline, although you can take the parent volume offline after it has been cloned.
- You cannot perform a `vol copy` command using a FlexClone volume or its parent as the destination volume.
- If the parent volume is a SnapLock Compliance volume, the FlexClone volume inherits the expiration date of the parent volume at the time of the creation of the FlexClone volume. The FlexClone volume cannot be deleted before its expiration date.
- There are some limitations on how you use SnapMirror with FlexClone volumes.

Related concepts

How you use volume SnapMirror replication with FlexClone volumes on page 212

FlexClone volumes and space guarantees

A FlexClone volume inherits its initial space guarantee from its parent volume. For example, if you create a FlexClone volume from a parent volume with a space guarantee of `volume`, then the FlexClone volume's initial space guarantee will be `volume` also. You can change the FlexClone volume's space guarantee.

For example, suppose that you have a 100-MB FlexVol volume with a space guarantee of `volume`, with 70 MB used and 30 MB free, and you use that FlexVol volume as a parent volume for a new FlexClone volume. The new FlexClone volume has an initial space guarantee of `volume`, but it does not require a full 100 MB of space from the aggregate, as it would if you had copied the volume. Instead, the aggregate needs to allocate only 30 MB (100 MB minus 70 MB) of free space to the clone.

If you have multiple clones with the same parent volume and a space guarantee of volume, they all share the same shared parent space with each other, so the space savings are even greater.

Note: The shared space depends on the existence of the shared Snapshot copy (the base Snapshot copy that was used to create the FlexClone volume). If you delete this shared Snapshot copy, you lose the space savings provided by the FlexClone volume.

Related concepts

FlexClone volumes and shared Snapshot copies on page 212
What space guarantees are on page 279

FlexClone volumes and shared Snapshot copies

When space guarantees are in effect, a new FlexClone volume uses the Snapshot copies it shares with its parent to minimize its space requirements. If you delete the shared Snapshot copies, you might increase the space requirements of the FlexClone volume.

For example, suppose that you have a 100-MB FlexVol volume that has a space guarantee of `volume`, with 70 MB used and 30 MB free, and you use that FlexVol volume as a parent volume for a new FlexClone volume. The new FlexClone volume has an initial space guarantee of `volume`, but it does not require a full 100 MB of space from the aggregate, as it would if you had copied the volume. Instead, the aggregate needs to allocate only 30 MB (100 MB – 70 MB) of free space to the clone.

Now, suppose that you delete a shared Snapshot copy from the FlexClone volume. The FlexClone volume can no longer optimize its space requirements, and the full 100 MB is required from the containing aggregate.

Note: If you are prevented from deleting a Snapshot copy from a FlexClone volume due to "insufficient space in the aggregate" it is because deleting that Snapshot copy requires the allocation of more space than the aggregate currently has available. You can either increase the size of the aggregate, or change the space guarantee of the FlexClone volume.

How you can identify shared Snapshot copies in FlexClone volumes

You can identify a shared Snapshot copy by listing the Snapshot copies *in the parent volume* with the `snap list` command. Any Snapshot copy that appears as `busy, vclone` in the parent volume and is also present in the FlexClone volume is a shared Snapshot copy.

How you use volume SnapMirror replication with FlexClone volumes

Because both volume SnapMirror replication and FlexClone volumes rely on Snapshot copies, there are some restrictions on how the two features can be used together.

Next topics

About creating a volume SnapMirror relationship using an existing FlexClone volume or its parent on page 213
About creating a FlexClone volume from volumes currently in a SnapMirror relationship on page 213

About creating a volume SnapMirror relationship using an existing FlexClone volume or its parent

You can create a volume SnapMirror relationship using a FlexClone volume or its parent as the *source* volume. However, you cannot create a new volume SnapMirror relationship using either a FlexClone volume or its parent as the *destination* volume.

About creating a FlexClone volume from volumes currently in a SnapMirror relationship

You can create a FlexClone volume from the source or destination volume in an existing volume SnapMirror relationship. However, doing so could prevent future SnapMirror replication operations from completing successfully.

Replication might not work because when you create the FlexClone volume, you might lock a Snapshot copy that is used by SnapMirror. If this happens, SnapMirror stops replicating to the destination volume until the FlexClone volume is destroyed or is split from its parent. You have two options for addressing this issue:

- If your need for the FlexClone volume is temporary, and you can accept the temporary cessation of SnapMirror replication, you can create the FlexClone volume and either delete it or split it from its parent when possible. At that time, the SnapMirror replication continues normally.
- If you cannot accept the temporary cessation of SnapMirror replication, you can create a Snapshot copy in the SnapMirror source volume, and then use that Snapshot copy to create the FlexClone volume. (If you are creating the FlexClone volume from the destination volume, you must wait until that Snapshot copy replicates to the SnapMirror destination volume.) This method allows you to create the clone without locking a Snapshot copy that is in use by SnapMirror.

How splitting a FlexClone volume from its parent works

Splitting a FlexClone volume from its parent removes any space optimizations that are currently employed by the FlexClone volume. After the split, both the FlexClone volume and the parent volume require the full space allocation determined by their space guarantees. The FlexClone volume becomes a normal FlexVol volume.

The following list contains facts about the clone splitting operation that you should know:

- When you split a FlexClone volume from its parent, all existing Snapshot copies of the FlexClone volume are deleted.
- No new Snapshot copies can be created of the FlexClone volume for the duration of the split operation.
- Because the clone-splitting operation is a copy operation that might take considerable time to carry out, Data ONTAP provides the `vol clone split stop` and `vol clone split status` commands to stop or check the status of a clone-splitting operation.

- The clone-splitting operation proceeds in the background and does not interfere with data access to either the parent or the clone volume.
- If you take the FlexClone volume offline while splitting is in progress, the operation is suspended; when you bring the FlexClone volume back online, the splitting operation resumes.
- After a FlexClone volume and its parent volume have been split, they cannot be rejoined.

Related tasks

Splitting a FlexClone volume from its parent on page 216

FlexClone volumes and LUNs

You can clone FlexVol volumes that contain LUNs and LUN clones.

Note: LUNs in this context refer to the LUNs that Data ONTAP serves to clients, not to the array LUNs used for storage on a storage array.

When you create a FlexClone volume, LUNs in the parent volume are present in the FlexClone volume but they are not mapped and they are offline. To bring the LUNs in the FlexClone volume online, you need to map them to igroups. When the LUNs in the parent volume are backed by Snapshot copies, the FlexClone volume also inherits the Snapshot copies.

If the parent volume contains LUN clones (LUNs created by using the `lun clone` command), the FlexClone volume inherits the LUN clones and their base Snapshot copies.

Note: The LUN clone's base Snapshot copy in the parent volume shares blocks with the base Snapshot copy in the FlexClone volume. You cannot delete the LUN clone's base Snapshot copy in the parent volume while the base Snapshot copy in the FlexClone volume still exists.

FlexClone volume operations

Operations you can perform with FlexClone volumes include creating a FlexClone volume and splitting it from its parent volume.

Next topics

Creating a FlexClone volume on page 215
Splitting a FlexClone volume from its parent on page 216
Determining the parent volume and base Snapshot copy for a FlexClone volume on page 217
Determining the space used by a FlexClone volume on page 217

Related concepts

How FlexClone volumes work on page 209

Creating a FlexClone volume

If you need a temporary copy of your data that can be made quickly and without using a lot of disk space, you can create a FlexClone volume. FlexClone volumes save data space because all unchanged data blocks are shared between the FlexClone volume and its parent.

Before you begin

Ensure that you have the flex_clone license installed.

Step

1. Enter the following command to clone the volume:

   `vol clone create` ***`clone_name`*** **`[-s {volume|file|none}] -b`** ***`parent_name`*** **`[`*** ***`parent_snap`*** **`]`**

   `clone_name` is the name of the FlexClone volume that you want to create.

   `-s {volume | file | none}` specifies the space guarantee setting for the new FlexClone volume. If no value is specified, the FlexClone volume is given the same space guarantee setting as its parent.

   `parent_name` is the name of the FlexVol volume that you intend to clone.

   `parent_snap` is the name of the base Snapshot copy of the parent FlexVol volume. If no name is specified, Data ONTAP creates a base Snapshot copy with the name `clone_`*`cl_name_prefix.id`*, where *`cl_name_prefix`* is up to 16 characters of the name of the new FlexClone volume and *`id`* is a unique digit identifier (for example 1, 2, and so on).

Note: The base Snapshot copy cannot be deleted as long as any clones based on that Snapshot copy exist.

Result

The FlexClone volume is created and, if NFS is in use, an entry is added to the `/etc/exports` file for every entry found for the parent volume.

The base Snapshot copy becomes a shared Snapshot copy between the FlexClone volume and its parent.

Example

To create a FlexClone volume named newclone from the parent FlexVol volume flexvol1, you would enter the following command:

```
vol clone create newclone -b flexvol1
```

Note: The Snapshot copy created by Data ONTAP is named clone_newclone.1.

After you finish

You can verify the status of the new FlexClone volume by using the `vol status -v` command.

Related concepts

About FlexClone volumes on page 209
What space guarantees are on page 279

Splitting a FlexClone volume from its parent

If you want the FlexClone volume to have its own disk space, rather than using that of its parent, you can split it from its parent.

Steps

1. Determine the approximate amount of free space required to split a FlexClone volume from its parent by entering the following command:

   ```
   vol clone split estimate clone_name
   ```

2. Verify that enough free space exists in the containing aggregate to support the split by entering the following command:

   ```
   df -A aggr_name
   ```

   The `avail` column tells you how much available space you have in your aggregate.

3. Enter the following command to split the volume:

`vol clone split start *clone_name*`

The clone-splitting operation begins. All existing Snapshot copies of the clone are deleted, and the creation of Snapshot copies of the clone is prevented for the duration of the split operation.

Note: If an online data migration operation is in progress, this command might fail. In this case, wait and retry the command when the online data migration operation is complete.

This operation could take some time to complete, depending on how much space is shared between the FlexClone volume and its parent.

If you take no further action, when all shared data has been copied, the clone will be split from its parent volume and become a regular FlexVol volume.

4. If you want to check the status of a clone-splitting operation, enter the following command:

   `vol clone split status *clone_name*`

5. If you want to stop the progress of an ongoing clone-splitting operation, enter the following command:

   `vol clone split stop *clone_name*`

   The clone-splitting operation halts; the original and FlexClone volumes remain clone partners, but they no longer share the disk space that was duplicated by the split.

6. You can display the status of the newly split FlexVol volume and verify the success of the clone-splitting operation by using the `vol status -v` command.

Related concepts

How splitting a FlexClone volume from its parent works on page 213

Determining the parent volume and base Snapshot copy for a FlexClone volume

You can determine the parent volume and base Snapshot copy for a FlexClone volume by using the `vol status` command.

Determining the space used by a FlexClone volume

You use a different method to determine the actual space used by FlexClone volumes than for other types of volumes, because a FlexClone volume shares data with its parent volume.

About this task

When a FlexClone volume is created, it shares all of its data with its parent volume. So even though its nominal size is the same as its parent's size, it uses very little free space from the aggregate. The

free space used by a newly-created FlexClone volume is approximately 0.5% of its nominal size. This space is used to store the FlexClone volume's metadata.

New data written to either the parent or the FlexClone volume is not shared between the volumes. The more new data that is written to the FlexClone volume, the more space the FlexClone volume requires from its containing aggregate.

Steps

1. Determine the nominal size of the FlexClone volume by entering the following command:

   df -m *clone_name*

2. Determine how much space is being shared between the parent and FlexClone volumes by entering the following command:

   vol clone split estimate *clone_name*

3. Subtract the size of the shared space from the nominal size of the FlexClone volume to determine the amount of free space being used by the FlexClone volume.

About FlexClone files and FlexClone LUNs

FlexClone files and FlexClone LUNs are writable, space-efficient clones of parent files and parent LUNs.

The Data ONTAP block-sharing mechanism is used for creating FlexClone files and LUNs. Clones use a small amount of storage space to store their metadata. Clones share the data blocks of their parent files and parent LUNs and occupy negligible storage space until clients write new data either to the parent file or LUN, or to the clone.

You can create FlexClone files and LUNs in the same FlexVol volume as their parent files and LUNs.

Clients can perform all normal file and LUN operations on both parent entities and clone entities.

Next topics

Related concepts

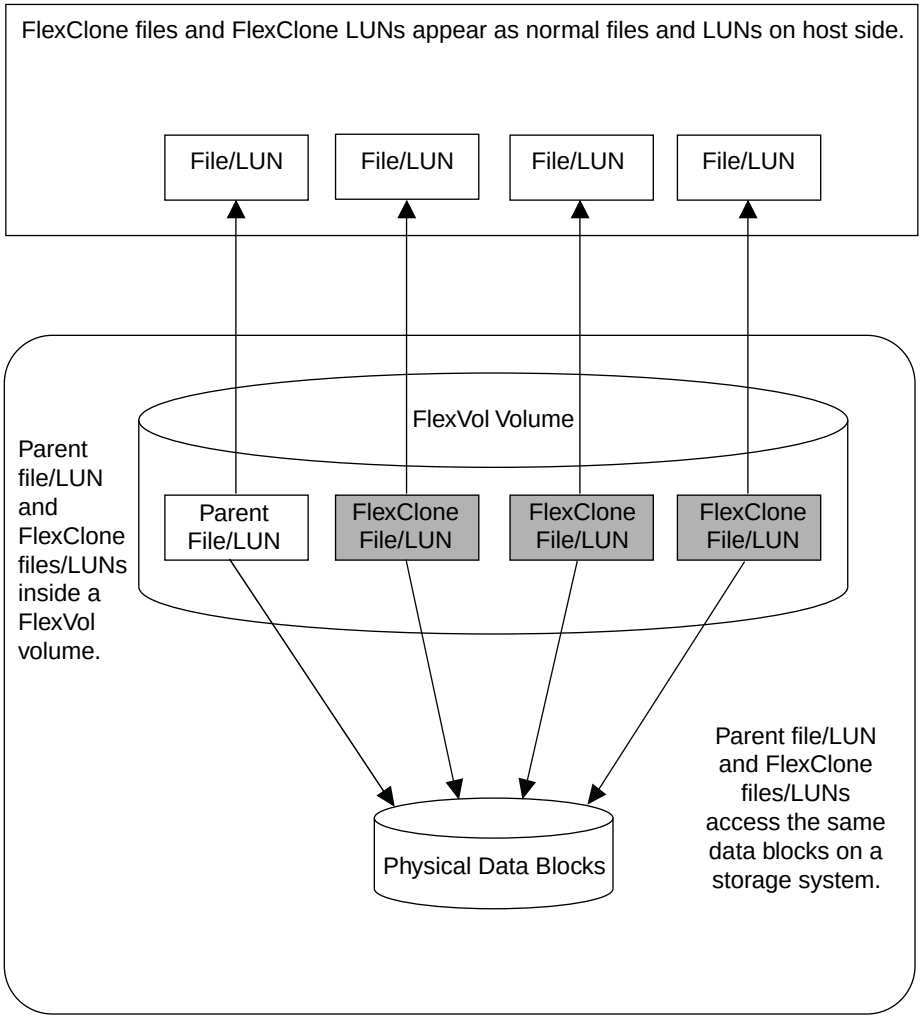
How FlexClone files and FlexClone LUNs work

Creating FlexClone files or FlexClone LUNs is highly space-efficient and time-efficient because the cloning operation does not involve physically copying any data.

You can create a clone of a file that is present in a FlexVol volume in a NAS environment, and you can also clone a complete LUN without the need of a backing Snapshot copy in a SAN environment.

The cloned copies initially share the same physical data blocks with their parents and occupy negligible extra space in the storage system for their initial metadata.

The following illustration shows the parent files or LUNs and FlexClone files or LUNs accessing the same data blocks on the storage system. On the host side, the parent files or LUNs and FlexClone files or LUNs appear as normal files and LUNs.



Unlike FlexClone volumes and LUN clones, the FlexClone files and FlexClone LUNs do not depend on a backing Snapshot copy. However, by default the cloning operation creates a temporary Snapshot copy of the FlexVol volume in which the cloning operation is being carried out. The temporary Snapshot copy is deleted immediately after a FlexClone file or LUN is created. You can stop the creation of a temporary Snapshot copy by using the `-n` option of the `clone start` command, but

you should do so only when you are certain that no writes will happen to the parent file or LUN during the cloning operation.

The cloning operation has no impact on client access to the parent file or LUN, either during the creation of clones or after the cloning operation is complete. Clients that are accessing the parent file or LUN do not experience any disruption or outage during the cloning operation. Clients can write to the source file or LUN while the cloning operation is in progress. Once the cloning operation is complete, clients see the FlexClone files or FlexClone LUNs as normal files and LUNs. Clients can perform all normal operations on them as they can on standard files and LUNs.

When clients write new data to a parent or clone, then the entity on which new data is written starts occupying extra storage space.

Related concepts

Differences between FlexClone LUNs and LUN clones on page 224
What happens when FlexClone file or LUN operation fails on page 248
FlexClone file and FlexClone LUN interoperability with Data ONTAP features on page 231
Considerations when creating FlexClone files or FlexClone LUNs on page 248

Related tasks

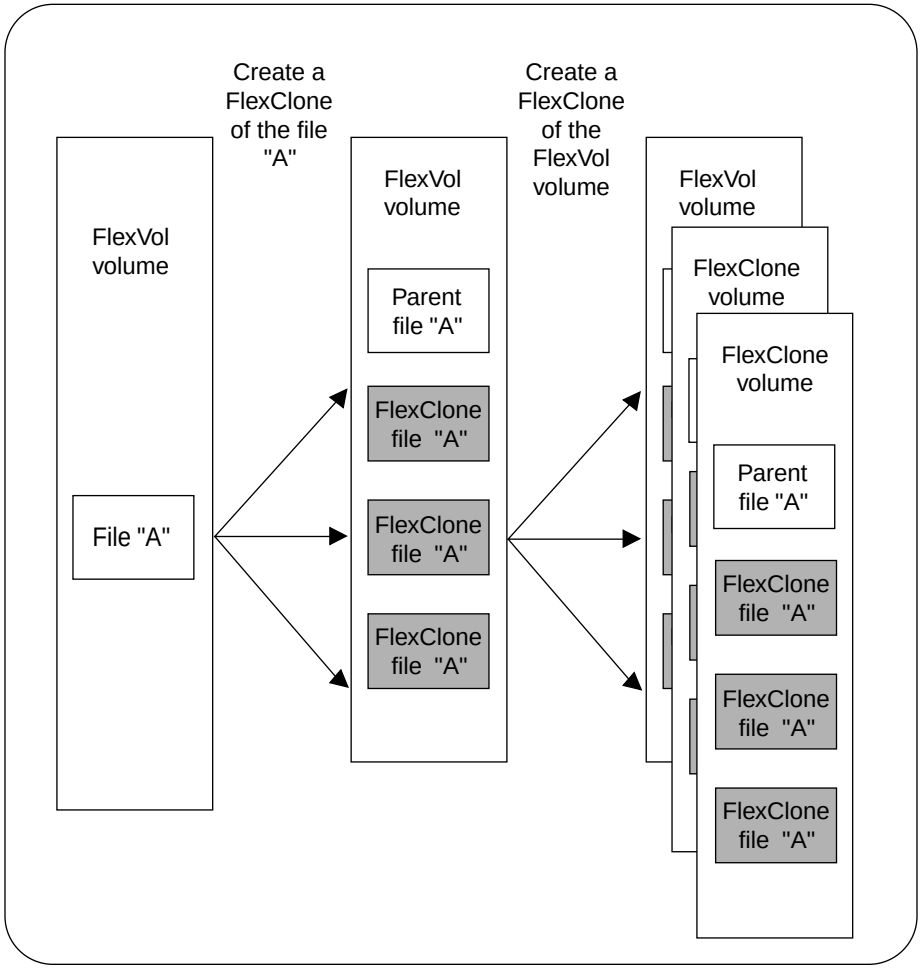
Creating a FlexClone file or FlexClone LUN on page 242
Viewing the status of a FlexClone file or FlexClone LUN operation on page 244

Collective usage of FlexClone at file, LUN, and volume level

You can use the FlexClone feature at file, LUN, and volume level to optimize storage space utilization.

The collective usage of the FlexClone feature at file, LUN, and volume level is a space-efficient and time-efficient solution for maintaining large number of duplicate copies of the same data.

As shown in the following illustration, you can create multiple FlexClone files of the parent file "A". For example, you might create three FlexClone files of the parent file. The illustration shows three FlexClone files of the parent file "A" in the FlexVol volume. The three FlexClone files share same data blocks of the parent file. Now you can clone at the FlexVol volume level and create multiple FlexClone volumes. For example, you might create two FlexClone volumes of the FlexVol volume. The two FlexClone volumes share data blocks with the parent FlexVol volume.



Now you have created multiple FlexClone files of the parent file "A", but all the FlexClone files access the same underlying physical storage. Thus, the storage space is used optimally.

Similarly, you can clone LUNs and files in a FlexVol volume and optimize the storage space utilization.

The FlexClone files or LUNs start occupying extra space only when the data is overwritten or when new writes begin.

Related concepts

About FlexClone volumes on page 209

Related tasks

Creating a FlexClone file or FlexClone LUN on page 242

Uses of FlexClone files and FlexClone LUNs

FlexClone files and FlexClone LUNs can help save time and storage space in a variety of situations.

You can quickly create space-efficient copies of your data by using FlexClone files and FlexClone LUNs in the following situations:

- When you need to deploy, upgrade, or redeploy thousands of standardized virtual desktops or servers
- When you need to test video, sound, or image processing applications
  You can use the cloned files for testing the applications.
- When you need to boot servers in a server farm
  You can create FlexClone LUNs of the parent boot LUN, then use the FlexClone LUN to boot a server in a server farm.

Considerations when planning FlexClone files or FlexClone LUNs

You should keep several considerations in mind when planning how to deploy FlexClone files and FlexClone LUNs.

- You can create FlexClone files and LUNs only in the same FlexVol volume as the parent files and LUNs.
- The following hardware platforms support FlexClone files and FlexClone LUNs:
  - N3300, N3400, and N3600 series
  - N5000 series
  - N6040, N6060, or N6070
  - N7600, N7700, N7800, or N7900
- You can create a FlexClone file or LUN only of a file or LUN that is part of the active file system. If you want to clone a file or LUN inside a Snapshot copy, you must first restore the entity to the active file system.
- You can clone a complete file, sub-file, LUN, or sub-LUN.
  To clone a sub-file or sub-LUN, you should know the block range of the parent entity and clone entity.
- The time required for creating a FlexClone file or FlexClone LUN depends on the size of the parent file or LUN.
- The `sis` attribute is added to a FlexVol volume when a FlexClone file or FlexClone LUN is created for the first time.

FlexVol volumes with deduplication enabled also show the `sis` attribute when you run the `vol status` command.

Related concepts

About FlexClone files and FlexClone LUNs on page 219
How deduplication works with FlexClone files and FlexClone LUNs on page 233

Related tasks

Creating a FlexClone file or FlexClone LUN on page 242

Differences between FlexClone LUNs and LUN clones

Data ONTAP provides two LUN cloning capabilities—LUN clone with the support of a Snapshot copy and FlexClone LUN. However, there are a few differences between these two LUN cloning techniques.

The following table lists the key differences between the two LUN cloning features.

FlexClone LUN	LUN clone
To create a FlexClone LUN, you should use the `clone start` command.	To create a LUN clone, you should use the `lun clone create` command.
You need not create a Snapshot copy manually.	You need to create a Snapshot copy manually before creating a LUN clone, because a LUN clone uses a backing Snapshot copy
A temporary Snapshot copy is created during the cloning operation. The Snapshot copy is deleted immediately after the cloning operation. However, you can prevent the Snapshot copy creation by using the `-n` option of the `clone start` command.	A LUN clone is coupled with a Snapshot copy.
A FlexClone LUN is independent of Snapshot copies. Therefore, no splitting is required.	When a LUN clone is split from the backing Snapshot copy, it uses extra storage space. The amount of extra space used depends on the type of clone split.
You can clone a complete LUN or a sub-LUN. To clone a sub-LUN, you should know the block range of the parent entity and clone entity.	You can only clone a complete LUN.
FlexClone LUNs are best for situations where you need to keep the clone for a long time.	LUN clones are best when you need a clone only for a short time.

FlexClone LUN	LUN clone
No Snapshot copy management is required.	You need to manage Snapshot copies if you keep the LUN clones for a long time.

For more information about LUN clones, see the *Data ONTAP Block Access Management Guide for iSCSI and FC*.

Related concepts

How FlexClone files and FlexClone LUNs work on page 219
How Snapshot copies work with FlexClone files and FlexClone LUNs on page 231

Related tasks

Creating a FlexClone file or FlexClone LUN on page 242

Operational limits for FlexClone files and FlexClone LUNs

There are limits on the number of FlexClone files or LUNs you can create, and on the amount of shared data in a volume.

Maximum number of FlexClone files or FlexClone LUNs

You can create a maximum of 255 FlexClone files or FlexClone LUNs from a parent file or LUN without creating a physical copy of the parent entity. If you try to create more than 255 clones, Data ONTAP automatically creates a new physical copy of the parent file or LUN.

Note: The block-sharing mechanism used by FlexClone files and LUNs is also used by deduplication. Therefore, if deduplication was enabled or is currently enabled on a FlexVol volume, you might end up creating a new physical copy of the parent entity even before creating the maximum of 255 FlexClone files or LUNs of a file or LUN.

Maximum limit on shared data in a FlexVol volume with FlexClone files and FlexClone LUNs

The total logical size of all FlexClone files and FlexClone LUNs in a FlexVol volume is 16 TB. If you attempt to create FlexClone file or LUN after the maximum size is reached, Data ONTAP automatically creates a new physical copy of the parent file or LUN.

Note: The 16 TB limit is on the sum of logical sizes of the FlexClone files or FlexClone LUNs. The total physical space actually used in the FlexVol volume by the parent entities and clone entities might be less, because the parent entities and clone entities share the same physical data blocks with little extra space required for the metadata of each clone.

For example, if you have a parent file of size 4 TB in a FlexVol volume, you can create four FlexClone files of the parent file. The sum of logical sizes of the FlexClone files is 16 TB. If you try

to create a fifth FlexClone file of the parent file, Data ONTAP instead creates a physical copy of the file by copying the complete file to the destination location of the clone. Similarly, if you try to clone any other file or LUN in the same FlexVol volume, Data ONTAP creates a physical copy instead of a clone.

Maximum FlexVol volume size for FlexClone files and FlexClone LUNs

A FlexVol volume with FlexClone files or LUNs has a smaller maximum size than a FlexVol volume without FlexClone files or LUNs. The maximum depends on your hardware platform, as shown in the table.

Storage system model	Maximum FlexVol volume size for FlexClone files or FlexClone LUNs (TB)
N3300	1
N3400	3
N3600	2
N5200	2
N5300	4
N5500	3
N5600	16
N6040	4
N6060	16
N6070	16
N7600	16
N7700	16
N7800	16
N7900	16

Note:

- After creating a FlexClone file or FlexClone LUN on a FlexVol volume, you cannot manually grow the volume size more than its limit for that particular platform.
- When enabling `vol autosize` on a FlexVol volume that has a FlexClone file or FlexClone LUN, you should ensure that the `autosize` setting is less than the maximum allowed volume size for that platform.

Maximum simultaneous FlexClone file or LUN operations

You can simultaneously run a maximum of 16 FlexClone file or FlexClone LUN operations on a single FlexVol volume. Any new FlexClone operation beyond this limit fails to start.

Maximum number of status entries in the metadata file

For managing cloning operations a small amount of metadata is stored on a disk in the metadata file for each running and failed cloning operations. The metadata file can have information about a maximum of 31 running and failed FlexClone file or FlexClone LUN operations. Once this limit is reached, you cannot start a new FlexClone file or FlexClone LUN operation.

When you start a new clone operation on a FlexVol volume that contains the maximum number of status entries in the metadata file, Data ONTAP displays an error message saying that no free slot is available to log the cloning operation. Before you can start a new clone operation, you must clear entries of failed cloning operations in the metadata file. To clear the metadata file, you use the `clone clear` command. Entries of successfully completed cloning operations are automatically cleared from the metadata file.

Maximum simultaneous FlexClone file or FlexClone LUN operations per storage system

You can simultaneously run a maximum of 500 FlexClone file or FlexClone LUN operations on a storage system.

Related concepts

How volume autosize works with FlexClone files and FlexClone LUNs on page 237

Related tasks

Creating a FlexClone file or FlexClone LUN on page 242
Clearing the status of a failed FlexClone file or FlexClone LUN operation on page 246

What happens when clients write new data to parent or FlexClone files and FlexClone LUNs

When new data is written either to a FlexClone file or FlexClone LUN, or to a parent file or LUN, the new data occupies additional storage space.

When the FlexClone file or LUN is first created, the parent file or LUN shares the same physical data blocks with the cloned file or LUN. However, when clients write new data to the parent file or LUN, or to its clones, then they start using extra storage space.

The parent file or LUN and its clones do not share the newly written data. The new data is stored separately for the parent file or LUN and for clones. Even if the same data is written to both parent

files or LUNs and clones, the data is written on different blocks on the disk and these data blocks are not shared between clones and parents.

Related concepts

About FlexClone files and FlexClone LUNs on page 219
How FlexClone files and FlexClone LUNs work on page 219

What happens when FlexClone files, FlexClone LUNs, or parents are deleted

FlexClone files or FlexClone LUNs and their parent files or LUNs can be deleted. Deleting parents or clones free the space they are using.

Deleting a parent file or LUN has no impact on the FlexClone file or FlexClone LUN. Clients can still see the clone files or LUNs as normal files and LUNs. Similarly, deleting a FlexClone file or FlexClone LUN has no impact on the parent file or LUN.

When a file or LUN or its clones that use shared blocks are deleted, then any remaining file or LUN (FlexClone or parent) continues to use the shared blocks.Therefore, deleting FlexClone files or FlexClone LUNs frees the space that is being used by their metadata, and any data that was overwritten in or newly written to the clone. However, if the parent file or LUN and all corresponding FlexClone files or FlexClone LUNs are deleted, then all the data blocks are freed.The freed storage space is added to the free storage space pool.

Related concepts

About FlexClone files and FlexClone LUNs on page 219
How FlexClone files and FlexClone LUNs work on page 219

Space savings achieved by using FlexClone files and FlexClone LUNs

You can use the `df -s` command to view the amount of storage space saved by creating FlexClone files and FlexClone LUNs. When you create a clone, you save the amount of space that is occupied by its parent.

Note: If you run the `df -s` command on a FlexVol volume with deduplication enabled, the output displays the space saved by both deduplication and FlexClone files or FlexClone LUNs.

Example

If you have a FlexVol volume of 100 GB with 50 GB used space and then create a file of 10 GB and a clone of it, the total used physical space is about 60 GB (50 GB + 10 GB for file and its clone). If the clone were a full physical copy, you would be using 70 GB (50 GB + 10 GB

for file + 10 GB for the clone). Therefore, you saved space of 10 GB by creating a FlexClone file. Your savings are 14% ((10/70)*100).

Related concepts

How deduplication works with FlexClone files and FlexClone LUNs on page 233

Related tasks

Viewing the space savings due to FlexClone files and FlexClone LUNs on page 246

File space utilization report

The file space utilization report enables you to see the files and the amount of space that they occupy in a deduplicated volume. You can choose to either move or delete the files to reclaim the space.

This report provides a view of the total number of blocks in a file and the number of blocks that are shared by non-deduplicated or non-cloned files.

Note: Total blocks refer to the number of blocks in a file, including blocks that are required for storing the file metadata.

What the FlexClone log file is

The FlexClone log file (clone log file) provides history of all the FlexClone file or FlexClone LUN cloning operations performed on the storage system. In the clone log file, you can view the details of all successful, unsuccessful, or stopped cloning operations.

The clone log file reside in the `/etc/log/clone` directory.

The clone log file records the following information:

- Cloning operation ID
- The name of the volume in which the cloning operation was performed
- Start time of the cloning operation
- End time of the cloning operation
- Parent file/LUN and clone file/LUN names
- Parent file/LUN ID
- Status of the clone operation: successful, unsuccessful, or stopped and some other details

Data ONTAP maintains seven weeks' worth of information about FlexClone file and LUN operations in the `clone` log file. Every Sunday at 12:00 a.m., the `clone` log file at `/etc/log/clone` is renamed. A suffix 0 to 5 is added to the `clone` log file name. First Sunday at 12:00 a.m., the clone log file is renamed as `clone.0`; and next Sunday at 12:00 a.m. it is renamed as `clone.1` and so on up to `clone.5`. The oldest `clone.5` log file is deleted at the end of seventh week.

Sample of clone log file

```
 Sun Jun 21 00:12:17 GMT 2009 Volume: mam Clone Start ID: 1095,
Clone File: f3,  Clone File ID: 4729, Clone File Generation Count
429265099,
Source File: f3, Source  File ID: 4729, Source File Generation Count:
429265099,
Total Blocks: 135, Entry Index: 0, Snap Index: -1, Snap ID: 0, Snap
CP Count : 0,
Change Log: true, Block Ranges : 0:50:30:0:122880 20:70:25:0:102400
100:0:80:0:327680 Jun 21 00:12:17 GMT 2009 Volume: mam Clone End ID:
1095,
Clone File: f3, Source File: f3 (Operation succeeded), Total Blocks:
135,
Blocks Copied: 0
```

Related concepts

About FlexClone files and FlexClone LUNs on page 219
FlexClone file and FlexClone LUN operations on page 241

Rapid Cloning Utility for VMware

Rapid Cloning Utility helps you to quickly create multiple clones of virtual machines in the VMware environment.

Using FlexClone technology, the Rapid Cloning Utility (RCU) allows users to quickly create and deploy VMware virtual machines (VMs) across new or existing NFS-based datastores. You can efficiently create virtual machine clones in VMware Virtual Center, power up virtual machines, and apply customized specifications to the guest operating system. The utility can deploy virtual machines for both server and desktop use.

The Rapid Cloning Utility can theoretically create up to 8,000 virtual machine clones and 32 datastores in a single execution. In practice, however, multiple executions of smaller requests is recommended. The exact size of these requests will depend on the size of the Virtual Infrastructure 3 or VMware VSphere deployment and the hardware configuration of the vCenter Server managing the ESX hosts.

For more information about the Rapid Cloning Utility, see the IBM NAS support site.

Related concepts

About FlexClone files and FlexClone LUNs on page 219

Related information

www.ibm.com/storage/support/nas/

FlexClone file and FlexClone LUN interoperability with Data ONTAP features

FlexClone file and FlexClone LUN work with most but not all of the Data ONTAP features.

Next topics

How Snapshot copies work with FlexClone files and FlexClone LUNs

You can perform all Snapshot copy operations on a FlexVol volume that contains FlexClone files or FlexClone LUNs.

The following are important points that you should know:

- If a Snapshot copy is created when the cloning operation is in progress, the partially cloned file or LUN is locked within the Snapshot copy. However, the FlexClone file or LUN is created successfully at the end of the cloning operation.
- The partially cloned file that is locked in the Snapshot copy has all its permissions set to zero. Therefore, when you restore a volume from the Snapshot copy, the partially cloned files are also

restored, which are of no use. You can identify a partially cloned file by the zero permission set on it when the volume is mounted using NFS.

For more information about Snapshot copies, see the *Data ONTAP Data Protection Online Backup and Recovery Guide*.

How volume SnapMirror works with FlexClone files and FlexClone LUNs

You should know a few important points when using volume SnapMirror with a FlexVol volume that has FlexClone files and LUNs.

If a FlexVol volume is a SnapMirror source and contains FlexClone files or FlexClone LUNs, volume SnapMirror transfers only the physical block and a small amount of metadata. On the destination only one copy of the physical block is stored, and the block is shared among the source and its clones. Therefore, the destination volume is an exact copy of the source volume and all the clone files or LUNs on the destination volume share the same physical blocks.

Volume SnapMirror locks all the volume Snapshot copies during the transfer. Volume SnapMirror can also lock temporary Snapshot copies created for cloning purposes. If volume SnapMirror transfer starts while a cloning operation is in progress, then the Snapshot copy taken is not deleted at the end of the cloning operation if SnapMirror transfer is still in progress and you must wait until the volume SnapMirror is complete before starting a new cloning operation.

You can suppress the creation of a Snapshot copy when cloning by using the `-n` option of the `clone start` command.

When using volume SnapMirror with FlexClone files and LUNs, you should take precautions in the following cases:

- There is a volume with FlexClone files and LUNs already on it, and you want to replicate this volume using SnapMirror.

  For the SnapMirror destination system, ensure that the size of the volume is within the size limit for volumes with FlexClone files or LUNs.
- There is a volume that is already a source for a volume SnapMirror relationship, and you want to create FlexClone files or LUNs inside such a volume.

  For both the SnapMirror source and destination systems, ensure that the size of the volume is within the size limit for volumes with FlexClone files or LUNs.

For more information about volume SnapMirror, see the *Data ONTAP Data Protection Online Backup and Recovery Guide*.

Related concepts

About FlexClone files and FlexClone LUNs on page 219
Operational limits for FlexClone files and FlexClone LUNs on page 225

How synchronous SnapMirror works with FlexClone files and FlexClone LUNs

You should not use a FlexVol volume that has FlexClone files and FlexClone LUNs as a source for synchronous SnapMirror.

Synchronous SnapMirror is not qualified on a FlexVol volume with FlexClone files or FlexClone LUNs.

For more information about synchronous SnapMirror, see the *Data ONTAP Data Protection Online Backup and Recovery Guide*.

How qtree SnapMirror and SnapVault work with FlexClone files and FlexClone LUNs

Qtree SnapMirror and SnapVault are not aware that FlexClone files and FlexClone LUNs are logical files that share physical blocks with their parents. Therefore, they mirror all the FlexClone files and LUNs to the destination as individual physical files and LUNs.

The destination FlexVol volume must have enough capacity to store the FlexClone files or LUNs, as separate files or LUNs.

Running deduplication on the destination volume after the qtree SnapMirror or SnapVault transfer is complete reduces the amount of used space on the destination FlexVol volume.

For more information about qtree SnapMirror and SnapVault, see the *Data ONTAP Data Protection Online Backup and Recovery Guide*.

How deduplication works with FlexClone files and FlexClone LUNs

You can create a FlexClone file or FlexClone LUN on a FlexVol volume with deduplication enabled.

The block-sharing mechanism used by FlexClone files and LUNs is also used by deduplication. Therefore, if deduplication was enabled or is currently enabled on a FlexVol volume, you might end up creating a new physical copy of the parent entity even before reaching maximum shared limit for FlexClone files and LUNs.

The `-l` option of the `clone start` command enables change logging. The change log information is used by deduplication. Enabling change logging ensures that there is an appropriate entry for both the parent and clone files in the deduplication metadata. When the data in the parent is overwritten, the newer data in written to a different block on the disk and the old data block is referenced only by the clone file and is no longer shared. If the cloning operation was performed with the `-l` option, and deduplication is run on the volume, the older block, which is now referenced only by the clone, can also be shared with other logical blocks in any other files across the volume that has the same data.

Note: FlexClone file and FlexClone LUN operations cannot be performed on a FlexVol volume that has a `sis undo` operation currently running on the volume.

Related concepts

Considerations when planning FlexClone files or FlexClone LUNs on page 223
Operational limits for FlexClone files and FlexClone LUNs on page 225

Related tasks

Creating a FlexClone file or FlexClone LUN on page 242

How quotas work with FlexClone files and FlexClone LUNs

Quota limits are applied on the total logical size of the FlexClone files or FlexClone LUNs. When you create a FlexClone file or FlexClone LUN, quotas do not recognize any space savings. For example, if you create a FlexClone file of a parent file of 10 GB, you are only using 10 GB of physical space, but the quota utilization is recorded as 20 GB (10 GB for the parent and 10 GB for the FlexClone file).

The effects of exceeding quota limits are different for qtree quota and user or group quota. If the FlexClone files or LUNs are part of a UNIX or mixed security style qtree, the quota of the user or group owning the parent file or LUN applies. If the FlexClone files or LUNs are part of an NTFS qtree, the root user quota applies.

If the creation of a FlexClone file or LUN would result in the qtree quota's being exceeded, the FlexClone operation fails.

If the creation of a FlexClone file or LUN would result in the group or user quota's being exceeded, the clone operation succeeds, provided the FlexVol volume has enough space to hold the metadata for the clone. However, the quota for that user or group is oversubscribed.

How space reservation works with FlexClone files and FlexClone LUNs

A FlexClone file does not inherit the space reservation attribute from the parent file. A FlexClone LUN inherits the space reservation setting of the parent LUN.

To enable space reservation on the FlexClone file, you can use the `file reservation` command.

FlexClone LUNs inherit the space reservation settings of the parent LUN. Therefore, if there is not enough space in the FlexVol volume to create a FlexClone LUN with the same space reservation as that of the parent, then the cloning operation fails.

Note: The space required according to space reservation attribute is separate for parent LUN and FlexClone LUN.

How MultiStore works with FlexClone files and FlexClone LUNs

Starting with Data ONTAP 7.3.3, FlexClone files and LUN commands are available in the default and nondefault vfiler contexts. You can use the FlexClone files and LUNs feature to create writable, space-efficient clones of parent files and parent LUNs within a vFiler unit.

The following are considerations for creating FlexClone files and LUNs on vFiler units:

- Both MultiStore and FlexClone licenses must be enabled on the storage system.
- A vFiler unit unit administrator can perform all FlexClone file and LUN operations only on vFiler units that you are authorized to manage.
- A storage system administrator can perform FlexClone file operations on storage resources owned by all vFiler units from the default vfiler context.
- A storage system administrator cannot run FlexClone LUN operations from the default vfiler context on a LUN owned by a nondefault vFiler unit.
- If you are running a FlexClone LUN operation in the default vfiler context and if the volume or qtree on which the FlexClone LUN operation is running is moved to a nondefault vFiler unit, then the FlexClone LUN operation fails.
- A storage system administrator can see all clone operations running on different vFiler units from the default vfiler context.
- FlexClone file and LUN operations are visible only from the vfiler context on which the operations are being run and from the default vfiler context.

  You cannot view clone operations being run on other vFiler units.
- Storage owned by a vFiler unit cannot be accessed or discovered from other vFiler units by using the FlexClone file or LUN commands.
- During reboot or takeover, if the storage is moved between vFiler units, the clone operation fails. However, this does not happen if the file clone operation was started from the default vfiler context.
- If a storage system reboots, then all the clone operations are restarted on the same vFiler unit after reboot.
- You can run a maximum of 500 FlexClone file and LUN operations on a storage system.
- All FlexClone file and LUNs commands are supported on vFiler units.

  The FlexClone file or LUN operations do not interrupt offline vfiler migration and disaster recovery operations.
- The logs of all clone operations performed in vFiler units are stored at `/vol/vol0/etc/log/clone`.
- You can run the following FlexClone file and LUN commands using CLI or Data ONTAP APIs in a vfiler context:

  Note: You must switch to the vfiler context of the vFiler unit that owns the FlexVol volume or qtree.

  - `clone start`
  - `clone status`
  - `clone stop`
  - `clone clear`

  Note: During online migration of a vFiler unit, you cannot use the `clone start` command on volumes that are owned by that vFiler unit.

For more information about MultiStore, see the *Data ONTAP MultiStore Management Guide*.

How volume move affects FlexClone files and FlexClone LUNs

You cannot run FlexClone files and FlexClone LUNs operations during the cutover phase of a volume move operation.

The following FlexClone files and LUNs commands are not allowed:

- `clone start`
- `clone status`
- `clone clear`
- `clone stop`

If you run any of these commands, the system generates one of these error messages `"Volume state transition is in progress"` or `"Volume does not exist"`.

If the `clone start` operation is in progress and the volume move operation enters the cutover phase, the volume move operation is paused.

For more information about volume move, see the *Data ONTAP Block Access Management Guide for iSCSI and FC*.

How NDMP and dump works with FlexClone files and FlexClone LUNs

NDMP and `dump` work at the logical level with FlexClone files and FlexClone LUNs. All FlexClone files or LUNs are backed up as separate files or LUNs.

When you use NDMP services to back up a qtree or FlexVol volume that contains FlexClone files or FlexClone LUNs, block sharing between parent entities and clone entities is disabled and clone entities are backed up to tape as separate files or LUNs. The space saving is lost. Therefore, the tape onto which you are backing up should have sufficient space to store the expanded amount of data.

When you restore, all files and LUNs are restored as separate physical files and LUNs.

If a `dump` backup is triggered while the cloning operation is in progress, the `dump` Snapshot copy contains a partially cloned file. The Snapshot copy with the partially cloned file is backed up. The `dump` backup is capable of managing the partially cloned file.

For more information about tape backup, see the *Data ONTAP Data Protection Tape Backup and Recovery Guide*.

How single file SnapRestore works with FlexClone files and FlexClone LUNs

You cannot run FlexClone file or FlexClone LUN and single file SnapRestore operations simultaneously on a FlexVol volume.

For more information about single file SnapRestore, see the *Data ONTAP Data Protection Online Backup and Recovery Guide*.

How file folding works with FlexClone files and FlexClone LUNs

File folding and FlexClone file or FlexClone LUN operations cannot run in parallel on the same FlexVol volume.

For more information about file folding, see the *Data ONTAP Data Protection Online Backup and Recovery Guide*.

How volume SnapRestore works with FlexClone files and FlexClone LUNs

You cannot run FlexClone file or FlexClone LUN and volume SnapRestore operations simultaneously on a FlexVol volume.

For more information about volume SnapRestore, see the *Data ONTAP Data Protection Online Backup and Recovery Guide*.

How volume autosize works with FlexClone files and FlexClone LUNs

The maximum volume autosize option setting should be less than the maximum recommended volume size for FlexClone files and LUNs. The maximum size depends on the hardware platform on which you are running the cloning operation.

When you enable `vol autosize` on a FlexVol volume that contains a FlexClone file or LUN, the maximum autosize setting must be less than the maximum recommended volume size for FlexClone files and LUNs for that hardware platform. If the maximum autosize setting is higher, then volume autosize is not enabled and an error message is displayed.

If you run the cloning operation on a FlexVol volume with `vol autosize` enabled, if the FlexVol volume runs out of space while creating the metadata required for the cloning operation, then the autosize operation is not activated and the cloning operation fails.

Related concepts

Operational limits for FlexClone files and FlexClone LUNs on page 225

How volume-copy works with FlexClone files and FlexClone LUNs

You can perform a volume-copy operation on a FlexVol volume that has FlexClone files and FlexClone LUNs in it.

After the volume-copy operation is done, the FlexClone files and FlexClone LUNs and their parents on the destination FlexVol volume share the same data blocks as they did on the source FlexVol volume.

The destination FlexVol volume has the attribute `sis` attached to it, which shows up in the output of the `vol status` command.

Note: If you run both `vol copy` transfer and cloning operations simultaneously on a FlexVol volume and if the cloning operation ends before the `vol copy` transfer, then the temporary Snapshot copy created for cloning purpose is not deleted after the cloning operation is complete.

You must wait until the volume-copy operation is complete before starting a new cloning operation using a temporary Snapshot copy.

For more information about volume-copy, see the *Data ONTAP Data Protection Online Backup and Recovery Guide*.

How FlexClone files and FlexClone LUNs work when the system reboots

If a FlexClone file or FlexClone LUN operation is in progress and the system reboots, then the FlexClone operation restarts automatically after reboot.

How an active/active configuration works with FlexClone files and FlexClone LUNs

FlexClone file and FlexClone LUN operations are supported in an active/active configuration.

If takeover occurs when a FlexClone file or FlexClone LUN operation is in progress, then the running clone operation is terminated and automatically restarted after the takeover operation is complete.

Similarly, if giveback starts when a FlexClone file or FlexClone LUN operation belonging to the partner node is in progress, then the running clone operation of the partner node is terminated and automatically restarted after the giveback operation is complete.

How role-based access control lists work with FlexClone files and FlexClone LUNs

You can use Data ONTAP role-based access capabilities for managing FlexClone file and FlexClone LUN operations.

You can create roles that have access only to the commands that are needed to perform the FlexClone file and LUN operations.

You can also restrict access to the FlexClone operations by using Data ONTAP role-based access control capabilities.

For more information about role-based access control list, see the *Data ONTAP System Administration Guide*.

How access control lists and streams work with FlexClone files and FlexClone LUNs

FlexClone files do not inherit the access control lists or streams of their parent files. FlexClone LUNs do inherit the access control list or streams of their parent LUNs.

If you want the FlexClone files to have the same ACL (access control list) as their parents, or if you want to attach the streams to the FlexClone files, then you must set ACLs individually on the FlexClone file after completing the cloning operation.

For more information about access control lists, see the *Data ONTAP File Access and Protocols Management Guide*.

How FlexShare works with FlexClone files and FlexClone LUNs

You can set a priority for FlexClone file and FlexClone LUN operations using FlexShare.

FlexShare treats the workload generated by FlexClone files or LUNs as system workload. You can use FlexShare to set a priority for the workload generated by the cloning operation. The impact on the storage system can be adjusted according to the priority set for system operations in FlexShare.

For more information about FlexShare, see the *Data ONTAP System Administration Guide*.

How volume clone works with FlexClone files and FlexClone LUNs

You can create a FlexClone volume of a FlexVol volume that has both a FlexClone file and FlexClone LUN and its parent file or LUN in it.

The FlexClone files or FlexClone LUNs and their parent files or LUNs that are present in the FlexClone volume continue to share blocks the same way they do in the parent FlexVol volume. In fact, all the FlexClone entities and their parents share the same underlying physical data blocks, minimizing physical disk space usage.

If the FlexClone volume is split from its parent volume, then the FlexClone files or FlexClone LUNs and their parent files or LUNs stop sharing the blocks in the child FlexClone volume. Thereafter they exist as independent files or LUNs. This means that the child volume uses more space than it did before the split operation.

FlexClone file and FlexClone LUN operations

You can start and stop a clone operation, view the status of a clone operation, and clear the status of a failed clone operation.

- You can create a FlexClone file or FlexClone LUN using the `clone start` command.

  Note: The maximum number of FlexClone file or FlexClone LUN operations that can run simultaneously on a volume is 16.
- You can view the status of all running and failed FlexClone file or FlexClone LUN operations using the `clone status` command. The command shows the status of all running and failed FlexClone file or FlexClone LUN operations. Each operation has a unique clone operation ID within a FlexVol volume.
- You can stop a running FlexClone file or FlexClone LUN operation using the `clone stop` command. To run this command you should know the unique ID of the clone operation.
- You can clear the status of a failed FlexClone file or FlexClone LUN clone operation using the `clone clear` command.

Next topics

Creating a FlexClone file or FlexClone LUN on page 242
Viewing the status of a FlexClone file or FlexClone LUN operation on page 244
Stopping a FlexClone file or FlexClone LUN operation on page 245
Clearing the status of a failed FlexClone file or FlexClone LUN operation on page 246
Viewing the space savings due to FlexClone files and FlexClone LUNs on page 246
Viewing the file space utilization report on page 247
Considerations when creating FlexClone files or FlexClone LUNs on page 248

Related concepts

Uses of FlexClone files and FlexClone LUNs on page 223
Considerations when planning FlexClone files or FlexClone LUNs on page 223
When a FlexClone file or LUN is moved or renamed during cloning operation on page 249
FlexClone file and FlexClone LUN interoperability with Data ONTAP features on page 231

Creating a FlexClone file or FlexClone LUN

You can create a FlexClone file or a FlexClone LUN of a parent file or LUN inside a FlexVol volume using the `clone start` command. You can also use this command to clone a sub-file or sub-LUN.

Before you begin

- You must install a FlexClone license on your storage system to create FlexClone files or FlexClone LUNs.
- In an active/active configuration, you must install the FlexClone license on both systems.
- To clone a sub-file or sub-LUN, you should know the block range of the parent entity and clone entity.

About this task

The FlexClone file or LUN must be in the same FlexVol volume as the parent.

Step

1. To create a FlexClone file or FlexClone LUN or to clone a sub-file or sub-LUN, choose one of actions from the following table.

If you want to create...	Then...
A FlexClone file or FlexClone LUN of a parent file or LUN inside a FlexVol volume.	Enter the following command: **clone start *src_path dest_path* [-n] [-l]** • *src_path*—Source path in the /vol/*volname*/... format • *dest_path*—Destination path in the /vol/*volname*/... format • -n— This option prevents creation of a temporary Snapshot copy of a FlexVol volume during the cloning operation. You should use the *-n* option only when you are certain that no modifications will happen to the parent file or LUN during the cloning operation. • -l—This option enables change logging for clone blocks. • You can use the -l option only on a deduplication enabled FlexVol volume. **Note:** • When you run the clone start command with -l option, the cloning operation succeeds if the aggregate is full but the volume has space. However, change logging stops and an EMS message is displayed. • When you run the clone start command with -l option and the sis is turned off, the cloning operation succeeds. However, further change logging stops.
A sub-file or sub-LUN clone.	Enter the following command: **clone start *src_path* [*dest_path*] [-n] [-l] -r *src_fbn*:*dest_fbn*:*fbn_cnt* ...** • *src_path*—Source path in the /vol/*volname*/... format • *dest_path*—Destination path in the /vol/*volname*/... format • -n— This option prevents creation of a temporary Snapshot copy of a FlexVol volume during the clone operation. You should use the *-n* option only when you are certain that no modifications will happen to the parent file or LUN during the cloning operation. • -l—This option enables change logging for clone blocks. You can use the -l option only on a deduplication enabled FlexVol volume. • -r—Specifies block ranges. • *src_fbn*—Starting fbn of the source block range. For a LUN, the fbn is considered as LBA (Logical block address). • *dest_fbn*—Starting fbn of the destination block range. The fbn is the destination address where the blocks will be referenced. • *fbn_cnt*—Number of blocks to be cloned.

Example

The following command creates a clone of `testfile` on the `toaster` storage system.

```
toaster> clone start/vol/testvol/testfile /vol/testvol/clonetestfile
Clone operation started successfully. ID: 10.
toaster> Fri May 29 14:09:14 IST [wafl.snap.delete:info]: Snapshot
copy dense_clone.0.ce7807da-4692- 11de-9242-00a098076602 on volume
testvol was deleted by the Data ONTAP function
dense_clone_delete_snapshot. The unique ID for this
Snapshot copy is (56, 53575). Fri May 29 14:09:14 IST
[dense.clone.finish:info]: Clone operation on file
'/vol/testvol/clonetestfile'  completed successfully.
The clone operation ID was 10
```

Related concepts

Considerations when planning FlexClone files or FlexClone LUNs on page 223
What happens when FlexClone file or LUN operation fails on page 248
What the FlexClone log file is on page 229
Considerations when creating FlexClone files or FlexClone LUNs on page 248
How deduplication works with FlexClone files and FlexClone LUNs on page 233
Operational limits for FlexClone files and FlexClone LUNs on page 225

Viewing the status of a FlexClone file or FlexClone LUN operation

You can view the status of all FlexClone file or FlexClone LUN operations currently running, the FlexClone operations that failed and the reason for the failure using the `clone status` command. You can also use this command to view the status of a stopped cloning operation if the stop operation is in progress. The command does not display information about successfully completed or successfully stopped cloning operations.

Step

1. To view the status of a FlexClone file or FlexClone LUN operation, enter the following command:

   `clone status [`*`vol-name [ID]`*`]`

   - *`vol-name`*—Volume name. If *`vol-name`* is not specified, the command displays the status of all clone operations on the storage system.
   - *`ID`*—Clone operation ID. If the *`ID`* is not specified, the command displays the status of all clone operations on the volume.

- If both *vol-name* and *ID* are specified, the command displays the status of the specific clone operation.

Example

You can view the status of the FlexClone operation of the `test_file` on the`toaster` storage system using the following command.

```
toaster > clone status testvol 538
ID: 538
Source: /vol/testvol/test_file
Destination: /vol/testvol/clone_test_file
Block ranges:
State: running (49% done)
Total blocks: 2621441
Blocks copied: 0
Type: file
```

Stopping a FlexClone file or FlexClone LUN operation

You can stop a FlexClone file or FlexClone LUN operation by using the `clone stop` command. The stop operation might take some time to complete. Stopping the clone operation deletes any temporary Snapshot copy created. The `clone status` command does not show any status after the cloning operation is stopped.

Before you begin

You need to know the ID of the FlexClone operation you want to stop. You can learn the ID by using the `clone status` command.

Step

1. To stop a FlexClone file or FlexClone LUN operation, enter the following command:

   clone stop *vol-name ID*

   - *vol_name*—Volume name
   - *ID*—Clone operation ID

Example

You can stop a FlexClone file operation on the `toaster` storage system using the following command.

```
toaster > clone stop testvol1 508
```

Clearing the status of a failed FlexClone file or FlexClone LUN operation

You can clear the status of a failed FlexClone file or FlexClone LUN operation by using the `clone clear` command.

Before you begin

You should know the ID of the failed FlexClone operation. You can find the ID by using the `clone status` command.

Step

1. To clear the status of a failed FlexClone file or FlexClone LUN operation, enter the following command:

   clone clear *vol-name ID*

   - *vol-name*—Volume name
   - *ID*—Clone operation ID

   Note: Status update is not displayed for successful FlexClone file or FlexClone LUN operations.

Example

You can clear the status of a failed FlexClone file operation on the `toaster` storage system using the following command.

```
toaster > clone clear testvol 804
```

Viewing the space savings due to FlexClone files and FlexClone LUNs

You can view the space saved by FlexClone files and LUNs using the `df-s` command.

Step

1. To view the space saving due to FlexClone files and LUNs, enter the following command:

   df -s *volname*

volname is the name of the FlexVol volume. For example, *test1*.

For more information about the `df` command, see the df(1) man page.

Example

The following example shows the space saving on the `test1` FlexVol volume.

```
toaster> df -s test1

Filesystem          used   saved   %saved
/vol/test1/         4828    5744     54%
```

Related concepts

Space savings achieved by using FlexClone files and FlexClone LUNs on page 228

Viewing the file space utilization report

You can view the file space utilization report by using the `du` command. This command enables you to determine the minimum number of blocks, excluding those that are trapped in Snapshot copies, that can be freed when a deduplicated or cloned file is deleted.

Step

1. Enter the following command to view the file space utilization report:

   `du [-u][-k][-m][-h][-r {`*`start-range:end-range`*` |`*`file_path`*`}]`

   The `-u` option displays the unique blocks in the file.

   The `-k` option displays the output in KB.

   The `-m` option displays the output in MB.

   The `-h` option displays the output in the appropriate unit of measurement. It scales the file size and displays the output appropriately in KB, MB, or GB.

   The `-r` option displays the number of total and unique blocks present in the specified range.

Examples

The following command displays the number of blocks in the file:

```
SystemA> du /vol/vol1/file_2t

382 /vol/vol1/file_2t
```

The following command displays the unique blocks in the file:

```
SystemA> du -u /vol/vol1/file_3t
382 127 /vol/vol1/file_3t
```

The following command displays the output in the appropriate unit of measurement:

```
SystemA> du -u -h /vol/vol1/file_4t
2101304KB 4120KB /vol/vol1/file_4t
```

The following command displays the output in MB:

```
SystemA> du -u -m /vol/vol1/file_5t
2052 4 /vol/vol1/file_5t
```

The following command displays the output in KB:

```
SystemA> du -u -k /vol/vol1/file_6t
2 1 /vol/vol1/file_5t
```

The following command displays the output in KB:

```
SystemA> du -r 1:32 /vol/vol1/file_7t
4 /vol/vol1/file_7t
```

Considerations when creating FlexClone files or FlexClone LUNs

You should know what happens when the cloning operation fails or when a FlexClone file or LUN is moved or renamed during the cloning operation.

Next topics

What happens when FlexClone file or LUN operation fails on page 248
When a FlexClone file or LUN is moved or renamed during cloning operation on page 249

What happens when FlexClone file or LUN operation fails

When cloning operations fail, messages and log entries are generated.

If the FlexClone file or FlexClone LUN cloning operation fails in the middle, all the changes made up to that point are reverted. The partially created FlexClone files or FlexClone LUNs and the temporary Snapshot copy are deleted.

If you try to clone a sub-file or sub-LUN and the operation fails, then the partially cloned file or LUN is not deleted.

In either of the preceding cases, an error message is displayed on the storage system console about the failed cloning operation. Also, the failed information is logged in the EMS log file and clone log file of the `/etc/log` directory.

Note: The `clone status` command displays the failed cloning operations. The status of failed cloning operation is displayed only if the failure status metadata is stored on the disk. If the cloning operation metadata is cleared using the `clone clear` command, then the status is not displayed.

Related concepts

What the FlexClone log file is on page 229

Related tasks

Creating a FlexClone file or FlexClone LUN on page 242
Viewing the status of a FlexClone file or FlexClone LUN operation on page 244
Clearing the status of a failed FlexClone file or FlexClone LUN operation on page 246

When a FlexClone file or LUN is moved or renamed during cloning operation

Renaming FlexClone files or FlexClone LUNs or moving them to another directory during the cloning operation does not impact the operation. However, if the cloning operation fails or is stopped, then the partially cloned files and LUNs are not deleted.

You must manually delete the partially cloned files or LUNs from the location. Also the `clone status` command shows the old path from where the cloning operation was started.

Related concepts

About FlexClone files and FlexClone LUNs on page 219

Related tasks

Viewing the status of a FlexClone file or FlexClone LUN operation on page 244

Space savings with deduplication

Deduplication is an optional feature of Data ONTAP that significantly improves physical storage space by eliminating duplicate data blocks within a FlexVol volume.

Deduplication works at the block level on the active file system, and uses the WAFL block-sharing mechanism. Each block of data has a digital signature that is compared with all other signatures in a data volume. If an exact block match exists, the duplicate block is discarded and its disk space is reclaimed.

You can configure deduplication operations to run automatically or on a schedule. You can deduplicate new and existing data, or only new data, on a FlexVol volume.

Deduplication removes data redundancies, as shown in the following illustration:

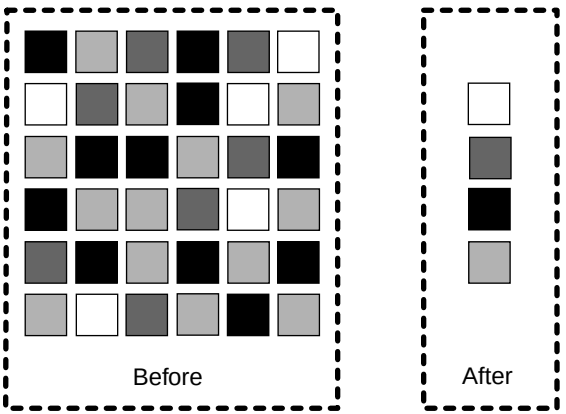


Figure 1: How deduplication removes data redundancies

Next topics

How deduplication works on page 252
What deduplication metadata is on page 252
Activating the deduplication license on page 253
Guidelines for using deduplication on page 253
Deduplication schedules on page 256
How deduplication works with other features and products on page 264
Common troubleshooting procedures for volumes with deduplication on page 275

Related tasks

Activating the deduplication license on page 253

How deduplication works

Deduplication operates at the block level within the entire FlexVol volume, eliminating duplicate data blocks and storing only unique data blocks.

Data ONTAP writes all data to a storage system in 4-KB blocks. When deduplication runs for the first time on a FlexVol volume with existing data, it scans all the blocks in the FlexVol volume and creates a digital fingerprint for each of the blocks. Each of the fingerprints is compared to all other fingerprints within the FlexVol volume. If two fingerprints are found to be identical, a byte-for-byte comparison is done for all data within the block. If the byte-for-byte comparison detects identical fingerprints, the pointer to the data block is updated, and the duplicate block is freed.

Deduplication runs on the active file system. Therefore, as additional data is written to the deduplicated volume, fingerprints are created for each new block and written to a change log file. For subsequent deduplication operations, the change log is sorted and merged with the fingerprint file, and the deduplication operation continues with fingerprint comparisons as previously described.

What deduplication metadata is

Deduplication uses fingerprints, which are digital signatures for every 4-KB data block in a FlexVol volume. The fingerprint database and the change logs form the deduplication metadata.

In Data ONTAP 7.3 and later, the fingerprint database and the change logs used by the deduplication operation are located outside the volume and in the aggregate. Therefore, the deduplication metadata is not included in the FlexVol volume Snapshot copies.

This approach enables deduplication to achieve higher space savings than in Data ONTAP 7.2. However, some of the temporary metadata files created during the deduplication operation are still placed inside the volume and are deleted only after the deduplication operation is complete. The temporary metadata files, which are created during a deduplication operation, can be locked in the Snapshot copies. These temporary metadata files remain locked until the Snapshot copies are deleted.

While deduplication can provide substantial space savings, a percentage of storage overhead is associated with it, which you should consider when sizing a FlexVol volume.

The deduplication metadata can occupy up to 6 percent of the total logical data of the volume, as follows:

- Up to 2 percent of the total logical data of the volume is placed inside the volume.
- Up to 4 percent of the total logical data of the volume is placed in the aggregate.

Related concepts

Deduplication and Snapshot copies on page 264

Activating the deduplication license

You can activate the deduplication license using the `license add` command after installing Data ONTAP.

Before you begin

Before installing the deduplication license, the near-line functionality license must be installed on the system.

About this task

The deduplication license is supported on systems that support the near-line functionality license.

Step

1. Enter the following command:

   `license add` ***`license_key`***

   `license_key` is the code for the deduplication license.

   For more information about the `license` command, see the na_license(1) man page.

   Note: The deduplication license is only supported with Data ONTAP 7.2.2 or later releases (for the IBM N3300, IBM N3400, and IBM N3600 systems, Data ONTAP 7.2.2 L1 or later releases).

Related references

Common troubleshooting procedures for volumes with deduplication on page 275

Guidelines for using deduplication

You must remember certain guidelines about system resources and free space when using deduplication.

The guidelines are as follows:

- Deduplication is a background process that consumes system resources while it is running. If the data does not change very often in a FlexVol volume, it is best to run deduplication less frequently. Multiple concurrent deduplication operations running on a storage system lead to a higher consumption of system resources.
- You must ensure that sufficient free space exists for deduplication metadata in the volumes and aggregates.

- For releases earlier than Data ONTAP 8.1, you cannot increase the size of a volume that contains deduplicated data beyond the maximum supported size limit, either manually or by using the `autogrow` option.
- For releases earlier than Data ONTAP 8.1, you cannot enable deduplication on a volume if it is larger than the maximum volume size. However, starting with Data ONTAP 7.3.1, you can enable deduplication on a volume after reducing its size within the supported size limits.

Next topics

Maximum volume size with deduplication on page 254
Performance considerations for deduplication on page 255
Deduplication and read reallocation on page 256
Deduplication and extents on page 256

Related concepts

Deduplication and volume SnapMirror on page 265
Deduplication must be disabled before removing the deduplication license on page 263
Deduplication and DataFabric Manager on page 269

Related tasks

Default schedule for deduplication on page 257

Maximum volume size with deduplication

There are limits on the volume size and the amount of data in a volume with deduplication.

The following table lists the maximum volume sizes supported for different systems, beginning with Data ONTAP 7.3.1, with and without deduplication:

Table 4: Maximum volume size supported for different systems, with and without deduplication

Model	Maximum size of volume without deduplication (TB)	Maximum size of volume with deduplication (TB)*	Total data size of volume with deduplication (TB)*
N3300	16	1	17
N3600	16	2	18
N5200	16	2	18
N5300	16	4	20
N5500	16	3	19
N5600	16	16	32

Model	Maximum size of volume without deduplication (TB)	Maximum size of volume with deduplication (TB)*	Total data size of volume with deduplication (TB)*
N6040	16	4	20
N6060	16	16	32
N6070	16	16	32
N7600	16	16	32
N7700	16	16	32
N7800	16	16	32
N7900	16	16	32

* This information is not applicable in the case of reservation-enabled LUNs.

Note:

- All systems listed in the preceding table must have the near-line functionality license enabled to use deduplication.
- A volume that exceeds the maximum supported size will not go offline. However, deduplication will be disabled on that volume.

Performance considerations for deduplication

Certain factors affect the performance of deduplication. You should check the performance impact of deduplication in a test setup, including sizing considerations, before deploying deduplication in performance-sensitive or production environments.

The following factors affect the performance of deduplication:

- Application and the type of data used
- The data access pattern (for example, sequential versus random access, the size and pattern of the input and output)
- The amount of duplicate data, the amount of total data, and the average file size
- The nature of data layout in the volume
- The amount of changed data between deduplication operations
- The number of concurrent deduplication operations
- Hardware platform (system memory and CPU module)
- Load on the system (for example, MBps)
- Disk types (for example, ATA/FC, and RPM of the disk)

Deduplication and read reallocation

Because read reallocation does not predictably improve the file layout and the sequential read performance when used on deduplicated volumes, you should not perform read reallocation on deduplicated volumes.

Read reallocation might conflict with deduplication by adding new blocks that were previously consolidated during the deduplication process. A deduplication scan might also consolidate blocks that were previously rearranged by the read allocation process, thus separating chains of blocks that were sequentially laid out on disk.

For more information about read reallocation, see the *Data ONTAP System Administration Guide*.

Related concepts

Improved sequential read performance for deduplicated FlexVol volumes on page 272

Deduplication and extents

Because enabling extents does not predictably optimize sequential data block layout when used on deduplicated volumes, you should not enable extents on deduplicated volumes.

Extents might conflict with deduplication by adding new blocks that were previously consolidated during the deduplication process. A deduplication scan might also consolidate blocks that were previously rearranged by extents, thus separating chains of blocks that were sequentially laid out on disk.

For more information about enabling extents, see the *Data ONTAP System Administration Guide*.

Related concepts

Improved sequential read performance for deduplicated FlexVol volumes on page 272

Deduplication schedules

You can run deduplication on a volume using the command-line interface at any point in time. You can also create a schedule to run deduplication at specified times.

If deduplication operations are enabled for a FlexVol volume, they run in the following situations:

- The default schedule (at midnight every day)
- According to a schedule you create, for specific days and at specific times
- Manually through the command-line interface
- Automatically, when 20 percent new or changed data has been written to the volume

Next topics

Default schedule for deduplication on page 257

Default schedule for deduplication

Deduplication operations run on enabled FlexVol volumes once a day at midnight by default. When deduplication is enabled for the first time on a FlexVol volume, a default schedule is configured. This default schedule runs deduplication every day at midnight.

Creating a deduplication schedule

Deduplication operations run on enabled FlexVol volumes once a day at midnight by default. If you wish to run deduplication at another time, you can create a deduplication schedule using the `sis config -s` command.

Step

1. Enter the following command:

   `sis config -s schedule path`

   `schedule` lists the days and hours of the day when deduplication runs. The schedule can be of the following types:

   - *`day_list`*[`@`*`hour_list`*]
     If *`hour_list`* is not specified, deduplication runs at midnight on each scheduled day.
   - *`hour_list`*[`@`*`day_list`*]
     If *`day_list`* is not specified, deduplication runs every day at the specified hours.
   - `-`
     A hyphen (-) disables deduplication operations for the specified FlexVol volume.

   `path` is the complete path to the FlexVol volume—for example, `/vol/vol1`.

   Example
   The following command starts deduplication operations at 11 p.m., Monday through Friday.

   ```
   systemA> sis config -s mon-fri@23 /vol/vol1
   ```

   For more information about scheduling deduplication operations, see the na_sis(1) man page.

Running deduplication manually on existing data

You can manually scan and eliminate duplicate blocks on an existing FlexVol volume using the `sis start` command.

Steps

1. To start deduplication operations, enter the following command:

   `sis start -s` ***`path`***

   `path` is the complete path to the FlexVol volume. For example, `/vol/vol1`.

   For more information, see the sis(1) man pages.

   If deduplication operations are already running on the volume when you run the `sis start -s` command (for example, if a scheduled deduplication operation has begun), the command fails. To eliminate duplicate blocks that existed before the previous operation, run the preceding command after the previous deduplication operation is complete.

   Note: You should disable deduplication schedules before running the `sis start -s` command on a large volume.

   ### Example

   ```
   systemA> sis start -s /vol/vol1
   ```

2. To start deduplication outside the preset schedule, enter the following command:

   `sis start` ***`path`***

   `path` is the complete path to the FlexVol volume. For example, `/vol/vol1`.

   For more information, see the sis(1) man pages.

   ### Example

   ```
   systemA> sis start /vol/vol1
   ```

   Note: You can run this command when you want to start deduplication outside the preset schedule, such as when your system is idle, or when you want to test the impact of deduplication on a particular operation.

When deduplication runs automatically

Deduplication runs automatically when the number of blocks added or changed since the last deduplication operation (performed either manually or automatically) exceeds a specified percentage (20 percent by default) of the total number of blocks that deduplication operations has already processed.

You can configure this value by using the `sis config -s auto@`*`num`* `/`*`vol`*`/`*`volname`* command.

`num` is a two-digit number to specify the percentage.

Example

The following command starts deduplication operations automatically when the specified threshold value is reached:

```
systemA> sis config -s auto@20 /vol/vol1
```

Deduplication operations

You can enable, start, stop, view, and disable deduplication operations.

You can perform the following deduplication tasks:

- Enable deduplication operations.
- Start deduplication operations.
- View the deduplication status of a volume.
- View deduplication space savings.
- Stop deduplication operations.
- Disable deduplication operations.

Next topics

Enabling deduplication operations

To enable deduplication, you use the `sis on` command and specify the FlexVol volume on which you want the deduplication feature to work.

Before you begin

You need to activate the deduplication license before enabling deduplication.

Step

1. Enter the following command:

   `sis on` ***`path`***

   `path` is the complete path to the FlexVol volume.

Example

```
systemA> sis on /vol/vol1
```

Starting a deduplication operation

You can start a deduplication operation on a volume by using the `sis start` command.

Step

1. Enter the following command:

   `sis start [-s] [-f] [-d] [-sp] /vol/`*`volname`*

   The `-s` option scans the volume completely and you are prompted to confirm if deduplication should be started on the volume.

   The `-f` option starts deduplication on the volume without any prompts.

   The `-d` option starts a new deduplication operation after deleting the existing checkpoint information.

   The `-sp` option initiates a deduplication operation by using the previous checkpoint regardless of how old the checkpoint is.

   Note: You can run a maximum of eight concurrent deduplication operations on a system. If any more consecutive deduplication operations are scheduled, the operations are queued. The N5500 and IBM N series N3400 (2859-A11) platforms support a maximum of five concurrent deduplication operations.

Viewing the deduplication status for a volume

You can view the status of deduplication operations on a volume by using the `sis status` command.

Step

1. Enter the following command to view the deduplication status for a volume:

   `sis status -l` ***`path`***

   `path` is the complete path to the FlexVol volume. For example, `/vol/vol1`.

   The `sis status` command is the basic command to view the status of deduplication operations on a volume. For more information about the `sis status` command, see the sis(1) man page.

   The following table lists and describes status and progress messages that you might see after running the `sis status -l` command.

Message	Message type	Description
Idle	status and progress	No active deduplication operation is in progress.
Pending	status	The limit of maximum concurrent deduplication operations allowed for a storage system or a vFiler unit is reached. Any deduplication operation requested beyond this limit is queued.
Active	status	Deduplication operations are running.
size Scanned	progress	A scan of the entire volume is running, of which *size* is already scanned.
size Searched	progress	A search of duplicated data is running, of which *size* is already searched.
size (*pct*) Done	progress	Deduplication operations have saved *size* amounts of data. *pct* is the percentage saved of the total duplicated data that was discovered in the search stage.
size Verified	progress	A verification of the metadata of processed data blocks is running, of which *size* is already verified.
pct% merged	progress	Deduplication operations have merged *pct%* (percentage) of all the verified metadata of processed data blocks to an internal format that supports fast deduplication operations.

Viewing deduplication space savings

You can check how much space you have saved with deduplication by using the `df -s` command.

About this task

The `df -s` command displays the space savings in the active file system only. Space savings in Snapshot copies is not included in the calculation.

Step

1. Enter the following command to view space savings with deduplication:

df -s *volname*

volname is the name of the FlexVol volume. For example, *vol1*.

For more information about the df command, see the df(1) man page.

Note: Using deduplication does not affect volume quotas. Quotas are reported at the logical level, and remain unchanged.

Stopping a deduplication operation

Deduplication consumes system resources during processing. In some situations, it might be advisable to stop currently active deduplication operations using the sis stop command when performance-critical operations such as replication, backup, archiving, or restoration are underway.

Step

1. Enter the following command to stop the deduplication operation:

   sis stop *path*

   path is the complete path to the FlexVol volume. For example, /vol/vol1.

Result

This command stops only the currently active deduplication operation. As long as deduplication operations remain enabled, other deduplication operations will run at their scheduled times.

Disabling deduplication

If deduplication on a specific volume has a performance impact greater than the space savings achieved, you might want to disable deduplication on that volume. You must disable deduplication before removing the deduplication license.

Steps

1. If deduplication is in progress on the volume, enter the following command to abort the operation:

   sis stop *path*

   path is the complete path to the FlexVol volume. For example, /vol/vol1.

2. Enter the following command to disable the deduplication operation:

   sis off *path*

   This command stops all future deduplication operations. For more information about the sis command, see the sis(1) man page.

Deduplication must be disabled before removing the deduplication license

Before removing the deduplication license, you must disable deduplication on all the FlexVol volumes, using the `sis off` command. Otherwise, you will receive a warning message asking you to disable this feature.

Note: Any deduplication operation that occurred before removing the license will remain unchanged.

The deduplication checkpoint feature

The checkpoint is used to periodically log the execution process of a deduplication operation. When a deduplication operation is stopped for any reason (such as system halt, panic, reboot, or last deduplication operation failed or stopped) and checkpoint data exists, the deduplication process can resume from the latest checkpoint file.

You can restart from the checkpoint by using the following commands:

- `sis start -s`
- `sis start` (manually or automatically)

You can view the checkpoint by using the following command:

- `sis status -l`

The checkpoint is created at the end of each stage or sub-stage of the deduplication process. For the `sis start -s` command, the checkpoint is created at every hour during the scanning phase.

If a checkpoint corresponds to the scanning stage (the phase when the `sis start -s` command is run) and is older than 24 hours, the deduplication operation will not resume from the previous checkpoint automatically. In this case, the deduplication operation will start from the beginning. However, if you know that significant changes have not occurred in the volume since the last scan, you can force continuation from the previous checkpoint using the `-sp` option

Related tasks

Starting a deduplication operation on page 260

Starting a deduplication operation with the checkpoint feature

You can start a deduplication operation with the checkpoint feature by using the `sis start` command.

Step

1. Enter the following command:

   `sis start [-s] [-f] [-d] [-sp] /vol/`*`volname`*

The `-s` option scans the volume completely and you are prompted to confirm if deduplication should be started on the volume.

The `-f` option starts deduplication on the volume without any prompts.

The `-d` option starts a new deduplication operation after deleting the existing checkpoint information.

The `-sp` option initiates a deduplication operation using the previous checkpoint, regardless of how old the checkpoint is.

How deduplication works with other features and products

You must keep certain considerations in mind when using deduplication with other features.

Next topics

Deduplication and Snapshot copies

You can run deduplication only on the active file system. However, this data can get locked in Snapshot copies created before you run deduplication, resulting in reduced space savings.

Data can get locked in Snapshot copies in two ways:

- If the Snapshot copies are created before the deduplication operation is run.
  You can avoid this situation by always running deduplication before Snapshot copies are created.
- When the Snapshot copy is created, a part of the deduplication metadata resides in the volume and the rest of the metadata resides in the aggregate outside the volume. The fingerprint files and

the change-log files that are created during the deduplication operation are placed in the aggregate and are not captured in Snapshot copies, which results in higher space savings. However, some temporary metadata files that are created during a deduplication operation are still placed inside the FlexVol volume; these files are deleted after the deduplication operation is complete. These temporary metadata files can get locked in Snapshot copies if the copies are created during a deduplication operation. The metadata remains locked until the Snapshot copies are deleted.

To avoid conflicts between deduplication and Snapshot copies, you should follow these guidelines:

- Run deduplication before creating new Snapshot copies.
- Remove unnecessary Snapshot copies stored in deduplicated volumes.
- Reduce the retention time of Snapshot copies stored in deduplicated volumes.
- Schedule deduplication only after significant new data has been written to the volume.
- Configure appropriate reserve space for the Snapshot copies.
- If snap reserve is 0, you should turn off the schedule for automatic creation of Snapshot copies (which is the case in most LUN deployments).

Deduplication and volume SnapMirror

You can use volume SnapMirror to replicate a deduplicated volume.

When using volume SnapMirror with deduplication, you must consider the following information:

- You need to enable both the deduplication and SnapMirror licenses.
- You can enable deduplication on the source system, the destination system, or both systems.
  The deduplication and the near-line functionality licenses must be enabled on the primary storage system (source). Starting with Data ONTAP 7.3.1, neither the near-line functionality license nor the deduplication license is needed on the destination storage system.

  Note: A deduplication license is not required on the destination storage system. However, if the primary storage system is not available and the secondary storage system becomes the new primary, deduplication must be licensed on the secondary storage system for deduplication to continue. Therefore, you might want to license deduplication on both storage systems.

  You can enable, run, and manage deduplication only from the primary storage system. However, the FlexVol volume in the secondary storage system inherits all the deduplication attributes and storage savings through SnapMirror.
- The shared blocks are transferred only once. Therefore, deduplication also reduces the use of network bandwidth.
- If the source and destination volumes are on different storage system models, they might have different maximum volume sizes.The lower maximum applies. When creating a SnapMirror relationship between two different storage system models, you should ensure that the maximum volume size with deduplication is set to the lower maximum volume size limit of the two models. For example, an N5300 system supports a maximum volume size with deduplication of 4 TB, and an N5600 system supports 16 TB. When establishing a volume SnapMirror relationship between the N5300 and N5600 systems, you should ensure that the volume SnapMirror relationship is established for 4 TB. After a failover, if the volume size on the N5600 system is increased to more than 4 TB, you will not be able to perform a new baseline transfer or resynchronize the

storage systems, because the N5300 system has a maximum volume size with deduplication of 4 TB.

- The volume SnapMirror update schedule does not depend on the deduplication schedule.
  When configuring volume SnapMirror and deduplication, you should coordinate the deduplication schedule and the volume SnapMirror schedule. You should start volume SnapMirror transfers of a deduplicated volume after the deduplication operation is complete. This schedule prevents the sending of undeduplicated data and additional temporary metadata files over the network. If the temporary metadata files in the source volume are locked in Snapshot copies, these files consume extra space in the source and destination volumes.

Starting with Data ONTAP 7.3.1, volumes whose size has been reduced to within the limit supported by deduplication can be part of the SnapMirror primary storage system and the secondary storage system.

Related references

Maximum volume size with deduplication on page 254

Deduplication and qtree SnapMirror

You can use deduplication for volumes that use qtree SnapMirror.

In Data ONTAP 7.3 and later, deduplication operations are supported with qtree SnapMirror. Qtree SnapMirror does not automatically initiate a deduplication operation at the completion of every individual qtree SnapMirror transfer. You can set up a deduplication schedule independent of your qtree SnapMirror transfer schedule.

When using qtree SnapMirror with deduplication, you must consider the following information:

- You need to enable both the deduplication and SnapMirror licenses.

  Note: You can enable deduplication on the source system, the destination system, or both systems.
- Even when deduplication is enabled on the source system, duplicate blocks are sent to the destination system. Therefore, no network bandwidth savings is achieved.
- To recognize space savings on the destination system, you should run deduplication on the destination after the qtree SnapMirror transfer is complete.
- You can set up a deduplication schedule independently of the qtree SnapMirror schedule. For example, on the destination system, the deduplication process does not start automatically after qtree SnapMirror transfers are finished.
- Qtree SnapMirror recognizes deduplicated blocks as changed blocks. Therefore, when you run deduplication on an existing qtree SnapMirror source system for the first time, all the deduplicated blocks are transferred to the destination system. This process might result in a transfer several times larger than the regular transfers.

When using qtree SnapMirror with deduplication, you should ensure that qtree SnapMirror uses only the minimum number of Snapshot copies that it requires. To ensure this minimum, you should retain only the latest Snapshot copies.

Related concepts

Deduplication and transfer of unchanged blocks on page 268

Reverting a SnapMirror destination system with volumes that use deduplication

For a volume SnapMirror relationship, the destination storage system should use the same release of Data ONTAP as the source system or a later release.

In releases earlier than Data ONTAP 7.3.1, when replicating volumes with deduplication, the near-line functionality license is required on the destination system. However, in Data ONTAP 7.3.1 and later,it is not essential that you enable the near-line functionality license on the destination for replicating such volumes.

Therefore, if you revert to a release earlier than Data ONTAP 7.3.1, you should ensure that the near-line functionality license is enabled on the destination system. Otherwise, after the revert operation, volume SnapMirror updates fail for volumes on the source that use deduplication.

Note: When using SnapMirror to replicate volumes that use deduplication, the destination system should support deduplication.

Deduplication and SnapVault

The deduplication feature is integrated with the SnapVault secondary license. This feature increases the efficiency of data backup and improves the use of secondary storage.

The behavior of deduplication with SnapVault is similar to the behavior of deduplication with qtree SnapMirror, with the following exceptions:

- Deduplication is also supported on the SnapVault destination volume.
- The deduplication schedule depends on the SnapVault update schedule on the destination system. However, the deduplication schedule on the source system does not depend on the SnapVault update schedule, and it can be configured independently on a volume.
- Every SnapVault update (baseline or incremental) starts a deduplication process on the destination system after the archival Snapshot copy is taken.
- A new Snapshot copy replaces the archival Snapshot copy after deduplication finishes running on the destination system. (The name of this new Snapshot copy is the same as that of the archival copy, but the Snapshot copy uses a new timestamp, which is the creation time.)
- You cannot configure the deduplication schedule on the destination system manually or run the `sis start` command. However, you can run the `sis start -s` command on the destination system.
- The SnapVault update does not depend on the deduplication operation. A subsequent incremental update is allowed to continue while the deduplication operation on the destination volume from

the previous backup is still in progress. In this case, the deduplication operation continues; however, the archival Snapshot copy is not replaced after the deduplication operation is complete.

- The SnapVault update recognizes the deduplicated blocks as changed blocks. Thus, when deduplication is run on an existing SnapVault source for the first time, all saved space is transferred to the destination system. The size of the transfer might be several times larger than the regular transfers. Running deduplication on the source system periodically will help prevent this issue for future qtree SnapMirror transfers. You should run deduplication before the SnapVault baseline transfer.

Note: You can run a maximum of eight concurrent deduplication operations on a system. This number includes the deduplication operations linked to SnapVault volumes and those that are not linked to SnapVault volumes. The N5500 and N3400 platforms support a maximum of five concurrent deduplication operations.

Deduplication and transfer of unchanged blocks

If qtree SnapMirror or SnapVault updates are performed before the completion of a deduplication operation on the source volume, after the next update, some unchanged blocks might also be transferred to the destination volume.

If deduplication is not running on the destination volume, then the redundant transferred data occupies more storage space on the destination volume.

Before you enable deduplication on the source volume, you should follow these guidelines:

- Ensure that deduplication is run on the qtree SnapMirror or SnapVault destination volume if deduplication is running on the source volume.
- Schedule the qtree SnapMirror update transfers in such a way that these transfers are run only after deduplication is completed on the source volume.

For more information, see the *SnapVault Best Practices Guide*.

Note: This technical report contains information about NetApp products that IBM licenses and in some cases customizes. Technical reports might contain information about models and features that are not supported by IBM.

Related information

TR-3487: SnapVault Best Practices Guide

Deduplication and synchronous SnapMirror

Synchronous SnapMirror is not supported for replicating volumes that use deduplication.

Deduplication and tape backups

Backup to a tape through the SMTape engine preserves deduplication on the restored volume. However, backups to a tape, either through NDMP or the native `dump` command, do not preserve

deduplication. Therefore, if you want to regain space savings on a volume restored from tape, you must run the `sis start -s` command on the restored volume.

Deduplication and SnapRestore

The metadata created during a deduplication operation is located in the aggregate. Therefore, when you initiate a SnapRestore operation on a volume, the metadata is not restored to the active file system. The restored data, however, retains the original space savings.

To run deduplication for all the data on the volume, you should use the `sis start -s` command.

This command builds the fingerprint database for all the data in the volume.

Deduplication and SnapLock volumes

Beginning with Data ONTAP 7.3.1, deduplication is seamlessly supported on SnapLock volumes.

However, if deduplication is enabled on a SnapLock volume and you attempt to revert to a Data ONTAP release that does not support deduplication on SnapLock volumes, the system will display an error message. To recover from this situation, contact technical support.

Deduplication and MetroCluster

Data ONTAP supports deduplication on stretch and fabric-attached MetroCluster. This support applies to filers and gateways.

For more information about deduplication support on active/active configurations and MetroCluster, see TR-3505, *NetApp Deduplication for FAS Deployment and Implementation Guide*.

Note: This technical report contains information about NetApp products that IBM licenses and in some cases customizes. Technical reports might contain information about models and features that are not supported by IBM.

Related information

TR-3505, NetApp Deduplication for FAS Deployment and Implementation Guide

Deduplication and DataFabric Manager

Starting with Data ONTAP 7.3.1, deduplication is supported with the Management Console data protection capability, the Management Console provisioning capability, and Operations Manager in DataFabric Manager 3.8.

Deduplication and the Management Console data protection capability in DataFabric Manager 3.8

In releases earlier than DataFabric Manager 3.8, the Management Console data protection capability waits for an active deduplication operation to complete before renaming the Snapshot copies. While the Management Console data protection capability waits, it does not allow clients to list the

Snapshot copies or restore from them. Therefore, in releases prior to DataFabric Manager 3.8, the use of deduplication with the Management Console data protection capability is not optimal.

However, this limitation is removed in DataFabric Manager 3.8.

Deduplication and the Management Console provisioning capability in DataFabric Manager 3.8

With the Management Console provisioning capability in DataFabric Manager 3.8, you can enable the provisioning policies to support all three modes of deduplication, namely, on-demand deduplication, automated deduplication, and scheduled deduplication.

For more information about using deduplication with the Management Console provisioning capability and the Management Console data protection capability, see the *Provisioning Manager and Protection Manager Guide to Common Workflows for Administrators*.

Deduplication and Operations Manager in DataFabric Manager 3.8

Deduplication is integrated into Operations Manager in DataFabric Manager 3.8.

You can configure deduplication on the system and generate reports or graphs summarizing space savings for file and LUN clones.

For more information about using deduplication with Operations Manager, see the *Operations Manager Administration Guide*.

Related information

IBM NAS documentation and support site - www.ibm.com/storage/support/nas

Deduplication and volume copy

Volume copy is a method of copying both data in the active file system and data in storage systems from one volume to another. The source and destination volumes must both be FlexVol volumes.

When deduplicated data is copied by using the `vol copy` command, the copy of the data at the destination inherits all the deduplication attributes and storage savings of the source data.

Starting with Data ONTAP 7.3, the metadata created during a deduplication operation (fingerprint files and changelog files) are located outside the volume in the aggregate. Therefore, when you run the volume copy operation on a volume, the fingerprint files and change-log files are not restored to the active file system. After a volume copy operation, if deduplication is enabled on the volume, any new data written to the volume continues to be deduplicated. However, space savings is only obtained for the new data.

To run deduplication for all the data on the volume, you should use the `sis start -s` command.

This command builds the fingerprint database for all the data in the volume. The amount of time this process takes depends on the size of the logical data in the volume. Before using the `sis start -s` command, you must ensure that the volume and the aggregate containing the volume have sufficient free space for deduplication metadata.

Deduplication and FlexClone volumes

Deduplication is supported on FlexClone volumes. FlexClone volumes are writable clones of a parent FlexVol volume.

The FlexClone volume of a deduplicated volume is a deduplicated volume. The cloned volume inherits the deduplication configuration of the parent volume (for example, deduplication schedules).

The FlexClone volume of a non-deduplicated volume is a non-deduplicated volume. If you run deduplication on a clone volume, the clone is deduplicated, but the original volume remains non-deduplicated.

Starting with Data ONTAP 7.3, the metadata created during a deduplication operation (fingerprint files and change-log files) are located outside the volume in the aggregate; therefore, they are not cloned. However, the data retains the space savings of the original data.

Any new data written to the destination volume continues to be deduplicated and fingerprint files for the new data are created. Space savings is only obtained for the new data.

To run deduplication for all the data on the cloned volume, you should use the `sis start -s` command. The time the process takes to finish depends on the size of the logical data in the volume.

When a cloned volume is split from the parent volume, deduplication of all data in the clone that was part of the parent volume is undone after the volume-split operation. However, if deduplication is running on the clone volume, the data is deduplicated in the subsequent deduplication operation.

Deduplication and an active/active configuration

You can activate deduplication in an active/active configuration.

The maximum number of concurrent deduplication operations allowed on each node of an active/active configuration is eight. If one of the nodes fails, the other node takes over the operations of the failed node. In takeover mode, the working node continues with its deduplication operations as usual. However, the working node does not start any deduplication operations on the failed node.

Note:

- The N5500 and N3400 platforms support a maximum of five concurrent deduplication operations.
- Change logging for volumes with deduplication continues for the failed node in takeover mode. Therefore, deduplication operations can be performed on data written during takeover mode after the failed node is active, and there is no loss in space savings. To disable change logging for volumes that belong to a failed node, you can turn off deduplication on those volumes. You can also view the status of volumes with deduplication for a failed node in takeover mode.

Deduplication and nondisruptive upgrade

When you upgrade nondisruptively to a Data ONTAP 7.3.2 release from an earlier release family, deduplication is enabled and deduplication schedules are maintained for all volumes, after the upgrade.

Deduplication and VMware

You can run deduplication in VMware environments for efficient space savings.

While planning the Virtual Machine Disk (VMDK) and data store layouts, you should follow these guidelines.

- Operating system VMDKs deduplicate efficiently because the binary files, patches, and drivers are highly redundant between virtual machines. You can achieve maximum savings by keeping these VMDKs in the same volume.
- Application binary VMDKs deduplicate to varying degrees. Applications from the same vendor commonly have similar libraries installed; therefore, you can achieve moderate deduplication savings. Applications written by different vendors do not deduplicate at all.
- Application datasets when deduplicated have varying levels of space savings and performance impact based on the application and intended use. You should carefully consider what application data needs to be deduplicated.
- Transient and temporary data, such as VM swap files, pagefiles, and user and system temp directories, does not deduplicate well and potentially adds significant performance impact when deduplicated. Therefore, it is best to keep this data on a separate VMDK and volume that are not deduplicated.

Application data has a major effect on the percentage of storage savings achieved with deduplication. New installations typically achieve large deduplication savings.

Note: In VMware environments, proper partitioning and alignment of the VMDKs is important. Applications whose performance is impacted by deduplication operations are likely to have the same performance impact when you run deduplication in a VMware environment.

Improved sequential read performance for deduplicated FlexVol volumes

In Data ONTAP 7.2.6, Data ONTAP 7.3.1, and later releases of these two release families, the performance of sequential read operations on highly-deduplicated data, including large deduplicated VMDK files, has been greatly improved.

In VMware environments, the VMDKs are created with a large number of duplicate blocks. This results in a high number of shared blocks after running deduplication. Therefore, applications that perform large sequential read operations, such as `dump`, when run on VMDK files might have low throughput.

Highly-shared blocks are efficiently cached and read from the cache instead of from the disk every time. This caching improves the sequential read performance on VMDK files.

Deduplication and MultiStore

Deduplication commands are available in all the vfiler contexts. Deduplication support on vFiler units allows users to reduce redundant data blocks within vFiler units.

You can enable deduplication only on FlexVol volumes in a vFiler unit. Deduplication support on vFiler units ensures that volumes owned by a vFiler unit are not accessible to another vFiler unit. Deduplication also supports disaster recovery and migration of vFiler units. If you enable deduplication on the volume in the source vFiler unit, the destination vFiler unit inherits all deduplication attributes.

You must license deduplication on the primary storage system. It is best that you also license deduplication on the secondary storage system. These licenses ensure that deduplication operations can continue without any disruption in case a failure causes the secondary vFiler unit to become the primary storage system.

To use the deduplication feature, you should activate the following licenses on the storage system:

- `multistore`
- `a_sis`

You can run deduplication commands using the RSH or SSH protocol. Any request is routed to the IP address and IP space of the destination vFiler unit.

Note: During an online migration of a vFiler unit, the following deduplication operations are not allowed on volumes that are owned by vFiler units:

- `sis start`
- `sis start -s`
- `sis on`
- `sis off`
- `sis config -s`

For more information about disaster recovery and migration of vFiler units, see the *Data ONTAP MultiStore Management Guide*.

Next topics

How to run deduplication on a vFiler unit using the CLI on page 273
How to set the maximum deduplication sessions per vFiler unit on page 275

How to run deduplication on a vFiler unit using the CLI

You can run deduplication on a vFiler unit by using the command-line interface (CLI).

The following deduplication commands are available from the `vfiler` context.

- `sis on`

- `sis off`
- `sis start`
- `sis config`
- `sis status`
- `sis stop`

All deduplication commands ensure boundary checks for each vFiler unit. This mechanism prevents any attempt to access volumes that do not belong to the requesting vFiler unit. You must switch to the `vfiler` context of the vFiler unit that owns the FlexVolvolume. Thereafter, you can run deduplication commands on the FlexVol volume.

Example:

The FlexVol volumes vola and volb are owned by vFiler units, vf1 and vf2, respectively. To switch context, you should issue the following commands.

- `vfiler context vf1`
- `sis on /vol/vola`
- `sis start -s /vol/vola`

When you switch context in this manner, you can run deduplication commands on vf1 and vola. However, you cannot run commands on vf2 or volb, because their context is vf2, not vf1. Therefore, the following command fails because the context is vf1: `vfiler context vf1; sis on /vol/volb`

This command fails because the context is vf1.

The output of these commands is specific to the `vfiler` context. These commands display information about all volumes that are contained within the current vFiler context.

- `sis config`—No volume name is specified.
- `sis status`—No volume name is specified.

Using the `vfiler run` command, you can specify the target vFiler unit of the command as an argument. This ensures that a proper `vfiler` context is assigned before a deduplication command is run.

`vfiler run` executes the command following it in the specified `vfiler` context.

`vfiler run [-q] vfilertemplate sis_command [args]`

The `run` subcommand runs the command on the vFiler units specified by *`vfilertemplate`*. If more than one vFiler unit is named, the command should be run for each vFiler unit. Any vFiler unit console command specific to vFiler units can be used. If the command is not specific to vFiler units, an error message is logged and the command fails.

How to set the maximum deduplication sessions per vFiler unit

You can specify the number of concurrent deduplication sessions that can be run per vFiler unit by using the `option sis.max_vfiler_active_ops` command.

Note: The maximum number of concurrent deduplication operations per storage system is eight. The command first checks the sis operations on the physical storage system, and then on the vFiler unit. On a 32 bit platform, the default number of concurrent deduplication sessions that can be run per vFiler unit is five.

Deduplication and volume move

During the cutover phase of a volume move operation, some of the deduplication operations are not allowed on the FlexVol volume that is being moved.

The following deduplication commands are not allowed:

- `sis start`
- `sis start -s`
- `sis on`
- `sis off`
- `sis config -s`
- `sis config -m`
- `sis revert_to`

If you try to nondisruptively move a FlexVol volume that has deduplication operations running, the volume move operation does not enter the cutover phase. The volume move operation is paused until the deduplication operations are completed.

For more information about volume move, see the *Data ONTAP Block Access Management Guide for iSCSI and FC*.

Related information

IBM NAS documentation and support site — www.ibm.com/storage/support/nas

Common troubleshooting procedures for volumes with deduplication

You need to know the common troubleshooting procedures for issues that might occur while configuring and running deduplication on FlexVol volumes.

Issues related to licensing

You should ensure that deduplication is licensed.

For all platforms, you should ensure that the near-line functionality license is also enabled.

You can check for active licenses by entering the `license` command. A license key must be displayed next to the installed license.

- `a_sis` *`license key`*
- `nearstore_options` *`license key`*

  Note: The `nearstore_options` license enables near-line functionality.

If the license is either removed or has expired, all `sis` commands fail and no additional deduplication occurs. However, the FlexVol volume remains deduplicated, and existing storage savings are retained.

Issues related to volume size

You must ensure that there is space available for the `sis on` command to complete successfully.

If you are running Data ONTAP 7.2, you need to leave approximately 6 percent additional space on the FlexVol volume that you plan to enable deduplication on. This is because the metadata in Data ONTAP 7.2 resides in the volume.

If you are running Data ONTAP 7.3, you need to leave approximately 2 percent additional space in the volume you are planning to enable deduplication on, and 4 percent outside the volume in the aggregate. Starting with Data ONTAP 7.3, some metadata resides in the volume and some outside the volume in the aggregate.

Issues related to space savings

If you run deduplication on a FlexVol volume with data that can be deduplicated efficiently, but the space savings after deduplication is not consistent with the data on the volume, data on the FlexVol volume might be locked by Snapshot copies. This normally occurs when deduplication is run on an existing FlexVol volume.

You should use the `snap list` command to check the Snapshot copies that exist, and use the `snap delete` command to delete them. Alternatively, you can wait for the Snapshot copies to expire, which results in space savings. You might also see less-than-expected savings if the aggregate runs out of space, not allowing additional deduplication metadata to be stored.

How space management works

The space management capabilities of Data ONTAP allow you to configure your storage systems to provide the storage availability required by the users and applications accessing the system, while using your available storage as effectively as possible.

Data ONTAP enables space management using the following capabilities:

- Space guarantees
- Space reservations
- Fractional reserve
- Automatic free space preservation

Next topics

What kind of space management to use on page 277
What space guarantees are on page 279
What space reservation is on page 280
How Data ONTAP can automatically provide more free space for full volumes on page 164
How aggregate overcommitment works on page 282

What kind of space management to use

The type of space management you should use depends on many factors, including your tolerance for out-of-space errors, whether you plan to overcommit your aggregates, and your rate of data overwrite.

The following table can help you determine which space management capabilities best suit your requirements.

Note: LUNs in this context refer to the LUNs that Data ONTAP serves to clients, not to the array LUNs used for storage on a storage array.

If...	Then use...	Typical usage	Notes
You want management simplicity	FlexVol volumes with a space guarantee of `volume` OR Traditional volumes	NAS file systems	This is the easiest option to administer. As long as you have sufficient free space in the volume, writes to any file in this volume will always succeed.

If...	Then use...	Typical usage	Notes
Writes to certain files must always succeed You want to overcommit your aggregate	FlexVol volumes with a space guarantee of `file` OR Traditional volume AND space reservation enabled for files that require writes to succeed	LUNS Databases	This option enables you to guarantee writes to specific files.
You need even more effective storage usage than file space reservation provides You actively monitor available space on your volume and can take corrective action when needed Snapshot copies are short-lived Your rate of data overwrite is relatively predictable and low	FlexVol volumes with a space guarantee of `volume` OR Traditional volume AND Space reservation on for files that require writes to succeed AND Fractional reserve < 100%	LUNs (with active space monitoring) Databases (with active space monitoring)	With fractional reserve <100%, it is possible to use up all available space, even with space reservations on. Before enabling this option, be sure either that you can accept failed writes or that you have correctly calculated and anticipated storage and Snapshot copy usage.
You want to overcommit your aggregate You actively monitor available space on your aggregate and can take corrective action when needed	FlexVol volumes with a space guarantee of `none`	Storage providers who need to provide storage that they know will not immediately be used Storage providers who need to allow available space to be dynamically shared between volumes	With an overcommitted aggregate, writes can fail due to insufficient space.

Related concepts

What space guarantees are on page 279
How volumes work on page 155

What space guarantees are

Space guarantees on a FlexVol volume ensure that writes to a specified FlexVol volume or writes to files with space reservations enabled do not fail because of lack of available space in the containing aggregate.

Space guarantee is an attribute of the volume. It is persistent across storage system reboots, takeovers, and givebacks. Space guarantee values can be `volume` (the default value), `file`, or `none`.

- A space guarantee of `volume` reserves space in the aggregate for the volume. The reserved space cannot be allocated to any other volume in that aggregate.

  The space management for a FlexVol volume that has a space guarantee of `volume` is equivalent to a traditional volume.
- A space guarantee of `file` reserves space in the aggregate so that any file in the volume with space reservation enabled can be completely rewritten, even if its blocks are being retained on disk by a Snapshot copy.

  Note: Writes to a volume with a space guarantee of `file` could fail. Because write errors are unexpected in a CIFS environment, do not set the space guarantee to `file` for volumes accessed using CIFS.
- A FlexVol volume that has a space guarantee of `none` reserves no extra space for user data; writes to LUNs or files contained by that volume could fail if the containing aggregate does not have enough available space to accommodate the write.

  Note: Because out-of-space errors are unexpected in a CIFS environment, do not set the space guarantee to `none` for volumes accessed using CIFS.

When space in the aggregate is reserved for space guarantee for an existing volume, that space is no longer considered free space. Operations that consume free space in the aggregate, such as creation of Snapshot copies or creation of new volumes in the containing aggregate, can occur only if there is enough available free space in that aggregate; these operations are prevented from using space already committed to another volume.

When the uncommitted space in an aggregate is exhausted, only writes to volumes or files in that aggregate with space guarantees are guaranteed to succeed.

Note: Space guarantees are honored only for online volumes. If you take a volume offline, any committed but unused space for that volume becomes available for other volumes in that aggregate. When you bring that volume back online, if there is not sufficient available space in the aggregate to fulfill its space guarantees, you must use the force (`-f`) option, and the volume's space guarantees are disabled. When a volume's space guarantee is disabled, the word `(disabled)` appears next to its space guarantees in the output of the `vol status` command.

Next topics

What kind of space guarantee traditional volumes provide on page 280

How you set space guarantees for new or existing volumes on page 280

What kind of space guarantee traditional volumes provide

Traditional volumes provide the same space guarantee as FlexVol volumes with space guarantee of `volume`. To guarantee that writes to a specific file in a traditional volume will always succeed, you need to enable space reservations for that file. (LUNs have space reservations enabled by default.)

How you set space guarantees for new or existing volumes

To set the space guarantee for an existing volume, you use the `vol options` command with the `guarantee` option. To set the space guarantee for a new volume, you use the `-s` option for the `vol create` command. Space guarantees can be `volume`, `file`, or `none`.

What space reservation is

When space reservation is enabled for one or more files or LUNs, Data ONTAP reserves enough space in the volume (traditional or FlexVol) so that writes to those files or LUNs do not fail because of a lack of disk space.

Note: LUNs in this context refer to the LUNs that Data ONTAP serves to clients, not to the array LUNs used for storage on a storage array.

For example, if you create a 100-GB space reserved LUN in a 500-GB volume, that 100 GB of space is immediately allocated, leaving 400 GB remaining in the volume. In contrast, if space reservation is disabled on the LUN, all 500 GB in the volume remain available until writes are made to the LUN.

Space reservation is an attribute of the file or LUN; it is persistent across storage system reboots, takeovers, and givebacks. Space reservation is enabled for new LUNs by default, but you can create a LUN with space reservations disabled or enabled. After you create the LUN, you can change the space reservation attribute by using the `lun set reservation` command. You can change the space reservation for files by using the `file reservation` command.

When a volume contains one or more files or LUNs with space reservation enabled, operations that require free space, such as the creation of Snapshot copies, are prevented from using the reserved space. If these operations do not have sufficient unreserved free space, they fail. However, writes to the files or LUNs with space reservation enabled will continue to succeed.

How Data ONTAP can automatically provide more free space for full volumes

Data ONTAP can automatically make more free space available for a FlexVol volume when that volume is nearly full. You can choose to make the space available by first allowing the volume size to increase, or by first deleting Snapshot copies.

You enable this capability for a FlexVol volume by using the `vol options` command with the `try_first` option.

Data ONTAP can automatically provide more free space for the volume by using one of the following methods:

- Increase the size of the volume when it is nearly full.
  This method is useful if the volume's containing aggregate has enough space to support a larger volume. You can increase the size in increments and set a maximum size for the volume.

  Note: The autosize capability is disabled by default, so you must enable and configure it by using the `vol autosize` command. You can use the `vol status -v` command to view the current autosize settings for a volume.
- Delete Snapshot copies when the volume is nearly full.
  For example, you can automatically delete Snapshot copies that are not linked to Snapshot copies in cloned volumes or LUNs, or you can define which Snapshot copies you want to delete first—your oldest or newest Snapshot copies. You can also determine when to begin deleting Snapshot copies—for example, when the volume is nearly full or when the volume's Snapshot reserve is nearly full.
  You use the `snap autodelete` command to configure automatic Snapshot copy deletion. For more information about deleting Snapshot copies automatically, see the *Data ONTAP Data Protection Online Backup and Recovery Guide*.

You can choose which method (increasing the size of the volume or deleting Snapshot copies) you want Data ONTAP to try first. If the first method does not provide sufficient extra free space to the volume, Data ONTAP will try the other method next.

Related tasks

Configuring a FlexVol volume to grow automatically on page 184
Configuring automatic free space preservation for a FlexVol volume on page 184

How aggregate overcommitment works

Using aggregate overcommitment, the storage administrator can appear to provide more storage than is actually available from a given aggregate, as long as not all of that storage is currently being used. Aggregate commitment is also called *thin provisioning*.

To use aggregate overcommitment, you create FlexVol volumes with a space guarantee of `none` or `file`. With a space guarantee of `none` or `file`, the volume size is not limited by the aggregate size. In fact, each volume could, if required, be larger than the containing aggregate. The storage provided by the aggregate is used up only as LUNs are created or data is appended to files in the volumes.

Note: LUNs in this context refer to the LUNs that Data ONTAP serves to clients, not to the array LUNs used for storage on a storage array.

Note: The aggregate must provide enough free space to hold the metadata for each FlexVol volume it contains. The space required for a FlexVol volume's metadata is approximately 0.5 percent of the volume's nominal size.

This could be useful if you are asked to provide greater amounts of storage than you know will be used immediately. Alternatively, if you have several volumes that sometimes need to grow temporarily, the volumes can dynamically share the available space with each other.

When the aggregate is overcommitted, it is possible for these types of writes to fail due to lack of available space:

- Writes to any volume with space guarantee of `none`
- Writes to any file that does not have space reservations enabled and that is in a volume with space guarantee of `file`

Therefore, if you have overcommitted your aggregate, you must monitor your available space and add storage to the aggregate as needed to avoid write errors due to insufficient space.

For more information about aggregate overcommitment, see Technical Reports 3563 & 3483.

Note: This technical report contains information about NetApp products that IBM licenses and in some cases customizes. Technical reports might contain information about models and features that are not supported by IBM.

Related information

TR-3563: NetApp Thin Provisioning
TR 3483: Thin Provisioning in a NetApp SAN or IP SAN Enterprise Environment

Considerations for bringing a volume online in an overcommited aggregate

When you take a FlexVol volume offline, it releases its allocation of storage space in its containing aggregate. Other volumes in that aggregate might start using that space while the volume is offline. When this happens, you cannot bring the volume back online as you normally would.

If you attempt to bring a FlexVol volume online when there is insufficient free space in the aggregate to honor its space guarantees, the `vol online` command fails. You can use the `-f` option to force the volume to come online; however, the space guarantees for that volume are disabled. If you later make more space available to the aggregate, the space guarantees will be automatically re-enabled.

Attention: Attempts to write to a volume with its space guarantees disabled could fail due to insufficient available space. For this reason, in environments that are sensitive to that error (such as CIFS or LUNs), it is best to avoid forcing a volume online.

Note: FlexCache volumes cannot be brought online if there is insufficient space in the aggregate to fulfill their space guarantee.

About qtrees

Qtrees enable you to partition your volumes into smaller segments that you can manage individually. You can set a qtree's size or security style, back it up, and restore it.

Next topics

When you use qtrees

Qtrees allow you to partition your data without incurring the overhead associated with a volume. You might create qtrees to organize your data, or to manage one or more of the following factors: quotas, backup strategy, security style, and CIFS oplocks setting.

The following list describes examples of qtree usage strategies:

- Quotas
  You can limit the size of the data used by a particular project, by placing all of that project's files into a qtree and applying a tree quota to the qtree.
- Backups
  You can use qtrees to keep your backups more modular, to add flexibility to backup schedules, or to limit the size of each backup to one tape.
- Security style
  If you have a project that needs to use NTFS-style security, because the members of the project use Windows files and applications, you can group the data for that project in a qtree and set its security style to NTFS, without requiring that other projects also use the same security style.
- CIFS oplocks settings
  If you have a project using a database that requires CIFS oplocks to be off, you can set CIFS oplocks to Off for that project's qtree, while allowing other projects to retain CIFS oplocks.

How qtrees compare with volumes

In general, qtrees are similar to volumes. However, they have some key differences.

The following table compares qtrees, FlexVol volumes, and traditional volumes.

Functionality	Qtree	FlexVol volume	Traditional volume
Enables organizing user data	Yes	Yes	Yes
Enables grouping users with similar needs	Yes	Yes	Yes
Accepts a security style	Yes	Yes	Yes
Accepts oplocks configuration	Yes	Yes	Yes
Can be backed up and restored as a unit using SnapMirror	Yes	Yes	Yes
Can be backed up and restored as a unit using SnapVault	Yes	No	No
Can be resized	Yes (using quota limits)	Yes	No (can be expanded but cannot be reduced in size)
Supports Snapshot copies	No (qtree data can be extracted from volume Snapshot copies)	Yes	Yes
Supports quotas	Yes	Yes	Yes
Can be cloned	No (except as part of a FlexVol volume)	Yes	No
Maximum number allowed	4,995 per volume	500 per system	100 per system

Qtree name restrictions

Using some special characters in qtree names, such as commas and spaces, can cause problems with other Data ONTAP capabilities, and should be avoided.

The following characters should be avoided in qtree names:

- Space
  Spaces in qtree names can prevent SnapMirror updates from working correctly.
- Comma
  Commas in qtree names can prevent quotas from working correctly for that qtree, unless the name is enclosed in double quotation marks.

Managing qtrees

You can create, delete, and rename qtrees. In addition, you can display their status and access statistics. You can also convert directories at the root of a volume into qtrees. You do many of these operations using your UNIX or Windows client.

About this task

Note: Many qtree commands cannot be performed while a volume move operation is in progress. If you are prevented from completing a qtree command for this reason, wait until the volume move is complete and then retry the command.

Next topics

Creating a qtree

You create qtrees using the `qtree create` command. You can also specify a UNIX-style permission for the new qtree.

Steps

1. Enter the following command:

   `qtree create` ***`path`*** **`[-m`** ***`mode`*** **`]`**

   `mode` is a UNIX-style octal number that specifies the permissions for the new qtree. If you do not specify a mode, the qtree is created with the permissions specified by the `wafl.default_qtree_mode` option.

   For more information about the format of the mode number, see your UNIX documentation.

   Note: If you are using this qtree in an NTFS-only environment, you can set the appropriate ACLs after creation using Windows tools.

   `path` is the path name of the qtree, with the following notes:

- If you want to create the qtree in a volume other than the root volume, include the volume in the name.
- If the path name does not begin with a slash (/), the qtree is created in the root volume.
- Qtree names can be up to 64 characters long. The entire path can be up to 1,024 characters long.

2. If you want to change the default security style or the default CIFS oplocks setting of the new qtree, you can change it now by using the `qtree security` or `qtree oplocks` commands.

Examples

The following command creates the news qtree in the users volume, giving the owner and the owner's group permission to read, write and execute the qtree:

```
qtree create /vol/users/news -m 770
```

The following command creates the news qtree in the root volume:

```
qtree create news
```

Related concepts

Qtree name restrictions on page 286
How security styles affect access to your data on page 162
About qtrees on page 285

Displaying qtree status

To find the security style, oplocks attribute, and SnapMirror status for all volumes and qtrees on the storage system or for a specified volume, you use the `qtree status` command.

Step

1. Enter the following command:

   ```
   qtree status [-i] [-v] [vol_name]
   ```

   The `-i` option includes the qtree ID number in the display.

   The `-v` option includes the owning vFiler unit, if the MultiStore license is enabled.

Displaying qtree access statistics

You display statistics on user accesses to files in qtrees on your system using the `qtree stats` command. This can help you determine which qtrees are incurring the most traffic. Determining traffic patterns helps with qtree-based load balancing.

About this task

The `qtree stats` command displays the number of NFS and CIFS accesses to the designated qtrees since the counters were last reset. The qtree stats counters are reset when one of the following actions occurs:

- The system is booted.
- The volume containing the qtree is brought online.
- The counters are explicitly reset using the `qtree stats -z` command.

Step

1. Enter the following command:

   `qtree stats [-z] [`*`vol_name`*`]`

   The `-z` option clears the counter for the designated qtree, or clears all counters if no qtree is specified.

   `vol_name` optionally specifies a volume. Statistics for all qtrees in that volume are displayed. If no volume is specified, statistics for all qtrees on the storage system are displayed.

Example output

```
system> qtree stats vol1
Volume        Tree         NFS ops       CIFS ops
--------      --------     -------       --------
vol1          proj1           1232             23
vol1          proj2             55            312
```

Converting a directory to a qtree

If you have a directory at the root of a volume that you want to convert to a qtree, you must migrate the data contained in the directory to a new qtree with the same name, using your client application.

About this task

The exact steps you take to convert a directory to a qtree depend on what client you use. The following process outlines the general tasks you need to complete.

Steps

1. Rename the directory to be made into a qtree.
2. Create a new qtree with the original directory name.
3. Use the client application to move the contents of the directory into the new qtree.
4. Delete the now-empty directory.

   Note: You cannot delete a directory if it is associated with an existing CIFS share.

Next topics

Converting a directory to a qtree using a Windows client on page 290
Converting a directory to a qtree using a UNIX client on page 291

Converting a directory to a qtree using a Windows client

To convert a directory to a qtree using a Windows client, you rename the directory, create a qtree on the storage system, and move the directory's contents to the qtree.

About this task

You must use Windows Explorer for this procedure. You cannot use the Windows command-line interface or the DOS prompt environment.

Steps

1. Open Windows Explorer.
2. Click the folder representation of the directory you want to change.

   Note: The directory must reside at the root of its containing volume.
3. From the File menu, select Rename to give this directory a different name.
4. On the storage system, use the `qtree create` command to create a new qtree with the original name of the directory.
5. In Windows Explorer, open the renamed directory folder and select the files inside it.
6. Drag these files into the folder representation of the new qtree.

   Note: The more subfolders contained in the folder that you are moving, the longer the move operation takes.
7. From the File menu, select Delete to delete the renamed, now-empty directory folder.

Converting a directory to a qtree using a UNIX client

To convert a directory to a qtree in UNIX, you rename the directory, create a qtree on the storage system, and move the directory's contents to the qtree.

Steps

1. Open a UNIX client window.
2. Use the `mv` command to rename the directory.

   Example

   ```
   client: mv /n/joel/vol1/dir1 /n/joel/vol1/olddir
   ```

3. From the storage system, use the `qtree create` command to create a qtree with the original name.

   Example

   ```
   system1: qtree create /n/joel/vol1/dir1
   ```

4. From the client, use the `mv` command to move the contents of the old directory into the qtree.

   Note: The more subdirectories contained in a directory that you are moving, the longer the move operation will take.

   Example

   ```
   client: mv /n/joel/vol1/olddir/* /n/joel/vol1/dir1
   ```

5. Use the `rmdir` command to delete the old, now-empty directory.

   Example

   ```
   client: rmdir /n/joel/vol1/olddir
   ```

After you finish

Depending on how your UNIX client implements the `mv` command, file ownership and permissions might not be preserved. If this occurs, update file owners and permissions to their previous values.

Deleting a qtree

You can delete a qtree using Windows Explorer or a UNIX client, if the qtree permissions allow.

Before you begin

Ensure that the following conditions are true:

- The volume that contains the qtree you want to delete is mounted (for NFS) or mapped (for CIFS).
- The qtree you are deleting is not directly mounted and does not have a CIFS share directly associated with it.
- The qtree permissions allow you to modify the qtree.

Steps

1. Find the qtree you want to delete.

   Note: The qtree appears as a normal directory at the root of the volume.

2. Delete the qtree using the method appropriate for your client.

   Example
   The following command on a UNIX host deletes a qtree that contains files and subdirectories:

   ```
   rm -Rf directory
   ```

   Note: On a Windows host, you must use Windows Explorer to delete a qtree.

Related concepts

How deleting a qtree affects tree quotas on page 315
About qtrees on page 285

Renaming a qtree

You can rename a qtree using Windows Explorer or a UNIX client, if the qtree permissions allow.

Before you begin

Ensure that the following conditions are true:

- The volume that contains the qtree you want to rename is mounted (for NFS) or mapped (for CIFS).
- The qtree you are renaming is not directly mounted and does not have a CIFS share directly associated with it.
- The qtree permissions allow you to modify the qtree.

Steps

1. Find the qtree you want to rename.

   Note: The qtree appears as a normal directory at the root of the volume.

2. Rename the qtree using the method appropriate for your client.

Example

The following command on a UNIX host renames a qtree:

```
mv old_name new_name
```

Note: On a Windows host, you must use Windows Explorer to rename a qtree.

After you finish

If you have quotas on the renamed qtree, update the quotas file to use the new qtree name.

Related concepts

How renaming a qtree affects quotas on page 315
About qtrees on page 285

Managing CIFS oplocks

CIFS oplocks reduce network traffic and improve storage system performance. However, in some situations, you might need to disable them. You can disable CIFS oplocks for the entire storage system or for a specific volume or qtree.

Next topics

About the CIFS oplocks setting on page 161
Enabling or disabling CIFS oplocks for the entire storage system on page 296
Enabling CIFS oplocks for a specific volume or qtree on page 296
Disabling CIFS oplocks for a specific volume or qtree on page 296

About the CIFS oplocks setting

Usually, you should leave CIFS oplocks on for all volumes and qtrees. This is the default setting. However, you might turn CIFS oplocks off under certain circumstances.

CIFS oplocks (opportunistic locks) enable the redirector on a CIFS client in certain file-sharing scenarios to perform client-side caching of read-ahead, write-behind, and lock information. A client can then work with a file (read or write it) without regularly reminding the server that it needs access to the file. This improves performance by reducing network traffic.

You might turn CIFS oplocks off on a volume or a qtree under either of the following circumstances:

- You are using a database application whose documentation recommends that CIFS oplocks be turned off.
- You are handling critical data and cannot afford even the slightest data loss.

Otherwise, you can leave CIFS oplocks on.

For more information about CIFS oplocks, see the CIFS section of the *Data ONTAP File Access and Protocols Management Guide*.

Related tasks

Enabling or disabling CIFS oplocks for the entire storage system on page 296
Enabling CIFS oplocks for a specific volume or qtree on page 296
Disabling CIFS oplocks for a specific volume or qtree on page 296

Enabling or disabling CIFS oplocks for the entire storage system

You use the `cifs.oplocks.enable` option to enable or disable CIFS oplocks for the entire storage system. If you set this option to On, then CIFS oplocks are enabled, and the individual setting for each qtree and volume takes effect.

Enabling CIFS oplocks for a specific volume or qtree

If you've previously disabled CIFS oplocks for a specific volume or qtree, and now you want to reenable them, you can do so by using the `qtree oplocks` command.

Steps

1. Ensure that the `cifs.oplocks.enable` option is set to `on`.

   Otherwise, enabling CIFS oplocks for a specific volume or qtree has no effect.

2. Enter the following command:

   `qtree oplocks` *`path`* `enable`

Example

To enable CIFS oplocks on the proj1 qtree in vol2, use the following commands:

```
sys1> options cifs.oplocks.enable on
sys1> qtree oplocks /vol/vol2/proj enable
```

After you finish

You can verify the update by using the `qtree status` command, using the name of the containing volume if you updated the CIFS oplocks for a qtree.

Disabling CIFS oplocks for a specific volume or qtree

If you want to disable CIFS oplocks for a specific volume or qtree, you can do so by using the `qtree oplocks` command.

Step

1. Enter the following command:

qtree oplocks *path* disable

Example

To disable CIFS oplocks on the proj1 qtree in vol2, use the following command:

qtree oplocks /vol/vol2/proj disable

After you finish

You can verify the update by using the `qtree status` command, using the name of the containing volume if you updated the CIFS oplocks for a qtree.

Changing security styles

You might need to change the security style of a new volume or qtree. Additionally, you might need to accommodate other users; for example, if you had an NTFS qtree and subsequently needed to include UNIX files and users, you could change the security style of that qtree from NTFS to mixed.

Before you begin

Make sure there are no CIFS users connected to shares on the qtree whose security style you want to change. If there are, you cannot change UNIX security style to mixed or NTFS, and you cannot change NTFS or mixed security style to UNIX.

About this task

You can set the security style of a volume or qtree. Setting the security style of a volume does not affect the security style of the qtrees contained by that volume. It only affects the security style for the files that are not contained by any qtree (these files are said to be in qtree 0).

Step

1. Enter the following command:

   qtree security *path* {unix | ntfs | mixed}

Examples

To change the security style of /vol/users/docs to Windows NT, use the following command:

qtree security /vol/users/docs ntfs

To change the security style of the root directory of the users volume to mixed (so that outside a qtree in the volume, one file can have NTFS security and another file can have UNIX security) use the following command:

qtree security /vol/users mixed

After you finish

If you have quotas in effect on the qtree or volume whose security style you just changed, reinitialize quotas on the volume containing this qtree.

Related concepts

How changing the security style of a qtree affects user quotas on page 316
How changing the security style of a qtree affects user quotas on page 316
How security styles affect access to your data on page 162

Related tasks

Reinitializing quotas on page 338

About quotas

Quotas provide a way to restrict or track the disk space and number of files used by a user, group, or qtree. You specify quotas using the `/etc/quotas` file. Quotas are applied to a specific volume or qtree.

Next topics

Why you use quotas

You can use quotas to limit resource usage, to provide notification when resource usage reaches specific levels, or simply to track resource usage.

You specify a quota for the following reasons:

- To limit the amount of disk space or the number of files that can be used by a user or group, or that can be contained by a qtree
- To track the amount of disk space or the number of files used by a user, group, or qtree, without imposing a limit
- To warn users when their disk usage or file usage is high

Overview of the quota process

Quotas can cause Data ONTAP to send a notification (soft quota) or to prevent a write operation from succeeding (hard quota) when quotas are exceeded.

When Data ONTAP receives a request to write to a volume, it checks to see whether quotas are activated for that volume. If so, Data ONTAP determines whether any quota for that volume (and, if the write is to a qtree, for that qtree) would be exceeded by performing the write operation. If any hard quota would be exceeded, the write operation fails, and a quota notification is sent. If any soft quota would be exceeded, the write operation succeeds, and a quota notification is sent.

About quota notifications

Quota notifications go to the console and the `/etc/messages` file. You can also configure SNMP traps to be triggered when a quota is exceeded.

When an attempt is made to exceed a hard quota, a console message is generated and an SNMP trap is triggered. The console messages and SNMP traps are sent only once every 60 minutes to avoid flooding the message file and console with redundant messages.

When a soft quota or threshold is exceeded, a console message is generated and an SNMP trap is triggered. For the soft quota, a console message is generated and an SNMP trap is triggered when the soft quota is no longer exceeded. For thresholds, there is no notification when the threshold is no longer exceeded.

SNMP traps can be used to arrange alternative methods of notification, such as e-mail. You can find details on SNMP traps in the `/etc/mib/netapp.mib` file.

Note: The syslog messages generated when a tree quota is reached contain qtree ID numbers rather than qtree names. You can correlate qtree names to ID numbers by using the `qtree status -i` command.

Quota targets and types

Quotas have a type: they can be either user, group, or tree. Quota targets specify the user, group, or qtree for which the quota limits are applied.

The following table lists the kinds of quota targets, what types of quotas each quota target is associated with, and how each quota target is represented.

Quota target	Quota type	How target is represented	Notes
user	user quota	UNIX user name UNIX UID A file or directory whose UID matches the user Windows user name in pre-Windows 2000 format Windows SID A file or directory with an ACL owned by the user's SID	User quotas can be applied for a specific volume or qtree.
group	group quota	UNIX group name UNIX GID A file or directory whose GID matches the group	Group quotas can be applied for a specific volume or qtree. **Note:** Data ONTAP does not apply group quotas based on Windows IDs.
qtree	tree quota	path name to the qtree For example, `vol/vol1/qtree2`	Tree quotas are applied to a particular volume and do not affect qtrees in other volumes.
*	user group tree	The asterisk character (*)	A quota target of * denotes a *default quota*. For default quotas, the quota type is determined by the value of the type field.

Special kinds of quotas

You use default, explicit, derived and tracking quotas to manage disk usage in the most efficient manner.

Next topics

How default quotas work on page 304

How default quotas work

You can use default quotas to apply a quota to all instances of a given quota type. For example, a default user quota affects all users on the system for the specified volume. In addition, default quotas enable you to modify your quotas easily.

You can use default quotas to automatically apply a limit to a large set of quota targets without having to create separate quotas for each target. For example, if you want to limit most users to 10 GB of disk space, you can specify a default user quota of 10 GB of disk space instead of creating a quota for each user. If you have specific users for whom you want to apply a different limit, you can create explicit quotas for those users. (Explicit quotas—quotas with a specific target or list of targets —override default quotas.)

Default quotas can be applied to all three types of quota target (users, groups, and qtrees).

Note: When a default user quota is in effect, Data ONTAP also tracks resource usage for the root user and the BUILTIN\Administrators group. Similarly, when a default group quota is in effect, Data ONTAP tracks resource usage for the group with GID 0.

Default user quota example

The following quotas file uses a default user quota to apply a 50-MB limit on each user for vol1:

```
#Quota target type           disk  files  thold  sdisk  sfile
#------------  ----           ----  -----  -----  -----  -----
*              user@/vol/vol1  50M
```

If any user on the system enters a command that would cause that user's data to take up more than 50 MB in vol1 (for example, writing to a file from an editor), the command fails.

How you use explicit quotas

You use explicit quotas to specify a quota for a specific quota target, or to override a default quota for a specific target.

An explicit quota specifies a limit for a particular user, group, or qtree. An explicit quota replaces any default quota in place for the same target.

Explicit quotas only affect default quotas at the same level (volume or qtree). For example, an explicit user quota for a qtree does not affect the default user quota for the volume that contains that qtree. However, the explicit user quota for the qtree overrides (replaces the limits defined by) the default user quota for that qtree.

Examples

The following quotas file contains a default user quota that limits all users in vol1 to 50 MB of space. However, one user, jsmith, is allowed 80 MB of space, because of the explicit quota (shown in bold):

```
#Quota target type            disk  files  thold  sdisk  sfile
#-----------  ----            ----  -----  -----  -----  -----
*             user@/vol/vol1  50M
jsmith        user@/vol/vol1  80M
```

The following quotas entry restricts the specified user, represented by four IDs, to 500MB of disk space and 10,240 files in the vol1 volume:

```
jsmith,corp\jsmith,engineering\”john smith”,
S-1-5-32-544   user@/vol/vol1         500M      10K
```

The following quotas entry restricts the eng1 group to 150 MB of disk space and an unlimited number of files in the /vol/vol2/proj1 qtree:

```
eng1        group@/vol/vol2/proj1  150M
```

The following quotas entry restricts the proj1 qtree in the vol2 volume to 750 MB of disk space and 76,800 files:

```
/vol/vol2/proj1   tree                  750M      75K
```

How derived quotas work

A quota applied as a result of a default quota, rather than an explicit quota (a quota with a specific target), is referred to as a *derived quota*.

Data ONTAP derives the quota information from the default quota and applies it if a write request affects the disk space or number of files used by an instance of the quota target. Derived quotas are applied for all quota target types (users, groups, and qtrees) unless an explicit quota is in effect for that target.

Data ONTAP tracks disk and file usage for quota targets of derived quotas, which means you can change the specifications of these derived quotas by resizing rather than having to perform a full quota reinitialization.

How you use tracking quotas

Tracking quotas generate reports of disk and file usage and do not limit resource usage. When tracking quotas are used, modifying quota values is less disruptive, because you can resize quotas rather than turning them off and back on.

To create a tracking quota, you specify a dash ("-") for the disk and files values. This tells Data ONTAP to monitor disk and files usage for that target and volume, without imposing any limits.

You can also specify a *default tracking quota*, which applies to all instances of the target. Default tracking quotas enable you to track usage for all instances of a quota type (for example, all qtrees or all users). In addition, they enable you use resizing rather than reinitialization when you want quota changes to take effect.

Examples

The following quotas file shows tracking quotas in place for a specific user, group, and qtree:

```
#Quota target   type           disk files thold sdisk sfile
#-----------    ----           ---- ----- ----- ----- -----
kjones          user@/vol/vol1    -   -
eng1            group@/vol/vol1   -   -
proj1           tree@/vol/vol1    -   -
```

The following quotas file contains the three possible default tracking quotas (users, groups, and qtrees):

```
#Quota target   type           disk  files thold sdisk sfile
#-----------    ----           ----  ----- ----- ----- -----
*               user@/vol/vol1    -    -
*               group@/vol/vol1   -    -
*               tree@/vol/vol1    -    -
```

Related concepts

About modifying quotas on page 324

How quotas are applied

Understanding how quotas are applied enables you to configure your quotas to get the limits that you expect.

Whenever an attempt is made to write data to a file in a volume that has quotas enabled, specific quota limits are checked before that write operation is allowed to proceed. If the write operation will exceed any of the quota limits, the operation is prevented and no further limits are checked.

For a file that is in qtree0 (not contained by any user-created qtree), the quota limits are checked in the following order:

1. The user quota for the user that owns the file on the volume
2. The group quota for the group that owns the file on the volume

For a file that is in a user-created qtree, the quota limits are checked in the following order:

1. The tree quota for that qtree

2. The user quota for the user that owns the file on the volume
3. The group quota for the group that owns the file on the volume
4. The user quota for the user that owns the file on the qtree
5. The group quota for the group that owns the file on the qtree

Note: The quota with the smallest limit may not be the one that is exceeded first. For example, if a user quota for volume vol1 specified 100 GB, and the user quota for qtree q2 was 20 GB, the volume limit could be reached first if that user had already written more than 80 GB of data in volume vol1 (but outside of qtree q2).

How quotas work with users and groups

When you specify a user or group as the target of a quota, the limits imposed by that quota are applied to that user or group. However, some special groups and users are handled differently. There are different ways to specify IDs for users, depending on your environment.

Next topics

How you specify UNIX users for quotas on page 307
How you specify Windows users for quotas on page 308
How quotas are applied to the root user on page 309
How quotas work with special Windows groups on page 310
How quotas are applied to users with multiple IDs on page 310
How Data ONTAP determines user IDs in a mixed environment on page 311
How quotas with multiple users work on page 311
How you link UNIX and Windows names for quotas on page 312

Related concepts

How default quotas work on page 304
How you use tracking quotas on page 305

How you specify UNIX users for quotas

You can specify a UNIX user for a quota using one of three formats: the user name, the UID, or a file or directory owned by the user.

To specify a UNIX user for a quota, you can use one of the following formats:

- The user name, as defined in the `/etc/passwd` file or the NIS password map, such as jsmith.

  Note: You cannot use a UNIX user name to specify a quota if that name includes a backslash (\) or an @ sign. This is because Data ONTAP treats names containing these characters as Windows names.

- The UID, such as 20.
- The path of a file or directory owned by that user, so that the file's UID matches the user.

Note:

If you specify a file or directory name, you should choose a file or directory that will last as long as the user account remains on the system.

Specifying a file or directory name for the UID does not cause Data ONTAP to apply a quota to that file or directory.

How you specify Windows users for quotas

You can specify a Windows user for a quota using one of three formats: the Windows name in pre-Windows 2000 format, the SID, or a file or directory owned by the SID of the user.

To specify a Windows user for a quota, you can use one of the following formats:

- The Windows name in pre-Windows 2000 format.
- The security ID (SID), as displayed by Windows in text form, such as S-1-5-32-544.
- The name of a file or directory that has an ACL owned by that user's SID.

Note:

If you specify a file or directory name, you should choose a file or directory that will last as long as the user account remains on the system.

For Data ONTAP to obtain the SID from the ACL, the ACL must be valid.

If the file or directory exists in a UNIX-style qtree, or if the storage system uses UNIX mode for user authentication, Data ONTAP applies the user quota to the user whose *UID*, not SID, matches that of the file or directory.

Specifying a file or directory name to identify a user for a quota does not cause Data ONTAP to apply a quota to that file or directory.

Next topics

How you specify a user name in pre-Windows 2000 format on page 308
How you specify a Windows domain using the QUOTA_TARGET_DOMAIN directive on page 309

How you specify a user name in pre-Windows 2000 format

The pre-Windows 2000 format, for example `engineering\john_smith`, is used by the quotas file for specifying Windows users.

Keep in mind the following rules when creating pre-Windows 2000 format user names:

- The user name must not exceed 20 characters
- The NetBIOS form of the domain name must be used.

How you specify a Windows domain using the QUOTA_TARGET_DOMAIN directive

Using the QUOTA_TARGET_DOMAIN directive in the quotas file enables you to specify the domain name only once for a group of Windows users.

The QUOTA_TARGET_DOMAIN directive takes an optional argument. This string, followed by a backslash (\), is prepended to the name specified in the quota entry. Data ONTAP stops adding the domain name when it reaches the end of the quotas file or another QUOTA_TARGET_DOMAIN directive.

Example

The following example illustrates the use of the QUOTA_TARGET_DOMAIN directive:

```
QUOTA_TARGET_DOMAIN corp
roberts    user@/vol/vol2         900M      30K
smith      user@/vol/vol2         900M      30K
QUOTA_TARGET_DOMAIN engineering
daly       user@/vol/vol2         900M      30K
thomas     user@/vol/vol2         900M      30K
QUOTA_TARGET_DOMAIN
stevens    user@/vol/vol2         900M      30K
```

The string corp\ is added as a prefix to the user names of the first two entries. The string engineering\ is added as a prefix to the user names of the third and fourth entries. The last entry is unaffected by the QUOTA_TARGET_DOMAIN entry because the entry contains no argument.

The following entries produce the same effects:

```
corp\roberts       user@/vol/vol2          900M      30K
corp\smith         user@/vol/vol2          900M      30K
engineering\daly   user@/vol/vol2          900M      30K
engineering\thomas user@/vol/vol2          900M      30K
stevens            user@/vol/vol2          900M      30K
```

How quotas are applied to the root user

The root user (UID=0) on UNIX clients is subject to tree quotas, but not user quotas or group quotas. This allows the root user to take actions on behalf of other users that would otherwise be prevented by a quota.

When root carries out a file or directory ownership change or other operation (such as the UNIX `chown` command) on behalf of a user with less privileges, Data ONTAP checks the quotas based on the new owner but does not report errors or stop the operation, even if the hard quota restrictions of the new owner are exceeded. This can be useful when an administrative action, such as recovering lost data, results in temporarily exceeding quotas.

Note: After the ownership transfer is carried out, however, a client system will report a disk space error if the user attempts to allocate more disk space while the quota is still exceeded.

How quotas work with special Windows groups

Quotas are applied to the Everyone group and the BUILTIN\Administrators group differently than to other Windows groups.

The following list describes what happens if the quota target is a special Windows group ID:

- If the quota target is the Everyone group, a file whose ACL shows that the owner is Everyone is counted under the SID for Everyone.
- If the quota target is BUILTIN\Administrators, the entry is considered a user quota, for tracking only. You cannot impose restrictions on BUILTIN\Administrators.

  If a member of BUILTIN\Administrators creates a file, the file is owned by BUILTIN\Administrators and is counted under the SID for BUILTIN\Administrators, not the user's personal SID.

Note: Data ONTAP does not support group quotas based on Windows group IDs. If you specify a Windows group ID as the quota target, the quota is considered to be a user quota.

How quotas are applied to users with multiple IDs

A user can be represented by multiple IDs. You can set up a single user quota for such a user by specifying a list of IDs as the quota target. A file owned by any of these IDs is subject to the restriction of the user quota.

Suppose a user has the UNIX UID 20 and the Windows IDs corp\john_smith and engineering\jsmith. For this user, you can specify a quota where the quota target is a list of the UID and Windows IDs. When this user writes to the storage system, the specified quota applies, regardless of whether the write originates from UID 20, corp\john_smith, or engineering\jsmith.

Note: Separate quota file entries are considered separate targets, even if the IDs belong to the same user.

For example, for the same user you can specify one quota that limits UID 20 to 1 GB of disk space and another quota that limits corp\john_smith to 2 GB of disk space, even though both IDs represent the same user. Data ONTAP applies quotas to UID 20 and corp\john_smith separately.

In this case, no limits are applied to engineering\jsmith, even though limits are applied to the other IDs used by the same user.

How Data ONTAP determines user IDs in a mixed environment

If you have users accessing your Data ONTAP storage from both Windows and UNIX clients, then both Windows and UNIX security are used to determine file ownership. Several factors determine whether Data ONTAP uses a UNIX or Windows ID when applying user quotas.

If the security style of the qtree or volume that contains the file is only NTFS or only UNIX, then the security style determines the type of ID used when applying user quotas. For qtrees with the mixed security style, the type of ID used is determined by whether the file has an ACL.

The following table summarizes what type of ID is used:

Security Style	ACL	No ACL
UNIX	UNIX ID	UNIX ID
Mixed	Windows ID	UNIX ID
NTFS	Windows ID	Windows ID

Note: If a file is owned by a user of the other type, and no mapping to the determined type exists, then Data ONTAP uses the default user ID for the determined type as defined in the following options:

- `wafl.default_nt_user`
- `wafl.default_unix_user`

For example, suppose the winfile file is in a qtree with the UNIX security style, and it is owned by Windows user corp\bob. If there is no mapping between corp\bob and a UNIX user id in the quotas file, the winfile file is charged against the user defined by the `wafl.default_nt_user` option.

Related concepts

How security styles affect access to your data on page 162

How quotas with multiple users work

When you put multiple users in the same quota target, the quota limits defined by that quota are not applied to each individual user; in this case, the quota limits are *shared* among all users listed in the quota target.

Note: You can combine multiple single quota user targets into one line by using the `quota resize` command. However, if you want to remove users from a quota target with multiple users, or add users to a target that already has multiple users, you must restart quotas before the change takes effect.

Example of more than one user in a quotas file entry

In the following example, there are two users listed in the quota entry:

```
#Quota       target type    disk files thold sdisk sfile
#----------  ----           ---- ----- ----- ----- -----
jsmith,chen  user@/vol/vol1 80M
```

The two users can use up to 80 MB of space combined. If one uses 75 MB, then the other one can use only 5 MB.

How you link UNIX and Windows names for quotas

In a mixed environment, users can log in as either Windows users or UNIX users. You can configure quotas to recognize that a user's UNIX id and Windows ID represent the same user.

Next topics

How you map names using the same quotas file entry on page 312
How you map names using the QUOTA_PERFORM_USER_MAPPING directive on page 312

How you map names using the same quotas file entry

You can map Windows to UNIX names by putting them together in the same entry in the quotas file. However, this requires a quotas file entry for every user.

Example

The following quotas file entry links the Windows ID corp\jroberts to the UNIX ID roberts for quotas:

```
roberts,corp\jroberts user@/vol/vol2 900M 30K
```

How you map names using the QUOTA_PERFORM_USER_MAPPING directive

If you have configured the system's `/etc/usermap.cfg` file with a one-to-one correspondence between UNIX names and Windows names, the QUOTA_PERFORM_USER_MAPPING directive in the quotas file automatically links the names. You do not have to add a separate entry for each user.

When you use this directive, Data ONTAP consults the `usermap.cfg` file to map the user names. When a UNIX and Windows name are mapped together, they are treated as the same person for determining quota usage.

For more information about the `usermap.cfg` file, see the *File Access and Protocols Management Guide*.

Note: This directive requires a one-to-one correspondence between Windows names and UNIX names. If a name maps to more than one name in the `usermap.cfg` file, there are duplicate entries in the quotas file and unpredictable results.

Note: If you are using this directive, when you make changes to the `usermap.cfg` file, you must turn quotas off and back on before your changes will take effect.

Example

The following example illustrates the use of the QUOTA_PERFORM_USER_MAPPING directive:

```
QUOTA_PERFORM_USER_MAPPING ON
roberts        user@/vol/vol2          900M      30K
corp\stevens   user@/vol/vol2          900M      30K
QUOTA_PERFORM_USER_MAPPING OFF
```

If the `usermap.cfg` file maps roberts to corp\jroberts, the first quota entry applies to the user whose UNIX name is roberts and whose Windows name is corp\jroberts. A file owned by either user name is subject to the restriction of this quota entry.

If the `usermap.cfg` file maps corp\stevens to cws, the second quota entry applies to the user whose Windows name is corp\stevens and whose UNIX name is cws. A file owned by either user name is subject to the restriction of this quota entry.

The effect of this example could also be achieved with multiple user names in a single quotas file entry, as in the following example:

```
roberts,corp\jroberts        user@/vol/vol2          900M
30K
corp\stevens,cws             user@/vol/vol2          900M
30K
```

About using wildcard entries in the usermap.cfg file

The use of wildcard entries in the `/etc/usermap.cfg` file causes ambiguity because all trusted domains are searched in an unspecified order for a match. To prevent this problem, you should specify the order in which Data ONTAP searches domains by using the `cifs.search_domains` option.

Unexpected results might occur if your `usermap.cfg` file contains the following entry:

```
*\*     *
```

If you use the QUOTA_PERFORM_USER_MAPPING directive in your quotas file with this wildcard entry in the `usermap.cfg` file, Data ONTAP tries to find users in one of the trusted domains. However, because Data ONTAP searches domains in an unspecified order, the results of this search can be unpredictable.

To address this issue, you can specify the order that Data ONTAP searches domain by using the `cifs.search_domains` option.

How quotas work with qtrees

You can create quotas with a qtree as their target; these quotas are called *tree quotas*. You can also create user and group quotas for a specific qtree. In addition, quotas for a volume are sometimes inherited by the qtrees contained by that volume.

Next topics

How tree quotas work on page 314
How user and group quotas work with qtrees on page 314
How default user quotas on a volume affect quotas for the qtrees in that volume on page 315

How tree quotas work

You can create a quota with a qtree as its target to limit how large the target qtree can become. These quotas are also called *tree quotas*.

When you apply a quota to a qtree, the result is similar to a disk partition, except that you can change the qtree's maximum size at any time by changing the quota. When applying a tree quota, Data ONTAP limits the disk space and number of files in the qtree, regardless of their owners. No users, including root and members of the BUILTIN\Administrators group, can write to the qtree if the write operation causes the tree quota to be exceeded.

Note: The size of the quota does not guarantee any specific amount of available space. The size of the quota can be larger than the amount of free space available to the qtree. You can use the `df` command to determine the true amount of available space in the qtree.

How user and group quotas work with qtrees

Tree quotas limit the overall size of the qtree. To prevent individual users or groups from consuming the entire qtree, you specify a user or group quota for that qtree.

Example user quota in a qtree

Suppose you have the following quotas file:

```
#Quota target type            disk files thold sdisk sfile
#------------  ----            ---- ----- ----- ----- -----
*              user@/vol/vol1  50M   -     45M
jsmith         user@/vol/vol1  80M   -     75M
```

It comes to your attention that a certain user, kjones, is taking up too much space in a critical qtree, qt1, which resides in vol2. You can restrict this user's space by adding the following line to the quotas file:

```
kjones        user@/vol/vol2/qt1  20M   -    15M
```

How default user quotas on a volume affect quotas for the qtrees in that volume

If a default user quota is defined for a volume, a default user quota is automatically created for every qtree contained by that volume for which a tree quota exists.

The automatically created default user quotas on the qtrees have the same limits as the default user quota you created for the volume, and they are each displayed on their own line in the quota report.

An explicit user quota for a qtree overrides (replaces the limits applied by) the automatically created default user quota, just as it would for a default user quota on that qtree that was created by an administrator.

How qtree changes affect quotas

When you delete, rename, or change the security style of a qtree, the quotas applied by Data ONTAP might change, depending on the current quotas being applied.

Next topics

How deleting a qtree affects tree quotas on page 315
How renaming a qtree affects quotas on page 315
How changing the security style of a qtree affects user quotas on page 316

How deleting a qtree affects tree quotas

When you delete a qtree, all quotas applicable to that qtree, whether they are explicit or derived, are no longer applied by Data ONTAP.

If you create a new qtree with the same name as the one you deleted, the quotas previously applied to the deleted qtree are not applied automatically to the new qtree until you reinitialize quotas. If a default tree quota exists, Data ONTAP creates new derived quotas for the new qtree.

If you don't create a new qtree with the same name as the one you deleted, you can delete the quotas that applied to that qtree to avoid getting errors when you reinitialize quotas.

How renaming a qtree affects quotas

When you rename a qtree, its ID does not change. As a result, all quotas applicable to the qtree continue to be applicable, without reinitializing quotas. However, before you reinitialize quotas, you must update the quota with the new qtree name to ensure that the quota continues to be applied for that qtree.

How changing the security style of a qtree affects user quotas

ACLs apply in qtrees using NTFS or mixed security style, but not in qtrees using UNIX security style. Therefore, changing the security style of a qtree might affect how quotas are calculated. You should always reinitialize quotas after you change the security style of a qtree.

If you change a qtree's security style from NTFS or mixed to UNIX, any ACLs on files in that qtree are ignored as a result, and file usage is charged against UNIX user IDs.

If you change a qtree's security style from UNIX to either mixed or NTFS, previously hidden ACLs become visible, any ACLs that were ignored become effective again, and the NFS user information is ignored.

Note: If no ACL existed before, the NFS information continues to be used in the quota calculation.

Attention: To make sure that quota usages for both UNIX and Windows users are properly calculated after you change the security style of a qtree, always reinitialize quotas for the volume containing that qtree.

Example

Suppose NTFS security is in effect on qtree A, and an ACL gives Windows user corp/joe ownership of a 5-MB file. User corp/joe is charged with 5 MB of disk space usage for qtree A.

Now you change the security style of qtree A from NTFS to UNIX. After quotas are reinitialized, Windows user corp/joe is no longer charged for this file; instead, the UNIX user that is mapped to the corp/joe user is charged for the file. If no UNIX user is mapped to corp/joe, then the default UNIX user is charged.

Note: Only UNIX group quotas apply to qtrees. Changing the security style of a qtree, therefore, does not affect the group quotas.

Related concepts

How Data ONTAP determines user IDs in a mixed environment on page 311

Differences among hard, soft, and threshold quotas

Hard quotas (Disk and Files fields) impose a hard limit on system resources; any operation that would result in exceeding the limit fails. The soft quotas (Threshold, Soft Disk, and Soft Files fields) send a warning message when resource usage reaches a certain level, but do not affect data access operations, so you can take appropriate action before the quota is exceeded.

Threshold quotas (quotas specified using the Threshold field) are equivalent to quotas specified using the Soft Disk field, except for how notifications are handled.

How the quotas file works

The quotas file, found in the /etc directory, contains one or more entries specifying limit or tracking quotas for qtrees, groups, and users. The file can contain default (general) and specific entries.

Next topics

The syntax of quota entries on page 317
How Data ONTAP reads the quotas file on page 322
What character encodings are supported by the quotas file on page 322
Sample quotas file on page 322

The syntax of quota entries

The syntax of a quota entry in the quotas file is *quota_target type*[@/vol/*dir/qtree_path*] *disk* [*files*] [*threshold*] [*soft_disk*] [*soft_files*]. Fields are separated by space characters or tabs.

Next topics

How the Quota Target field works on page 317
How the Type field works on page 318
How the Disk field works on page 318
How the Files field works on page 319
How the Threshold field works on page 320
How the Soft Disk field works on page 320
How the Soft Files field works on page 321

How the Quota Target field works

The Quota Target field specifies the name of the qtree, group, or user to which this quota is being applied. An asterisk (*) in this field denotes a default quota, which is applied to all members of the type specified in this entry that do not have an explicit quota.

If you create multiple explicit quotas with the same target, only the first quota with that target is accepted and applied. The others are rejected and do not take effect.

Related concepts

Quota targets and types on page 302

How the Type field works

The Type field specifies the type of entity (qtree, group, or user) to which this quota is being applied. If the type is user or group, this field can optionally restrict the quota to a specific volume, directory, or qtree.

The Type field specifies the quota type, which can be one of the following types:

- User or group quotas, which specify the amount of disk space and the number of files that particular users and groups can own.
- Tree quotas, which specify the amount of disk space and the number of files that particular qtrees can contain.

The following table summarizes the possible values for the Type field, along with examples.

Quota type	Value in the Type field	Sample Type field
User quota in a volume (explicit or default)	`user@/vol/volume`	`user@/vol/vol1`
User quota in a qtree (explicit or default)	`user@/vol/volume/qtree`	`user@/vol/vol0/home`
Group quota in a volume (explicit or default)	`group@/vol/volume`	`group@/vol/vol1`
Group quota in a qtree (explicit or default)	`group@/vol/volume/qtree`	`group@/vol/vol0/home`
Explicit tree quota	`tree`	`tree`
Default tree quota	`tree@/vol/volume`	`tree@/vol/vol0`

How the Disk field works

The Disk field specifies the maximum amount of disk space that the quota target can use. The value in this field represents a hard limit that cannot be exceeded.

The following list describes the rules for specifying a value in this field:

- You cannot leave the Disk field blank.
  The value that follows the Type field is always assigned to the Disk field; thus, for example, Data ONTAP regards the following two quotas file entries as equivalent:

```
#Quota Target          type            disk      files
 /export               tree            75K
 /export               tree                      75K
```

- K means 1,024 bytes, M means 2 to the 20th power or 1024 * 1024 bytes, and G means 2 to the 30th power or 1024 * 1024 * 1024 bytes.

  Note: The Disk field is not case-sensitive. Therefore, you can use K, k, M, m, G, or g.

- The maximum value you can enter in the Disk field is one of the following values (equivalent to 16 TB):

  - 16,383G
  - 16,777,215M
  - 17,179,869,180K

  Note: If you omit the K, M, or G, Data ONTAP assumes a default value of K. The value cannot be specified in decimal notation.

- The value in the Disk field should be a multiple of 4 KB.

  If it is not , the Disk field can appear incorrect in quota reports. This happens because the Disk field is always rounded up to the nearest multiple of 4 KB to match disk space limits, which are translated into 4-KB disk blocks.

- Your quota limit can be larger than the amount of disk space available in the volume.

  In this case, a warning message is printed to the console when quotas are initialized.

- To apply a tracking quota (which tracks disk usage without imposing a limit), type a hyphen (-).

How the Files field works

The Files field specifies the maximum number of files that the quota target can own. This field is optional. The value in this field represents a hard limit that cannot be exceeded.

The following list describes the rules for specifying a value in this field:

- K means 1,024 files, M means 2 to the 20th power or 1024 * 1024 files, and G means 2 to the 30th power or 1024 * 1024 * 1024 files.

  You can omit the K, M, or G. For example, if you type 100, it means that the maximum number of files is 100.

  Note: The Files field is not case-sensitive. Therefore, you can use K, k, M, m, G, or g.

- The maximum value you can enter in the Files field is `4G` or one of the following values:

  - 4,294,967,295
  - 4,194,304K
  - 4,096M

- To apply a tracking quota (which tracks file usage without imposing a limit), type a hyphen (-).

  Note: If the quota target is root, or if you specify 0 as the UID or GID, you *must* type a hyphen.

- A blank in the Files field means there is no restriction on the number of files that the quota target can use.

Note: If you leave the Files field blank, you cannot specify values for the Threshold, Soft Disk, or Soft Files fields.

- The Files field must be on the same line as the Disk field.
  Otherwise, the Files field is ignored.

How the Threshold field works

The Threshold field specifies the disk space threshold. If a write causes the quota target to exceed the threshold, the write still succeeds, but a warning message is logged to the storage system console and an SNMP trap is generated. This field is optional.

The following list describes the rules for specifying a value in this field:

- K means 1,024 bytes, M means 2 to the 20th power or 1024 * 1024 bytes, and G means 2 to the 30th power or 1024 * 1024 * 1024 bytes.

  Note: The Threshold field is not case-sensitive. Therefore, you can use K, k, M, m, G, or g.

- The maximum value you can enter in the Threshold field is one of the following values (equivalent to 16 TB):

  - 16,383G
  - 16,777,215M
  - 17,179,869,180K

  Note: If you omit the K, M, or G, Data ONTAP assumes a default value of K. The value cannot be specified in decimal notation.

- The value in the Threshold field, if any, should be a multiple of 4 KB.
  If it is not, the Threshold field can appear incorrect in quota reports. This happens because the Threshold field is always rounded up to the nearest multiple of 4 KB to match disk space limits, which are translated into 4-KB disk blocks.
- The Threshold field must be on the same line as the Disk field.
  Otherwise, the Threshold field is ignored.
- If you do not want to specify a threshold for the quota target, enter a hyphen (-) in this field or leave it blank.

How the Soft Disk field works

The Soft Disk field specifies the amount of disk space that the quota target can use before a warning is issued. If the quota target exceeds the soft limit, a warning message is logged to the storage system console and an SNMP trap is generated. This field is optional, and works the same way as the Threshold field.

The following list describes the rules for specifying a value in this field:

- K means 1,024 bytes, M means 2 to the 20th power or 1024 * 1024 bytes, and G means 2 to the 30th power or 1024 * 1024 * 1024 bytes.

  Note: The Soft Disk field is not case-sensitive. Therefore, you can use K, k, M, m, G, or g.

- The maximum value you can enter in the Soft Disk field is one of the following values(equivalent to 16 TB):
  - 16,383G
  - 16,777,215M
  - 17,179,869,180K

  Note: If you omit the K, M, or G, Data ONTAP assumes a default value of K. The value cannot be specified in decimal notation.
- The value in the Threshold field, if any, should be a multiple of 4 KB.
  If it is not, the Soft Disk field can appear incorrect in quota reports. This happens because the Soft Disk field is always rounded up to the nearest multiple of 4 KB to match disk space limits, which are translated into 4-KB disk blocks.
- The Soft Disk field must be on the same line as the Disk field.
  Otherwise, the Soft Disk field is ignored.
- If you do not want to specify a soft disk limit for the quota target, enter a hyphen (-) in this field or leave it blank.

How the Soft Files field works

The Soft Files field specifies the number of files that the quota target can use before a warning is issued. If the quota target exceeds the soft limit, a warning message is logged to the storage system console and an SNMP trap is generated. This is an optional field.

The following list describes the rules for specifying a value in the Soft Files field:

- K means 1,024 files, M means 2 to the 20th power or 1024 * 1024 files, and G means 2 to the 30th power or 1024 * 1024 * 1024 files.
  You can omit the K, M, or G. For example, if you type 100, it means that the soft limit on the number of files is 100.

  Note: The Soft Files field is not case-sensitive. Therefore, you can use K, k, M, m, G, or g.
- The maximum value you can enter in the Soft Files field is 4G or one of the following values:
  - 4,294,967,295
  - 4,194,304K
  - 4,096M
- A blank in the Soft Files field means there is no soft quota on the number of files that the quota target can use.
- The Soft Files field must be on the same line as the Disk field.
  Otherwise, the Soft Files field is ignored.

How Data ONTAP reads the quotas file

There are a few simple rules to follow to ensure that Data ONTAP can read your quotas file properly.

An entry in the quotas file can extend to multiple lines. However, the Files, Threshold, Soft Disk, and Soft Files fields must be on the same line as the Disk field; otherwise, they are ignored.

If you do not want to specify a value for a field in the middle of an entry, you can use a dash (-).

Any text after a pound sign (#) is considered a comment.

Entries in the quotas file can be in any order. After Data ONTAP receives a write request, it grants access only if the request meets the requirements specified by all quotas entries.

If you create multiple explicit quotas file entries with the same target, only the first quota with that target is accepted and applied. The others are rejected and do not take effect.

What character encodings are supported by the quotas file

The quotas file supports two types of character encoding: Unicode and root volume UNIX encoding (the language specified for the root volume using the `vol lang` command).

You can edit the quotas file from either a PC or a UNIX workstation. Data ONTAP can detect whether a file was edited and saved by a Unicode-capable editor, such as Notepad. If so, Data ONTAP considers all entries in the file to be in Unicode. Otherwise, Data ONTAP considers the entries to be in the root volume UNIX encoding.

Standard Generalized Markup Language (SGML) entities are allowed only in the root volume UNIX encoding.

Note: If you want to include non-ASCII characters in your quotas file, you must use Unicode or SGML.

Sample quotas file

A short example quotas file, together with explanations, can help you to understand the different types of quota entries and how they affect your quotas.

The following sample quotas file contains both default and explicit quotas:

```
#Quota Target type                disk  files  thold  sdisk  sfile
#------------ ----                 ----  -----  -----  -----  -----
*             user@/vol/vol1       50M   15K
*             group@/vol/vol1      750M  85K
*             tree@/vol/vol1       100M  75K
jdoe          user@/vol/vol1/proj1 100M  75K
msmith        user@/vol/vol1       75M   75K
msmith        user@/vol/vol1/proj1 75M   75K
```

This quotas file has the following effects:

- Any user not otherwise mentioned in this file can use 50 MB of disk space and 15,360 files in the vol1 volume.
- Any group not otherwise mentioned in this file can use 750 MB of disk space and 87,040 files in the vol1 volume.
- Any qtree in the vol1 volume not otherwise mentioned in this file can use 100 MB of disk space and 76,800 files.
- If a qtree is created in the vol1 volume (for example, a qtree named /vol/vol1/proj2), Data ONTAP enforces a derived default user quota and a derived default group quota that have the same effect as the following quota entries:

```
*                   user@/vol/vol1/proj2     50M       15K
*                   group@/vol/vol1/proj2   750M       85K
```

- If a qtree is created in the vol1 volume (for example, a qtree named /vol/vol1/proj2), Data ONTAP tracks the disk space and number of files owned by UID 0 and GID 0 in the /vol/vol1/proj2 qtree. This is due to the following quotas file entry:

```
*                   tree@/vol/vol1           100M       75K
```

- A user named msmith can use 75 MB of disk space and 76,800 files in the vol1 volume because an explicit quota for this user exists in the /etc/quotas file, overriding the default limit of 50 MB of disk space and 15,360 files.
- By giving jdoe and msmith 100 MB and 75 MB explicit quotas for the proj1 qtree, which has a tree quota of 100MB, that qtree becomes oversubscribed. This means that the qtree could run out of space before the user quotas are exhausted.

  Note: Quota oversubscription is supported; however, a warning is printed alerting you to the oversubscription.

About activating or reinitializing quotas

You use the `quota on` command to activate or reinitialize quotas, which causes all quotas for that volume to be recalculated. Knowing how quota initialization works can help you manage your quotas less disruptively.

The following list outlines some facts you should know about activating or reinitializing quotas:

- Changes to quotas do not take effect until quotas are either reinitialized or resized using the `quota resize` command.
- You activate or reinitialize quotas for only one volume at a time.
- Your quotas file does not need to be free of all errors to activate quotas. Invalid entries are reported and skipped. If the quotas file contains any valid entries, quotas are activated.
- Quota reinitialization can take some time, during which storage system data is available, but quotas are not enforced for the specified volume.
- Quota reinitialization is performed in the background by default; other commands can be performed while the reinitialization is proceeding.

Note: Errors or warnings from the reinitialization process could be interspersed with the output from other commands.

- Quota reinitialization can be invoked in the foreground with the `-w` option; this is useful if you are reinitializing from a script.
- Errors and warnings from the reinitialization process are logged to the console as well as to `/etc/messages`.
- Quota activation persists across halts and reboots. You should not activate quotas in the `/etc/rc` file.

Related concepts

When you can use resizing on page 324
When a full quota reinitialization is required on page 326

About modifying quotas

After you make changes to your quotas, you need to tell Data ONTAP to incorporate the changes. There are two ways to do this, depending on the nature of the changes and your existing quotas.

You can tell Data ONTAP to incorporate quota changes in one of the following two ways:

- Resize quotas
  Resizing quotas is faster than a full reinitialization; however, some quota changes might not be reflected.
- Reinitialize quotas
  Performing a full quota reinitialization recalculates all quotas. This process might take some time, but all quotas changes are guaranteed to be reflected after the initialization is complete.

Note: Your storage system functions normally while quotas are being initialized; however, quotas remain deactivated for the specified volume until the initialization is complete.

Next topics

When you can use resizing on page 324
When a full quota reinitialization is required on page 326

When you can use resizing

Because quota resizing is faster than quota initialization, you should use resizing whenever possible. However, resizing only works for certain types of quota changes.

You can use quota resizing for the following types of changes to the quotas file:

- You change an existing quota.
  For example, you change the size of an existing quota.
- You add a quota for a quota target for which a default or default tracking quota exists.

- You delete a quota for which a default or default tracking quota entry is specified.

  Attention: After you have made extensive quotas changes, you should perform a full reinitialization to ensure that all of the changes take effect.

  Note: If you attempt to resize and not all of your quota changes can be incorporated by using a resize operation, Data ONTAP issues a warning.

  You can determine from the quota report whether your storage system is tracking disk usage for a particular user, group, or qtree. If you see a quota in the quota report, it means that the storage system is tracking the disk space and the number of files owned by the quota target.

Example quotas file changes that can be made effective using the quota resize command

Some quotas file changes can be made effective using the `quota resize` command. Consider the following sample quotas file:

```
#Quota Target type             disk  files thold sdisk sfile
#------------ ----             ----  ----- ----- ----- -----
*            user@/vol/vol2    50M   15K
*            group@/vol/vol2  750M   85K
*            tree@/vol/vol2      -     -
jdoe         user@/vol/vol2/  100M   75K
kbuck        user@/vol/vol2/  100M   75K
```

Suppose you make the following changes:

- Increase the number of files for the default user target.
- Add a new user quota for a new user that needs more than the default user quota.
- Delete the kbuck user's explicit quota entry; the kbuck user now needs only the default quota limits.

These changes result in the following quotas file:

```
#Quota Target type             disk  files thold sdisk sfile
#------------ ----             ----  ----- ----- ----- -----
*            user@/vol/vol2    50M   25K
*            group@/vol/vol2  750M   85K
*            tree@/vol/vol2      -     -
jdoe         user@/vol/vol2/  100M   75K
bambi        user@/vol/vol2/  100M   75K
```

All of these changes can be made effective using the `quota resize` command; a full quota reinitialization is not necessary.

Example quotas file changes that cannot be made effective using the quota resize command

Some quotas file changes cannot be made effective using the `quota resize` command. For example, suppose your quotas file did not contain the default tracking tree quota, and you want to add a tree quota to the quotas file, resulting in this quotas file:

```
#Quota Target  type            disk files thold sdisk sfile
#------------  ----            ---- ----- ----- ----- -----
*              user@/vol/vol2   50M   25K
*              group@/vol/vol2 750M   85K
jdoe           user@/vol/vol2/ 100M   75K
bambi          user@/vol/vol2/ 100M   75K
/vol/vol2/proj1 tree           500M  100K
```

In this case, using the `quota resize` command does not cause the newly added entry to be effective, because there is no default entry for tree quotas already in effect. A full quota initialization is required.

Related concepts

How quota reports work on page 327

When a full quota reinitialization is required

Although resizing quotas is faster, you must do a full quota reinitialization if you make certain or extensive changes to your quotas.

A full quota reinitialization is necessary in the following circumstances:

- You create a quota for a target that has not previously had a quota
- You change user mapping in the `usermap.cfg` file and you use the QUOTA_PERFORM_USER_MAPPING entry in the quotas file.
- You change the security style of a qtree from UNIX to either mixed or NTFS.
- You change the security style of a qtree from mixed or NTFS to UNIX.
- You make extensive changes to your quotas.

Related concepts

How you map names using the QUOTA_PERFORM_USER_MAPPING directive on page 312

How quotas work with vFiler units

When you create vFiler units, or move resources between vFiler units, quotas for the containing volume are deactivated.

After you create vFiler units or reassign resources between vFiler units, you should ensure that quotas are on.

Note: If having quotas briefly deactivated is disruptive to any applications, you should disable those applications before assigning resources to vFiler units.

How quota reports work

Quota reports enable you to see what quotas Data ONTAP is applying. You can change the format of the quota report and how user IDs are displayed using the options for the `quota report` command.

Next topics

What fields quota reports contain on page 327
How quota report options affect quota reports on page 328
How the ID field is displayed in quota reports on page 330
How you can use the quota report to see what quotas are in effect on page 330

What fields quota reports contain

Some quota report fields are always displayed; others depend on what options you use for the `quota report` command.

The following table lists the headings that can appear in quota reports, with a description and the option required to display that heading if needed.

Quota report heading	Description
Type	Quota type: user, group, or tree.
ID	User ID, UNIX group name, qtree name. If the quota is a default quota, the value in this field is an asterisk.
Volume	Volume to which the quota is applied.
Tree	Qtree to which the quota is applied.
K-Bytes Used	Current amount of disk space used by the quota target. If the quota is a default quota, the value in this field is 0.

Quota report heading	Description
Limit	Maximum amount of disk space that can be used by the quota target (the value in the Disk field of the quotas file).
S-Limit	Maximum amount of disk space that can be used by the quota target before a warning is issued (the value in the Soft Disk field of the quotas file). This column is displayed only when you use the `-s` option for the `quota report` command.
T-hold	Disk space threshold (the value in the Threshold field of the quotas file). This column is displayed only when you use the `-t` option for the `quota report` command.
Files Used	Current number of files used by the quota target. If the quota is a default quota, the value in this field is 0.
Limit	Maximum number of files allowed for the quota target (the value in the File field of the quotas file).
S-Limit	Maximum number of files that can be used by the quota target before a warning is issued (the value in the Soft Files field of the quotas file). This column is displayed only when you use the `-s` option for the `quota report` command.
VFiler	Displays the name of the vFiler unit for this quota entry. This column is displayed only when you use the `-v` option for the `quota report` command. This option is available only on storage systems that have MultiStore licensed.
Quota Specifier	For an explicit quota, this field shows how the quota target is specified in the quotas file. For a derived quota, the field is blank.

How quota report options affect quota reports

What options you use for the `quota report` command affect how the report is formatted and how user IDs are displayed.

The following table lists the options for the `quota report` command with their results on the quota report:

Option	Result
none	Generates the default quota report. The ID field displays one of the IDs using the following formats: • For a Windows name, the first seven characters of the user name with a preceding backslash are displayed. The domain name is omitted. • For a SID, the last eight characters are displayed. The Quota Specifier field displays an ID that matches the one in the ID field, using the same format as the /etc/quotas file entry.
`-q`	Displays the quota target's UNIX UID, GID or Windows SID in the following formats: • UNIX UIDs and GIDs are displayed as numbers. • Windows SIDs are displayed as text. **Note:** Data ONTAP does not perform a lookup of the name associated with the target ID.
`-s`	The soft limit (S-limit) columns are included.
`-t`	The threshold (T-hold) column is included.
`-v`	The vFiler column is included.
`-u`	Displays multiple IDs for your quota targets. The ID field displays all the IDs listed in the quota target of a user quota in the following format: • On the first line, the format is the same as the default format. • Each additional name in the quota target is displayed, in its entirety, on a separate line. The Quota Specifier field displays the list of IDs specified in the quota target. **Note:** You cannot combine the `-u` and `-x` options.
`-x`	Displays all the quota target's IDs on the first line of that quota target's entry, as a comma separated list. **Note:** You cannot combine the `-u` and `-x` options. The threshold column is included.

How the ID field is displayed in quota reports

Usually, the ID field of the quota report displays a user name instead of a UID or SID. However, there are some exceptions to this rule.

The ID field does *not* display a user name in the following circumstances:

- For a quota with a UNIX user as the target, the ID field shows the UID instead of a name if either of the following conditions applies:
  - No user name for the UID is found in the password database.
  - You specifically request the UID by including the `-q` option for the `quota reports` command.
- For a quota with a Windows user as the target, the ID field shows the SID instead of a name if either of the following conditions applies:
  - The SID is specified as a quota target and the SID no longer corresponds to a user name.
  - Data ONTAP cannot find an entry for the SID in the SID-to-name map cache and cannot connect to the domain controller to ascertain the user name for the SID when it generates the quota report.

How you can use the quota report to see what quotas are in effect

Because of the various ways that quotas interact, more quotas are in effect than just the ones you have explicitly created. To see what quotas are in effect, you can view the quota report.

Example with no user quotas specified for the qtree

In this example, there is one qtree, q1, which is contained by the volume vol1. The administrator has created three quotas:

- A default tree quota limit on vol1 of 400 MB
- A default user quota limit on vol1 of 100 MB
- An explicit user quota limit on vol1 of 200 MB for the user jsmith

The quotas file for these quotas looks similar to the following excerpt:

```
#Quota target type             disk files  thold sdisk  sfile
#------------  ----             ---  -----  ----  -----  -----
*              tree@/vol/vol1   400M
*              user@/vol/vol1   100M
jsmith         user@/vol/vol1   200M
```

The quota report for these quotas looks similar to the following excerpt:

```
sys1> quota report
                                K-Bytes             Files
Type       ID    Volume    Tree  Used      Limit     Used    Limit   Quota Specifier
----- -------- -------- -------- --------- --------- ------- ------- ---------------
tree         *     vol1        -         0    409600       0       - *
```

```
user         *     vol1        -         0    102400       0       - *
user     jsmith    vol1        -       112    204800       7       - jsmith
tree         1     vol1       q1         0    409600       6       - /vol/vol1/q1
user         *     vol1       q1         0    102400       0       -
user     jsmith    vol1       q1         0    102400       5       -
user       root    vol1       q1         0         -       1       -
user       root    vol1        -         0         -       8       -
```

The first three lines of the quota report display the three quotas specified by the administrator.

The last two lines display the tracking quotas that are automatically created for the root user whenever a default user is specified for a UNIX or mixed-style qtree.

The fourth line displays the tree quota that is derived from the default tree quota for every qtree in vol1 (in this example, only q1).

The fifth line displays the default user quota that is created for the qtree as a result of the existence of the default user quota on the volume and the qtree quota.

The sixth line displays the derived user quota that is created for the qtree as a result of the existence of the explicit user quota on the volume and the default user quota for the qtree (line 5). Note that the limit applied to the user jsmith in the qtree q1 is not determined by the explicit user quota limit (200 MB). This is because the explicit user quota limit is on the volume, so it does not affect limits for the qtree. Instead, the derived user quota limit for the qtree is determined by the default user quota for the qtree (100 MB).

Example with user quotas specified for the qtree

This example is similar to the previous one, except that the administrator has added two quotas on the qtree.

There is still one volume, vol1, and one qtree, q1. The administrator has created the following quotas:

- A default tree quota limit on vol1 of 400 MB
- A default user quota limit on vol1 of 100 MB
- An explicit user quota limit on vol1 for the user jsmith of 200 MB
- A default user quota limit on qtree q1 of 50 MB
- An explicit user quota limit on qtree q1 for the user jsmith of 75 MB

The quotas file for these quotas looks like this:

```
#Quota target type             disk files  thold sdisk  sfile
#-----------  ----              ---  -----  ----  -----  -----
*             tree@/vol/vol1    400M
*             user@/vol/vol1    100M
jsmith        user@/vol/vol1    200M
*             user@/vol/vol1/q1 50M
jsmith        user@/vol/vol1/q1 75M
```

The quota report for these quotas looks like this:

```
sys1> quota report
                                K-Bytes             Files
Type       ID    Volume    Tree  Used      Limit     Used    Limit   Quota Specifier
----- -------- -------- -------- --------- --------- ------- ------- ---------------
tree         *     vol1        -         0    409600       0       - *
user         *     vol1        -         0    102400       0       - *
user     jsmith    vol1        -       112    204800       7       - jsmith
user         *     vol1       q1         0     51200       0       - *
user     jsmith    vol1       q1         0     76800       5       - jsmith
tree         1     vol1       q1         0    409600       6       - /vol/vol1/q1
user      root     vol1        -         0         -       2       -
user      root     vol1       q1         0         -       1       -
```

The first five lines of the quota report display the five quotas created by the administrator.

The last two lines display the tracking quotas that are automatically created for the root user whenever a default user is specified for a UNIX or mixed-style qtree.

The sixth line displays the tree quota that is derived from the default tree quota for every qtree in vol1 (in this example, only q1).

Note that for this example, Data ONTAP does not create the default user quota and the derived user quotas, because the administrator specified a default user quota and an explicit user quota for the qtree.

Progressive quota examples

Following through a series of progressive examples can help you to understand how to create your quotas file and read your quota reports.

For the following examples, assume that you have a storage system that has one volume, vol1.

Example 1: default quota

You decide to impose a hard limit of 50 MB for each user in vol1, using the following quotas file:

```
#Quota target type             disk files  thold sdisk  sfile
#------------  ----             ---  -----  ----  -----  -----
*              user@/vol/vol1   50M
```

If any user on the system enters a command that would use more than 50 MB in vol1, the command fails (for example, writing to a file from an editor).

Example 2: default quota override

Suppose that you have received a complaint from an important user, saying that she needs more space in vol1. To give this user more space, you update your quotas file as follows (her username is jsmith):

```
#Quota target type             disk files  thold sdisk  sfile
#-----------  ----             ---  -----  ----  -----  -----
*             user@/vol/vol1   50M
jsmith        user@/vol/vol1   80M
```

Now, jsmith can use up to 80 MB of space on vol1, even though all other users are still limited to 50 MB.

The quota report looks like this:

```
filer1> quota report
                                K-Bytes             Files
Type       ID    Volume    Tree  Used      Limit     Used    Limit   Quota Specifier
----- -------- -------- -------- --------- --------- ------- ------- ---------------
user       *      vol1        -         0     51200       0       -    *
user       jsmith vol1        -     63275     81920      37       -    jsmith
user       root   vol1        -         0         -       1       -
```

Note that an extra quota is shown, for the root user. Default user quotas do not apply to root, so the root user has no space limit on vol1, as shown in the quota report by the dash ("-") in the Limit column for the root user.

Example 3: thresholds

This example sets up a threshold for all users at 45 MB, except for jsmith, who will get a threshold at 75 MB. To set up a user-specific threshold, we change the quotas file to read as follows:

```
#Quota target   type             disk    files   thold
sdisk   sfile
#-----------    ----             ----    -----   -----
-----   -----
*               user@/vol/vol1   50M     -       45M
jsmith          user@/vol/vol1   80M     -       75M
```

Note that it was necessary to add a dash (-) in the Files field as a placeholder because the Threshold field comes after the Files field in the quotas file.

Now the quota report looks like this:

```
filer1> quota report -t
                          K-Bytes                        Files
Type    ID    Volume  Tree  Used    Limit     T-hold    Used    Limit   Quota Specifier
------ ----- -------- ----- ------ --------  -------- ------- -------- ----------------
user    *      vol1    -     0       51200    46080     0        -     *
user    jsmith vol1    -     63280   81920    76800     47       -     jsmith
user    root   vol1    -     0       -        -         51       -
```

Note that the `-t` flag is used to display threshold limits.

Example 4: quotas on qtrees

Suppose that you decide you need to partition some space for two projects. You create two qtrees, named proj1 and proj2, to accommodate those projects within vol1. Creating qtrees does not cause any change for your quotas, because the quotas file only applies quotas to the

volume so far. Users can use as much space in a qtree as they are allotted for the entire volume (provided they did not exceed the limit for the volume by using space in the root or another qtree). In addition, each of the qtrees can grow to consume the entire volume.

You decide that you want to make sure that neither qtree grows to more than 20 GB. Your quotas file now looks like this:

```
#Quota target   type            disk files thold sdisk   sfile
#------------   ----            ---- ----- ----- -----   -----
*               user@/vol/vol1  50M    -     45M
jsmith          user@/vol/vol1  80M    -     75M
*               tree@/vol/vol1  20G
```

Note that the correct type is *tree*, not qtree.

Now your quota report looks like this:

```
filer1> quota report -t
                          K-Bytes                        Files
Type   ID    Volume    Tree  Used      Limit     T-hold   Used    Limit  Quota Specifier
----- ------ -------- ------ --------- --------- -------- ------- ----- -------------
user   *      vol1      -          0     51200    46080     0       -  *
user   jsmith vol1      -       63280     81920    76800    55       -  jsmith
tree   *      vol1      -          0   20971520        -     0       -  *
tree   1      vol1      proj1      0   20971520        -     1       - /vol/vol1/proj1
user   *      vol1      proj1      0      51200    46080     0       -
user   root   vol1      proj1      0          -        -     1       -
tree   2      vol1      proj2      0   20971520        -     1       - /vol/vol1/proj2
user   *      vol1      proj2      0      51200    46080     0       -
user   root   vol1      proj2      0          -        -     1       -
user   root   vol1      -          0          -        -     3       -
```

Several new lines have appeared. The first new line is exactly what you added to the quotas file:

```
tree * vol1 - 0 20971520 - 0 - *
```

The next line shows what is called a *derived quota*. You did not add this quota directly. It is derived from the default tree quota that you just added. This new line means that a quota of 20 GB is being applied to the proj1 qtree:

```
tree 1 vol1 proj1 0 20971520 - 1 - /vol/vol1/proj1
```

The next line shows another derived quota. This quota is derived from the default user quota you added in an earlier example. Default user quotas on a volume are automatically inherited for all qtrees contained by that volume, if quotas are enabled for qtrees. When you added the first qtree quota, you enabled quotas on qtrees, so this derived quota was created:

```
user * vol1 proj1 0 51200 46080 0 -
```

The rest of the new lines are for the root user and for the other qtree.

Example 5: user quota on a qtree

You decide to limit users to less space in the proj1 qtree than they get in the volume as a whole. You want to keep them from using any more than 10 MB in the proj1 qtree. To do so, you update the quotas file as follows:

```
#Quota target   type                disk    files   thold   sdisk   sfile
#------------   ----                ----    -----   -----   -----   -----
*               user@/vol/vol1      50M      -       45M
jsmith             user@/vol/vol1      80m      -       75M
*               tree@/vol/vol1      20G
*               user@/vol/vol1/proj1 10M
```

Now a quota report looks like this:

```
filer1> quota report
                                K-Bytes             Files
Type       ID    Volume    Tree  Used      Limit     Used    Limit   Quota Specifier
----- -------- -------- -------- --------- --------- ------- ------- ---------------
user          *  vol1        -         0     51200       0       -     *
user     jsmith  vol1        -         0     81920      57       -     jsmith
tree          *  vol1        -         0  20971520       0       -     *
user          *  vol1    proj1         0     10240       0       -     *
tree          1  vol1    proj1         0  20971520       1       -     /vol/vol1/proj1
tree          2  vol1    proj2         0  20971520       1       -     /vol/vol1/proj2
user          *  vol1    proj2         0     51200       0       -
user       root  vol1    proj2         0         -       1       -
user       root  vol1        -         0         -       3       -
user       root  vol1    proj1         0         -       1       -
```

The new report entry that appears as a result of the line you added is this one:

```
user * vol1 proj1 0 10240 0 - *
```

However, now your phone is ringing. It's jsmith again, complaining that her quota has been decreased. You ask where she is trying to put data, and she says "in proj1." She is being prevented from writing more data to the proj1 qtree because the quota you created to override the default user quota (to give her more space) was on the volume. But now that you have added a default user quota on the proj1 qtree, that quota is being applied and limiting all users' space in that qtree, including jsmith. You must add a new line to the quotas file overriding the qtree default quota to give her more space in the proj1 qtree:

```
jsmith user@/vol/vol1/proj1 80M
```

This adds the following line to your quota report:

```
Type       ID    Volume    Tree  Used      Limit     Used    Limit   Quota Specifier
----- -------- -------- -------- --------- --------- ------- ------- ---------------
user    jsmith    vol1     proj1   57864     81920       57    -       jsmith
```

Related concepts

How default quotas work on page 304
How derived quotas work on page 305
How you use explicit quotas on page 304
How the quotas file works on page 317
How quota reports work on page 327
About quotas on page 301

Managing quotas

You create, delete, and modify quotas as your users and their storage requirements and limitations change. You can also manage how quota messages are logged, and view quota reports, which help you understand what quotas Data ONTAP is applying.

Next topics

Activating quotas

You activate quotas to turn quotas on and read the quotas file. You activate quotas using the `quota on` command, for one volume at a time.

Before you begin

If the quotas file contains user quotas that use Windows IDs as targets, CIFS must be running when you activate quotas.

Step

1. Enter the following command:

   `quota on [-w] vol_name`

   The `-w` option causes the command to return only after the entire quotas file has been scanned (synchronous mode). This is useful when activating quotas from a script.

   Example
   The following example activates quotas on a volume named vol2:

   `quota on vol2`

Quota reinitialization is started for the specified volume. Quota reinitialization can take some time, during which storage system data is available, but quotas are not enforced for the specified volume.

Result

When quota initialization is complete, quotas are on for the specified volume. This procedure does not modify or initialize quotas for any other volume.

After you finish

If a quota initialization is still running when the storage system is upgraded, Data ONTAP terminates the quota initialization, which must be manually restarted from the beginning. For this reason, you should allow any running quota initialization to complete before upgrading your storage system.

Related concepts

About activating or reinitializing quotas on page 323
About modifying quotas on page 324
About quotas on page 301

Reinitializing quotas

You reinitialize quotas by using the `quota off` command followed by the `quota on` command. This causes Data ONTAP to reread the quotas file. Reinitializing quotas takes time. In some cases resizing is more efficient.

Before you begin

If the quotas file contains user quotas that use Windows IDs as targets, CIFS must be running when you reinitialize quotas.

About this task

Depending on how many quotas you have and the size of the file system, quota reinitialization can take some time. During quota reinitialization, data access is not affected. However, quotas are not enforced until reinitialization completes.

Steps

1. If quotas are already activated for the volume on which you want to reinitialize quotas, enter the following command:

   `quota off` *`vol_name`*

   Quotas are turned off for the specified volume.

2. Enter the following command:

   `quota on [-w] vol_name`

   The `-w` option causes the command to return only after the entire quotas file has been scanned (synchronous mode). This is useful when activating quotas from a script.

   Quota reinitialization is started for the specified volume. Quota reinitialization can take some time, during which storage system data is available, but quotas are not enforced for the specified volume.

Result

When quota initialization is complete, quotas are back on for the specified volume.

Note: Quotas are not affected for any volume other than the volume specified in the `quota on` command.

Related concepts

About activating or reinitializing quotas on page 323
About modifying quotas on page 324
About quotas on page 301

Deactivating quotas

You use the `quota off` command to deactivate quotas for a specific volume.

About this task

If a quota initialization is almost complete, the `quota off` command can fail. If this happens, retry the command after a minute or two.

Canceling quota initialization

If you started a quota initialization and you now want to cancel it, you can use the `quota off` command.

About this task

If a quota initialization is almost complete, the `quota off` command can fail. If this happens, the `quota on` command should finish shortly.

Resizing quotas

You use the `quota resize` command to cause Data ONTAP to reread the quotas file for the specified volume. Resizing only works for certain types of changes to the quotas file. For other changes, you need to reinitialize quotas.

Related concepts

When you can use resizing on page 324
About quotas on page 301

Deleting quotas

You can remove quota restrictions for a quota target in two ways: by changing the quotas file entry so that there is no restriction on resource use for that quota target, or by deleting the quotas file entry for that quota target.

Next topics

Deleting a quota by removing resource restrictions on page 340
Deleting a quota by removing the quotas file entry on page 341

Deleting a quota by removing resource restrictions

You can remove a quota for a specific target by removing the resource restrictions for that target. This is equivalent to changing that quota entry to a tracking quota.

Steps

1. Open the quotas file with the editor of your choice and edit the quotas file entry for the specified target so that the quota entry becomes a tracking quota.

   Example
   Suppose your quotas file contained the following entry for the jdoe user:

   ```
   jdoe            user@/vol/vol2/       100M     75K
   ```

   To remove the restrictions for jdoe, you edit the entry as follows:

   ```
   jdoe            user@/vol/vol2/         -        -
   ```

2. Save and close the quotas file.

   The quotas file is updated but the change is not yet effective.

After you finish

Run the `quota resize` command to cause Data ONTAP to reread the quotas file; this will cause your change to become effective.

Related concepts

About modifying quotas on page 324

Deleting a quota by removing the quotas file entry

You can remove a quota for a specific target by removing the quotas file entry for that target. Depending on what other quotas you have set up, you then need to resize or reinitialize quotas.

Steps

1. Open the quotas file with the editor of your choice and remove the entry for the quota you want to delete.

   Note: If the change is temporary, you can disable the quota by prepending the pound sign (#) to the line. This causes Data ONTAP to treat the line as a comment.

2. Save and close the quotas file.

   The quotas file is updated but the change is not yet effective.

After you finish

If you have a default quota or default tracking quota in place for the quota type you modified, you can use the `quota resize` command to cause Data ONTAP to reread the quotas file. Otherwise, reinitialize quotas using the `quota off` and `quota on` commands for the volume for which you modified the quota.

Related concepts

About modifying quotas on page 324

Managing quota message logging

You turn quota message logging on or off, for a single volume or for all volumes, using the `quota logmsg` command. You can also specify a time interval during which quota messages are not logged. This interval defaults to 60 minutes.

About this task

For more information about the `quota logmsg` command, see the na_quota(1) man page.

Displaying a quota report

You display a quota report using the `quota report` command. You can display a quota report for all quotas or for a specific file, directory, qtree or volume by specifying a pathname.

Step

1. To display a quota report, enter the following command:

   `quota report [path]`

   You can display a quota report for all quotas or for a specific file, directory, qtree or volume by specifying a path.

   You can control the format and fields displayed using the `quota report` command options. For more information on the available options, see the na_quota(1) man page.

Related concepts

How quota reports work on page 327
About quotas on page 301

Using the quota report to determine which quotas limit writes to a specific file

You can use the `quota report` command with a specific file path to determine which quota limits affect whether a write to that file will be allowed. This can help you understand which quota is preventing a write operation.

Step

1. To determine which quota limits affect whether a write to a file will be allowed, enter the following command:

   `quota report` ***`filepath`***

Example

The following example shows the command and output to determine what quotas are in effect for writes to the file f4.txt, which resides in the qtree q1 in the volume vol1:

```
sys1> quota report /vol/vol1/q1/f4.txt
                                K-Bytes             Files
Type       ID    Volume    Tree  Used      Limit     Used    Limit   Quota Specifier
----- -------- -------- -------- --------- --------- ------- ------- ---------------
user    jsmith     vol1        -       112    204800       7       - jsmith
```

```
user    jsmith     vol1       q1         0     76800       5       - jsmith
tree         1     vol1       q1         0    409600       6       - /vol/vol1/q1
```

Storage limits

There are limits for aggregates, FlexVol volumes, traditional volumes, FlexCache volumes, FlexClone volumes, files, and LUNs, qtrees and RAID groups that you should consider when planning your storage architecture.

Limits are listed in the following sections:

Volume limits

Limit	Native storage	Back-end storage arrays	Notes
Aggregates and traditional volumes (combined) Maximum per system	100	100	In an active/active configuration, this limit applies to each node individually, so the overall limit for the pair is doubled.
Array LUNs Minimum size for root volume	N/A	Model-dependent	See the *Gateway Interoperability Matrix*.
Files Maximum size	16 TB	16 TB	
FlexCache volumes Maximum per system	100	100	

Limit	Native storage	Back-end storage arrays	Notes
FlexVol volumes Maximum per system	N3700, N3300, and IBM N series N3400 (2859-A11): 200 All other models: 500	N3700, N3300, and IBM N series N3400 (2859-A11): 200 All other models: 500	In an active/active configuration, these limits apply to each node individually, so the overall limit for the pair is doubled. If you plan to perform a non-disruptive upgrade, the limitation on the number of FlexVol volumes you can have might be smaller than the numbers listed here. For more information, see the *Data ONTAP Upgrade Guide*.
FlexVol volumes Minimum size	20 MB	20 MB	
FlexVol volumes Maximum size	16 TB	16 TB	
FlexVol root volumes Minimum size	Model-dependent	Model-dependent	See table below.
Links (hard) Maximum per parent directory	99,998	99,998	
Qtrees Maximum number per volume	4,995	4,995	
Subdirectories Maximum per parent directory	99,998	99,998	
Traditional volumes Maximum size	16 TB	16 TB	

Limit	Native storage	Back-end storage arrays	Notes
Traditional volumes and aggregates Maximum per system	100	100	In an active/active configuration, this limit applies to each node individually, so the overall limit for the pair is doubled.

Aggregate limits

Limit	Native storage	Back-end storage arrays	Notes
Aggregates and traditional volumes (combined) Maximum per system	100	100	In an active/active configuration, this limit applies to each node individually, so the overall limit for the pair is doubled.
Aggregates Maximum size	16 TB	16 TB	
Aggregates Minimum size	N/A	10 GB	
Array LUNs Maximum per aggregate	N/A	Model-dependent	See the *Gateway Interoperability Matrix*.
RAID groups Maximum per aggregate	150	150	
Traditional volumes Maximum size	16 TB	16 TB	

RAID group limits

Limit	Native storage	Back-end storage arrays	Notes
RAID groups Maximum per system	400	400	
RAID groups Maximum per aggregate	150	150	

RAID group sizes

RAID type	Default size	Maximum size	Minimum size
RAID-DP	ATA/BSAS/SATA: 14 FC/SAS: 16	ATA/BSAS/SATA: 16 FC/SAS: 28	3
RAID4	ATA/BSAS/SATA: 7 FC/SAS: 8	ATA/BSAS/SATA: 7 FC/SAS: 14	2
RAID0	8	14	1

FlexClone file and FlexClone LUN limits

Limit	Native storage	Back-end storage arrays	Notes
Maximum per file or LUN	255	255	If you try to create more than 255 clones, Data ONTAP automatically creates a new physical copy of the parent file or LUN. This limit could be lower for FlexVol volumes that use deduplication.
Maximum simultaneous operations per FlexVol volume	16	16	

Limit	Native storage	Back-end storage arrays	Notes
Maximum simultaneous operations per storage system	500	500	
Maximum size of FlexVol volume	Model-dependent	Model-dependent	See table below.
Maximum total shared data per FlexVol volume	16 TB	16 TB	Any subsequent attempts to create FlexClone files or FlexClone LUNs after the maximum size is reached cause Data ONTAP to create physical copies of the parent file or LUN.

Minimum sizes for root FlexVol volumes

Storage system model	Minimum root FlexVol volume size
N3700	10 GB
N3300	10 GB
N3400	16 GB
N3600	12 GB
N5200	12 GB
N5300	16 GB
N5500	16 GB
N5600	23 GB
N6040	16 GB
N6060	23 GB
N6070	37 GB
N6210	17 GB
N6240	22 GB
N6270	30 GB

Storage system model	Minimum root FlexVol volume size
N7600	37 GB
N7700	37 GB
N7800	69 GB
N7900	69 GB

Maximum FlexVol volume sizes for FlexClone files and FlexClone LUNs

Storage system model	Maximum FlexVol volume size for FlexClone files or FlexClone LUNs (TB)
N3300	1
N3400	3
N3600	2
N5200	2
N5300	4
N5500	3
N5600	16
N6040	4
N6060	16
N6070	16
N7600	16
N7700	16
N7800	16
N7900	16

Abbreviations

A list of abbreviations and their spelled-out forms are included here for your reference.

A
ABE (Access-Based Enumeration)
ACE (Access Control Entry)
ACL (access control list)
ACP (Alternate Control Path)
AD (Active Directory)
ALPA (arbitrated loop physical address)
ALUA (Asymmetric Logical Unit Access)
AMS (Account Migrator Service)
API (Application Program Interface)
ARP (Address Resolution Protocol)
ASCII (American Standard Code for Information Interchange)
ASP (Active Server Page)
ATA (Advanced Technology Attachment)

B
BCO (Business Continuance Option)
BIOS (Basic Input Output System
BCS (block checksum type)
BLI (block-level incremental)
BMC (Baseboard Management Controller)

C
CD-ROM (compact disc read-only memory)
CDDI (Copper Distributed Data Interface)
CDN (content delivery network)
CFE (Common Firmware Environment)
CFO (controller failover)
CGI (Common Gateway Interface)
CHA (channel adapter)
CHAP (Challenge Handshake Authentication Protocol)
CHIP (Client-Host Interface Processor)
CIDR (Classless Inter-Domain Routing)
CIFS (Common Internet File System)
CIM (Common Information Model)
CLI (command-line interface)
CP (consistency point)
CPU (central processing unit)
CRC (cyclic redundancy check)
CSP (communication service provider)

D
DAFS (Direct Access File System)
DBBC (database consistency checker)
DCE (Distributed Computing Environment)
DDS (Decru Data Decryption Software)
dedupe (deduplication)
DES (Data Encryption Standard)
DFS (Distributed File System)
DHA (Decru Host Authentication)
DHCP (Dynamic Host Configuration Protocol)
DIMM (dual-inline memory module)
DITA (Darwin Information Typing Architecture)
DLL (Dynamic Link Library)
DMA (direct memory access)
DMTD (Distributed Management Task Force)
DNS (Domain Name System)
DOS (Disk Operating System)
DPG (Data Protection Guide)
DTE (Data Terminal Equipment)

E

ECC (Elliptic Curve Cryptography) or (EMC Control Center)

ECDN (enterprise content delivery network)

ECN (Engineering Change Notification)

EEPROM (electrically erasable programmable read-only memory)

EFB (environmental fault bus)

EFS (Encrypted File System)

EGA (Enterprise Grid Alliance)

EISA (Extended Infrastructure Support Architecture)

ELAN (Emulated LAN)

EMU environmental monitoring unit)

ESH (embedded switching hub)

F

FAQs (frequently asked questions)

FAS (fabric-attached storage)

FC (Fibre Channel)

FC-AL (Fibre Channel-Arbitrated Loop)

FC SAN (Fibre Channel storage area network)

FC Tape SAN (Fibre Channel Tape storage area network)

FC-VI (virtual interface over Fibre Channel)

FCP (Fibre Channel Protocol)

FDDI (Fiber Distributed Data Interface)

FQDN (fully qualified domain name)

FRS (File Replication Service)

FSID (file system ID)

FSRM (File Storage Resource Manager)

FTP (File Transfer Protocol)

G
GbE (Gigabit Ethernet)
GID (group identification number)
GMT (Greenwich Mean Time)
GPO (Group Policy Object)
GUI (graphical user interface)
GUID (globally unique identifier)

H
HA (high availability)
HBA (host bus adapter)
HDM (Hitachi Device Manager Server)
HP (Hewlett-Packard Company)
HTML (hypertext markup language)
HTTP (Hypertext Transfer Protocol)

I
IB (InfiniBand)
IBM (International Business Machines Corporation)
ICAP (Internet Content Adaptation Protocol)
ICP (Internet Cache Protocol)
ID (identification)
IDL (Interface Definition Language)
ILM (information lifecycle management)
IMS (If-Modified-Since)
I/O (input/output)
IP (Internet Protocol)
IP SAN (Internet Protocol storage area network)
IQN (iSCSI Qualified Name)
iSCSI (Internet Small Computer System Interface)
ISL (Inter-Switch Link)
iSNS (Internet Storage Name Service)
ISP (Internet storage provider)

J
JBOD (just a bunch of disks)
JPEG (Joint Photographic Experts Group)

K
KB (Knowledge Base)
Kbps (kilobits per second)
KDC (Kerberos Distribution Center)

L
LAN (local area network)
LBA (Logical Block Access)
LCD (liquid crystal display)
LDAP (Lightweight Directory Access Protocol)
LDEV (logical device)
LED (light emitting diode)
LFS (log-structured file system)
LKM (Lifetime Key Management)
LPAR (system logical partition)
LREP (logical replication tool utility)
LUN (logical unit number)
LUSE (Logical Unit Size Expansion)
LVM (Logical Volume Manager)

M
MAC (Media Access Control)
Mbps (megabits per second)
MCS (multiple connections per session)
MD5 (Message Digest 5)
MDG (managed disk group)
MDisk (managed disk)
MIB (Management Information Base)
MIME (Multipurpose Internet Mail Extension)
MMC (Microsoft Management Console)
MMS (Microsoft Media Streaming)
MPEG (Moving Picture Experts Group)
MPIO (multipath network input/output)
MRTG (Multi-Router Traffic Grapher)
MSCS (Microsoft Cluster Service
MSDE (Microsoft SQL Server Desktop Engine)
MTU (Maximum Transmission Unit)

N
NAS (network-attached storage)
NDMP (Network Data Management Protocol)
NFS (Network File System)
NHT (N series Health Trigger)
NIC (network interface card)
NMC (Network Management Console)
NMS (network management station)
NNTP (Network News Transport Protocol)
NTFS (New Technology File System)
NTLM (NetLanMan)
NTP (Network Time Protocol)
NVMEM (nonvolatile memory management)
NVRAM (nonvolatile random-access memory)

O
OFM (Open File Manager)
OFW (Open Firmware)
OLAP (Online Analytical Processing)
OS/2 (Operating System 2)
OSMS (Open Systems Management Software)
OSSV (Open Systems SnapVault)

P
PC (personal computer)
PCB (printed circuit board)
PCI (Peripheral Component Interconnect)
pcnfsd (storage daemon)
(PC)NFS (Personal Computer Network File System)
PDU (protocol data unit)
PKI (Public Key Infrastructure)
POP (Post Office Protocol)
POST (power-on self-test)
PPN (physical path name)
PROM (programmable read-only memory)
PSU power supply unit)
PVC (permanent virtual circuit)

Q
QoS (Quality of Service)
QSM (Qtree SnapMirror)

R
RAD (report archive directory)
RADIUS (Remote Authentication Dial-In Service)
RAID (redundant array of independent disks)
RAID-DP (redundant array of independent disks, double-parity)
RAM (random access memory)
RARP (Reverse Address Resolution Protocol)
RBAC (role-based access control)
RDB (replicated database)
RDMA (Remote Direct Memory Access)
RIP (Routing Information Protocol)
RISC (Reduced Instruction Set Computer)
RLM (Remote LAN Module)
RMC (remote management controller)
ROM (read-only memory)
RPM (revolutions per minute)
rsh (Remote Shell)
RTCP (Real-time Transport Control Protocol)
RTP (Real-time Transport Protocol)
RTSP (Real Time Streaming Protocol)

S
SACL (system access control list)
SAN (storage area network)
SAS (storage area network attached storage) or (serial-attached SCSI)
SATA (serial advanced technology attachment)
SCSI (Small Computer System Interface)
SFO (storage failover)
SFSR (Single File SnapRestore operation)
SID (Secure ID)
SIMM (single inline memory module)
SLB (Server Load Balancer)
SLP (Service Location Protocol)
SNMP (Simple Network Management Protocol)
SNTP (Simple Network Time Protocol)
SP (Storage Processor)
SPN (service principal name)
SPOF (single point of failure)
SQL (Structured Query Language)
SRM (Storage Resource Management)
SSD (solid state disk
SSH (Secure Shell)
SSL (Secure Sockets Layer)
STP (shielded twisted pair)
SVC (switched virtual circuit)

T

TapeSAN (tape storage area network)

TCO (total cost of ownership)

TCP (Transmission Control Protocol)

TCP/IP (Transmission Control Protocol/Internet Protocol)

TOE (TCP offload engine)

TP (twisted pair)

TSM (Tivoli Storage Manager)

TTL (Time To Live)

U

UDP (User Datagram Protocol)

UI (user interface)

UID (user identification number)

Ultra ATA (Ultra Advanced Technology Attachment)

UNC (Uniform Naming Convention)

UPS (uninterruptible power supply)

URI (universal resource identifier)

URL (uniform resource locator)

USP (Universal Storage Platform)

UTC (Universal Coordinated Time)

UTP (unshielded twisted pair)

UUID (universal unique identifier)

UWN (unique world wide number)

V

VCI (virtual channel identifier)

VCMDB (Volume Configuration Management Database)

VDI (Virtual Device Interface)

VDisk (virtual disk)

VDS (Virtual Disk Service)

VFM (Virtual File Manager)

VFS (virtual file system)

VI (virtual interface)

vif (virtual interface)

VIRD (Virtual Router ID)

VLAN (virtual local area network)

VLD (virtual local disk)

VOD (video on demand)

VOIP (voice over IP)

VRML (Virtual Reality Modeling Language)

VTL (Virtual Tape Library)

W

WAFL (Write Anywhere File Layout)

WAN (wide area network)

WBEM (Web-Based Enterprise Management)

WHQL (Windows Hardware Quality Lab)

WINS (Windows Internet Name Service)

WORM (write once, read many)

WWN (worldwide name)

WWNN (worldwide node name)

WWPN (worldwide port name)

www (worldwide web)

Z
ZCS (zoned checksum)

Index

E

F

H

I

L

M

N

P

Q

R

S

T

U

V

W

Z

NA 210-05202_A0, Printed in USA

GC52-1277-04