Technical Report

NetApp ONTAP FlexGroup Volumes

Best Practices and Implementation Guide

Justin Parisi and Richard Jernigan, NetApp

January 2020 | TR-4571

Abstract

This document provides a brief overview of NetApp® ONTAP® FlexGroup and a set of best practices and implementation tips to use with this feature. The FlexGroup feature is an evolution of scale-out NAS containers that blends nearly infinite capacity with predictable, low-latency performance in metadata-heavy workloads. For information about FlexGroup volumes that is not covered in this document, email flexgroups-info@netapp.com, and we will add information to this technical report as necessary. For a more detailed technical overview of FlexGroup volumes, see TR-4557, NetApp ONTAP FlexGroup Volumes: A Technical Overview.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF BEST PRACTICES

Note: For a condensed version of this document, see TR-4571-a: NetApp ONTAP FlexGroup Volume Top Best Practices.

1 The Evolution of NAS in NetApp ONTAP

As hard-drive costs are driven down and flash hard-drive capacity grows exponentially, file systems are following suit. The days of file systems that number in the tens of gigabytes, or even terabytes, are over. Storage administrators face increasing demands from application owners for large buckets of capacity with enterprise-level performance.

Machine learning and artificial intelligence workloads involve storage needs for a single namespace that can extend into the petabyte range (with billions of files). With the rise in these technologies, along with the advent of big data frameworks such as Hadoop, the evolution of NAS file systems is overdue. NetApp® ONTAP® FlexGroup is the ideal solution for these types of architectures.

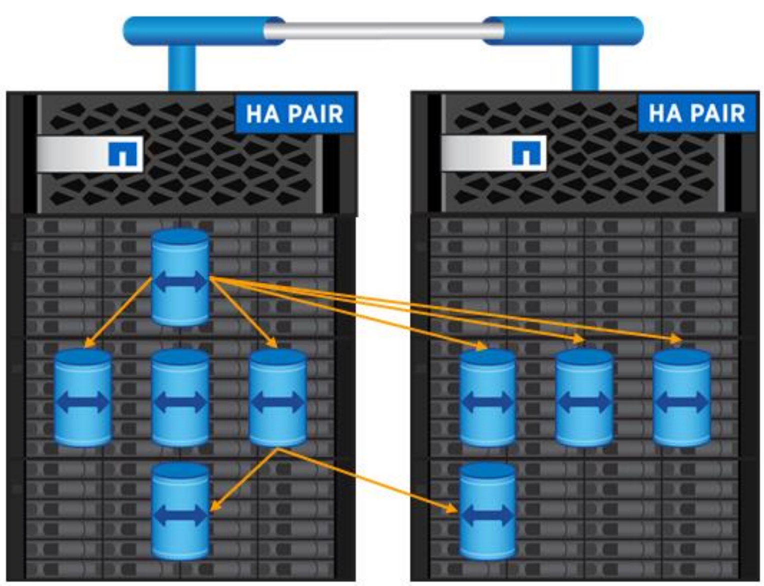
1.1 Flexible Volumes: A Tried-and-True Solution

The flexible volume, NetApp FlexVol® software, was introduced in NetApp Data ONTAP software in 2005 as part of the Data ONTAP 7.0 (Data ONTAP operating in 7-Mode) release. The concept was to take a storage file system and virtualize it across a hardware construct to provide flexible storage administration in an ever-changing data center.

FlexVol volumes could be grown or shrunk nondisruptively and be allocated to the storage operating system as thin-provisioned containers to enable overprovisioning of storage systems. Doing so allowed storage administrators the freedom to allocate space as consumers demanded it.

However, as data grew, file systems needed to grow. FlexVol can handle most storage needs with its 100TB capacity, and Data ONTAP provided a clustered architecture that those volumes could work with. But the use case for large buckets of storage in a single namespace required petabytes of storage. Before FlexGroup, ONTAP administrators could create junction paths to attach FlexVol volumes to one another. In this way, they created a file system on the cluster that could act as a single namespace. Figure 1 shows an example of what a FlexVol volume junction design for a large namespace would look like.

Figure 1) FlexVol design with junctioned architecture for >100TB capacity.



Although this architecture worked for many environments, it was awkward to manage and did not give a "single bucket" approach to the namespace, where the FlexVol volume's capacity and file count constraints are limiting factors.

1.2 Infinite Volumes: Massive Capacity with Limitations

In NetApp Data ONTAP 8.1.1, the Infinite Volume solution was presented as a potential solution to enterprises with massively large storage needs. With a 20PB maximum and the capability to grow a single namespace nondisruptively, the Infinite Volume solution provided a more than capable method of storing large amounts of data.

Single Namespace Metadata Volume Limitations: Infinite Volume

Because the Infinite Volume solution used a single namespace volume for all metadata operations, several limitations applied:

- Less than stellar performance, with large amounts of metadata because volume affinity limits and serial operations created CPU efficiencies
- A two-billion-file maximum due to the single FlexVol volume limit that was imposed by the metadata volume
- The inability to share storage virtual machines (SVMs) with FlexVol volumes
- No SMB2.x and 3.x support

Therefore, although Infinite Volume provided an excellent method to store archive data, it did not offer a way to cover multiple use cases in big data environments with predictable low latency.

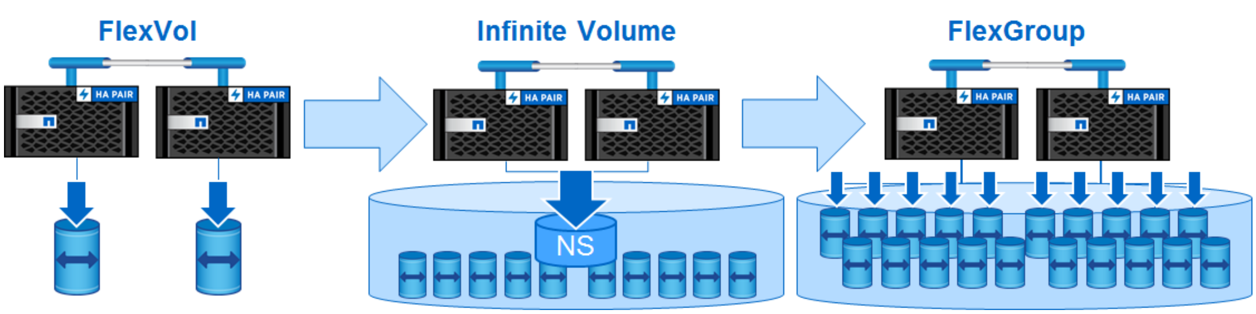
Note: In ONTAP 9.5 and later, the Infinite Volume feature is no longer supported. See Deprecation of Infinite Volume later in this document.

1.3 FlexGroup Volumes: An Evolution of NAS

ONTAP 9.1 brought innovation to scale-out NAS file systems: the NetApp ONTAP FlexGroup volume.

With FlexGroup volumes, a storage administrator can easily provision a massive single namespace in a matter of seconds. FlexGroup volumes have virtually no capacity or file count constraints outside of the physical limits of hardware or the total volume limits of ONTAP. Limits are determined by the overall number of constituent member volumes that work in collaboration to dynamically balance load and space allocation evenly across all members. There is no required maintenance or management overhead with a FlexGroup volume. You simply create the FlexGroup volume and share it with your NAS clients. ONTAP does the rest.

Figure 2) Evolution of NAS file systems in ONTAP.



2 Terminology

Many of the usual NetApp ONTAP terms (such as SVM, LIF, and FlexVol) are covered in TR-3982: NetApp Clustered Data ONTAP 8.3.x and 8.2.x. Terminology specific to NetApp ONTAP FlexGroup is covered in the following list.

- **Constituent/member volumes.** In a FlexGroup context, “constituent volume” and “member volume” are interchangeable terms. They refer to the underlying NetApp FlexVol volumes that make up a FlexGroup volume and provide the capacity and performance gains that are achieved only with a FlexGroup volume.
- **FlexGroup volume.** A FlexGroup volume is a single namespace that is made up of multiple constituent/member volumes. It is managed by storage administrators, and it acts like a NetApp FlexVol volume. Files in a FlexGroup volume are allocated to individual member volumes and are not striped across volumes or nodes.
- **Affinity.** Affinity describes the tying of a specific operation to a single thread.
- **Automated Incremental Recovery (AIR).** Automated Incremental Recovery is an ONTAP subsystem that repairs FlexGroup inconsistencies dynamically, with no outage or administrator intervention required.
- **Ingest.** Ingest is the consumption of data by way of file or folder creations.
- **Junction paths.** Junction paths were used to provide capacity beyond a FlexVol volume's 100TB limit prior to the simplicity and scale-out of FlexGroup. Junction paths join multiple FlexVol volumes together to scale out across a cluster and provide multiple volume affinities. The use of a junction path in ONTAP is known as “mounting” the volume within the ONTAP namespace.
- **Large files.** See the next section, What Are Large Files?.
- **Overprovisioning and thin provisioning.** Overprovisioning (or thin provisioning) storage is the practice of disabling a volume’s space guarantee (`guarantee = none`). This practice allows the virtual space allocation of the FlexVol volume to exceed the physical limits of the aggregate that it resides on. For example, with overprovisioning, a FlexVol volume can be 100TB on an aggregate that has a physical size of only 10TB. Overprovisioning allows storage administrators to grow volumes to large sizes to avoid the need to grow them later, but it does present the management overhead of needing to monitor available space closely.

  In overprovisioned volumes, the available space reflects the actual physical available space in the aggregate. Therefore, the usage percentage and capacity available values might seem off a bit. However, they simply reflect a calculation of the actual space that is available when compared with the virtual space that is available in the FlexVol volume. For a more accurate portrayal of space allocation when using overprovisioning, use the `aggregate show-space` command.
- **Remote access layer (RAL).** The remote access layer (RAL) is a feature in the NetApp WAFL® system that allows a FlexGroup volume to balance ingest workloads across multiple FlexGroup constituents or members.
- **Remote hard links.** Remote hard links are the building blocks of FlexGroup. These links act as normal hard links but are unique to ONTAP. The links allow a FlexGroup volume to balance workloads across multiple remote members or constituents. In this case, “remote” simply means “not in the parent volume.” A remote hard link can be another FlexVol member on the same aggregate or node.

2.1 What Are Large Files?

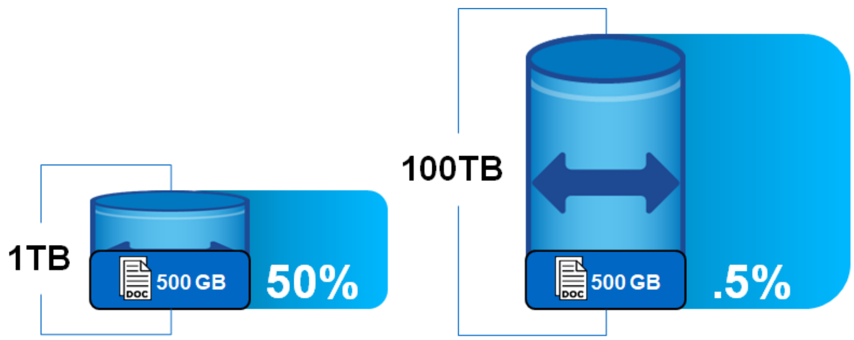
This document uses the term “large file” liberally. Therefore, it’s important to define exactly what a “large file” is in the context of FlexGroup.

A FlexGroup volume operates optimally when a workload is ingesting numerous small files, because FlexGroup volumes maximize the system resources to address those specific workloads that might bottleneck because of serial processing in a FlexVol volume. FlexGroup volumes also work well with various other workloads (as defined in section 4, Use Cases). One type of workload that can create problems, however, is a workload with larger files or files that grow over time, such as database files.

In a FlexGroup volume, a large file is a product of the percentage of allocated space, not of any specific file size. Thus, in some FlexGroup configurations—for example, in which the member volume size is only 1TB—a “large file” might be 500GB (50% of the member volume size). In other configurations, for example, in which the member volume size is 100TB, that same 500GB file size would only take up 0.5% of the volume capacity. This type of file could be large enough to throw off the ingest heuristics in the FlexGroup volume, or it could potentially create problems later when the member volume gets closer to full.

Starting in ONTAP 9.6, elastic sizing helps mitigate concerns with larger files: ONTAP borrows space from other member volumes to allow large files to complete their writes. ONTAP 9.7 also introduces ingest algorithm changes to help balance large files and/or datasets with mixed file sizes. Both of these features make FlexGroup volumes a realistic landing place for most workloads.

Figure 3) What is a large file?



3 NetApp ONTAP FlexGroup Advantages

NetApp ONTAP FlexGroup provides various advantages for different workloads. The advantages are described in the following sections.

3.1 Massive Capacity and Predictable Low Latency for High-Metadata Workloads

FlexGroup volumes offer a way for storage administrators to easily provision massive amounts of capacity with the ability to nondisruptively scale out that capacity. FlexGroup also enables parallel performance for high metadata workloads that can increase throughput and total operations while still providing low latency for mission-critical workloads.

3.2 Efficient Use of All Cluster Hardware

FlexGroup volumes allow storage administrators to easily span multiple physical aggregates and nodes with member FlexVol volumes, while maintaining a true single namespace for applications and users to dump data into. Although clients and users see the space as monolithic, ONTAP is working behind the scenes to distribute the incoming file creations evenly across the FlexGroup volume to provide efficient CPU and disk utilization.

3.3 Simple, Easy-to-Manage Architecture and Balancing

To make massive capacity easy to deploy, NetApp lets you manage FlexGroup volumes like NetApp FlexVol volumes. ONTAP handles the underlying member volume creation and balance across the cluster nodes and provides a single access point for NAS shares.

3.4 Superior Density for Big Data

A FlexGroup volume enables you to condense large amounts of data into smaller data center footprints by way of the superb storage efficiency features of ONTAP, including:

- Thin provisioning
- Data compaction
- Data compression
- Deduplication

In addition, ONTAP supports large SSDs, which can deliver massive amounts of raw capacity in a single 24-drive shelf enclosure. It is possible to get petabytes of raw capacity in just 10U of rack space, which cuts costs on cooling, power consumption, and rack rental space and offers excellent density in the storage environment. These features, combined with a FlexGroup volume's ability to efficiently use that capacity and balance performance across a cluster, give you a solution that was made for big data.

4 Use Cases

The NetApp ONTAP FlexGroup design is most beneficial in specific use cases (electronic design and automation, software development, and so on). They are listed in section 4.1, "Ideal Use Cases."

4.1 Ideal Use Cases

A FlexGroup volume works best with workloads that are heavy on ingest (a high level of new data creation), heavily concurrent, and evenly distributed among subdirectories:

- Electronic design automation (EDA)
- Artificial intelligence and machine learning log file repositories
- Software build/test environments (such as GIT)
- Seismic/oil and gas
- Media asset or HIPAA archives
- File streaming workflows
- Unstructured NAS data (such as home directories)
- Big data and data lakes (Hadoop with the NetApp NFS connector)

4.2 Non-Ideal Cases

Some workloads are currently not recommended for FlexGroup volumes. These workloads include:

- Virtualized workloads
- Workloads that require striping (large files spanning multiple nodes or volumes)
- Workloads that require specific control over the layout of the relationships of data to NetApp FlexVol volumes
- Workloads that require specific features and functionality that are not currently available with FlexGroup volumes

If you have questions, feel free to email ng-flexgroups-info@netapp.com.

FlexGroup Volume Use Case Examples

The following sections describe two examples of real-world use cases.

FlexGroup Use Case Example #1: NetApp Active IQ Infrastructure

The NetApp on NetApp program involves NetApp's backing infrastructure teams and their use of NetApp products. This program serves not only to provide NetApp IT and other groups with the best solutions for their problems, but also to show that NetApp has confidence in their own offerings.

In Episode 182 of the Tech OnTap® podcast, representatives from NetApp on NetApp describe how FlexGroup volumes are being used to power the NetApp Active IQ® infrastructure. For details about the solution, including statistics, read the blog post ONTAP FlexGroup Technology Powers NetApp's Massive Active IQ Data Lake.

FlexGroup Use Case Example #2: Back Up Repository for SQL Server

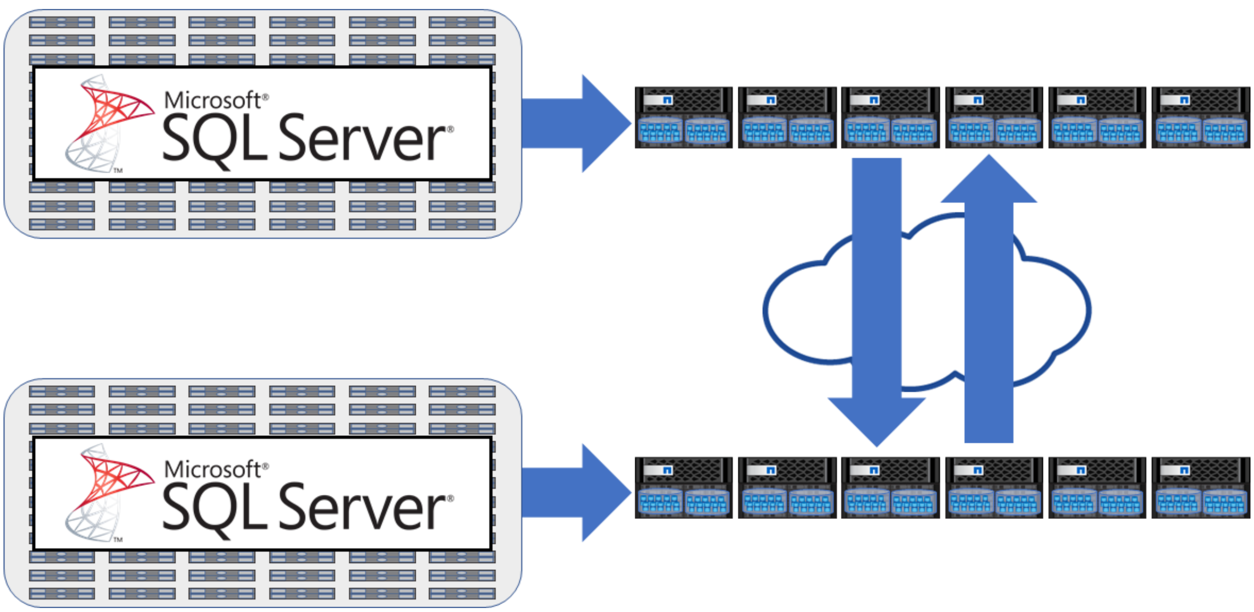
In this environment, the customer wanted to perform compressed backups of 5,000 Microsoft SQL Servers over SMB. This test was done with approximately 200 servers to vet out the solution, with a slow ramp up over the course of a few months.

But this database isn't only a backup target—it will also be replicated to a disaster recovery site by using NetApp SnapMirror® for extra data protection.

Each site has a six-node FAS8200 cluster running ONTAP 9.4 using 6TB near-line SAS (NL-SAS) encrypted drives. Each cluster holds 3PB of usable capacity. The clusters use 30 FlexGroup volumes and use qtrees within the volumes for data organization.

The FlexGroup volumes are 64TB each and the member volumes are 2.6TB each, with four members per node across six nodes (24 total members per FlexGroup volume).

Figure 4) SQL Server backup environment.



The Results

This customer needed a single namespace that could collect ~150TB worth of MSSQL backup data over a 12-hour period. That's ~12TB per hour at ~3.5GB per second.

During testing, we used 222 servers at site A and 171 servers at site B. During the test, each cluster's CPU was at 95% utilization and the backup jobs (sequential writes) were able to accomplish 8.4GB per second, which is ~2.4 times the amount of throughput the job needed. At this rate, the backups could complete in approximately 5 hours, rather than the 12-hour window. Also, this SMB workload performed approximately 120,000 IOPS. When more clients are added to this workload, we expect the throughput to max out at around 9GB/sec.

Figure 5) Throughput and total operations during test runs.

cpu avg	cpu busy	total ops	nfs-ops	cifs-ops	fcache ops	spin-ops	total recv	total sent	data busy	data recv	data sent	cluster busy	cluster recv	cluster sent	disk read	disk write	pkts recv	pkts sent
56%	81%	54530	0	54530	0	54420	6.16GB	2.65GB	44%	3.34GB	28.3MB	22%	2.82GB	2.62GB	128MB	3.31GB	968237	898917
65%	78%	70482	0	70482	0	70407	8.03GB	3.44GB	47%	4.33GB	30.9MB	24%	3.70GB	3.41GB	114MB	4.79GB	1178768	1102912
74%	87%	88725	0	88725	0	88105	10.2GB	4.30GB	49%	5.44GB	37.1MB	36%	4.78GB	4.26GB	157MB	5.56GB	1389743	1324559
86%	92%	111577	0	111577	0	110569	12.8GB	5.88GB	53%	6.84GB	41.9MB	31%	6.00GB	5.84GB	153MB	6.77GB	1726469	1679506
88%	92%	115036	0	115036	0	113509	13.2GB	6.44GB	51%	7.06GB	45.9MB	49%	6.14GB	6.40GB	142MB	7.65GB	1845760	1814549
92%	95%	118148	0	118148	0	117104	13.6GB	6.11GB	45%	7.26GB	49.9MB	42%	6.34GB	6.07GB	149MB	8.11GB	1802929	1769902
95%	98%	122953	0	122953	0	122123	14.3GB	7.10GB	47%	7.54GB	45.9MB	43%	6.75GB	7.06GB	135MB	8.29GB	1978205	1952416
96%	99%	126241	0	126241	0	125104	14.6GB	6.43GB	53%	7.75GB	54.3MB	44%	6.80GB	6.37GB	135MB	8.28GB	1865375	1849777
95%	97%	121948	0	121948	0	120719	13.9GB	7.25GB	44%	7.47GB	47.3MB	40%	6.41GB	7.20GB	108MB	8.30GB	1995908	1967271
95%	98%	123079	0	123079	0	121113	13.9GB	5.71GB	41%	7.56GB	49.0MB	38%	6.37GB	5.66GB	129MB	8.40GB	1761097	1712061
95%	97%	120567	0	120567	0	120493	13.7GB	7.01GB	42%	7.41GB	47.6MB	36%	6.34GB	6.96GB	116MB	8.48GB	1888934	1882711
95%	98%	119573	0	119573	0	119458	13.6GB	5.74GB	37%	7.33GB	44.4MB	35%	6.28GB	5.69GB	111MB	8.19GB	1702969	1671363
95%	97%	119538	0	119538	0	119829	13.5GB	6.98GB	41%	7.34GB	46.2MB	35%	6.17GB	6.93GB	120MB	8.44GB	1880298	1873821
95%	98%	118119	0	118119	0	118373	13.4GB	5.56GB	37%	7.25GB	45.4MB	37%	6.17GB	5.52GB	118MB	8.42GB	1666066	1630785
95%	98%	118862	0	118862	0	118327	13.6GB	6.29GB	39%	7.29GB	47.1MB	33%	6.30GB	6.24GB	114MB	8.31GB	1784134	1759266
96%	99%	121039	0	121039	0	121136	13.7GB	6.67GB	38%	7.44GB	44.5MB	34%	6.21GB	6.63GB	120MB	8.35GB	1832520	1827158
96%	99%	120852	0	120852	0	120920	13.7GB	5.77GB	39%	7.42GB	47.8MB	33%	6.24GB	5.72GB	111MB	8.51GB	1706939	1678778
94%	97%	119819	0	119819	0	120129	13.7GB	7.05GB	41%	7.36GB	42.6MB	35%	6.29GB	7.01GB	118MB	8.49GB	1882656	1877381

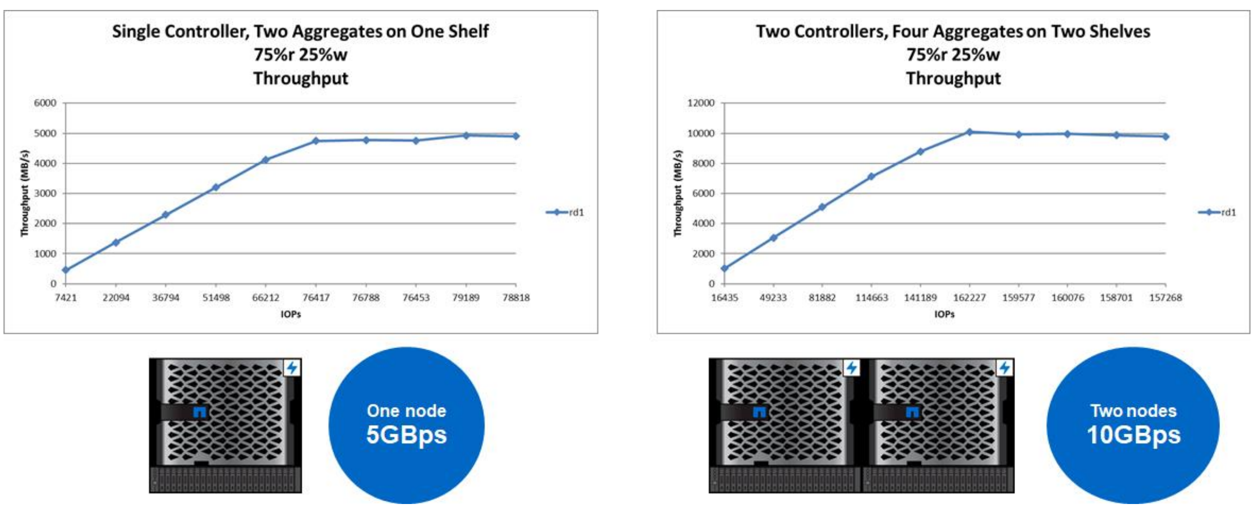
Data Protection

In addition to the performance seen on the FlexGroup volume for the production workload, this customer was also able to achieve a high rate of transfer for the SnapMirror relationships between sites—8.4GB per second for the SnapMirror transfer. This rate means that the replication window for a 150TB dataset would be about 5.5 hours for the initial transfer. After that, the deltas should be able to complete well within the required transfer window, providing a solid disaster recovery plan for these MSSQL backups.

Scale-Out Performance

This six-node cluster was able to push over 8.4GB per second to a FlexGroup volume. In NetApp Customer Proof of Concept (CPOC) labs, we've seen near-linear performance gains by adding nodes to a cluster. The following graphs show throughput results for a single-node NetApp AFF A700 all-flash storage system and a two-node AFF A700.

Figure 6) CPOC scale-out throughput results.



Note: If you want to add more performance to your backup workload, you can add more nodes.

Conclusion

Not only is a FlexGroup volume great for small or high-file-count workloads such as EDA and software builds, but it also can handle high throughput requirements for larger streaming files. It also reduces backup windows by scaling out storage across multiple nodes and applies all your cluster resources while maintaining performance even with spinning drive.

5 Supported Features

This section shows which NetApp ONTAP features are supported for use with FlexGroup volumes; it also notes the ONTAP version in which feature support was added. If a feature is not listed in this section, email flexgroups-info@netapp.com for information.

Table 1) General ONTAP feature support.

Supported Feature	Version of ONTAP First Supported
NetApp Snapshot™ technology	ONTAP 9.0
NetApp SnapRestore® software (FlexGroup level)	ONTAP 9.0
Hybrid aggregates	ONTAP 9.0
Constituent or member volume move	ONTAP 9.0
Postprocess deduplication	ONTAP 9.0
NetApp RAID-TEC™ technology	ONTAP 9.0
Per-aggregate consistency point	ONTAP 9.0
Sharing FlexGroup with FlexVol in the same SVM	ONTAP 9.0
Active IQ Unified Manager	ONTAP 9.1
Inline adaptive compression	ONTAP 9.1
Inline deduplication	ONTAP 9.1
Inline data compaction	ONTAP 9.1
Thin provisioning	ONTAP 9.1
NetApp AFF	ONTAP 9.1
Quota reporting	ONTAP 9.1
SnapMirror technology	ONTAP 9.1
User and group quota reporting (no enforcement)	ONTAP 9.1
Aggregate inline deduplication (cross-volume deduplication)	ONTAP 9.2
NetApp Volume Encryption (NVE)	ONTAP 9.2
NetApp SnapVault® technology	ONTAP 9.3
Qtrees	ONTAP 9.3
Automated deduplication schedules	ONTAP 9.3
Version-independent SnapMirror and unified replication	ONTAP 9.3
Antivirus scanning for SMB	ONTAP 9.3
Volume autogrow	ONTAP 9.3
QoS maximums/ceilings	ONTAP 9.3
FlexGroup expansion without SnapMirror rebaseline	ONTAP 9.3

Supported Feature	Version of ONTAP First Supported
Improved ingest heuristics	ONTAP 9.3
SMB change/notify	ONTAP 9.3
File audit	ONTAP 9.4
NetApp FPolicy™	ONTAP 9.4
Adaptive QoS	ONTAP 9.4
QoS minimums (AFF only)	ONTAP 9.4
Relaxed SnapMirror limits	ONTAP 9.4
SMB 3.x Multichannel	ONTAP 9.4
FabricPool	ONTAP 9.5
Quota enforcement	ONTAP 9.5
Qtree statistics	ONTAP 9.5
Inherited SMB watches and change notifications	ONTAP 9.5
SMB copy offload (offloaded data transfer)	ONTAP 9.5
Storage-Level Access Guard	ONTAP 9.5
NetApp FlexCache® (cache only; FlexGroup as origin supported in ONTAP 9.7)	ONTAP 9.5
Volume rename	ONTAP 9.6
Volume shrink	ONTAP 9.6
NetApp MetroCluster™	ONTAP 9.6
Elastic sizing	ONTAP 9.6
Continuously Available Shares (SMB)* * SQL Server and Hyper-V workloads only	ONTAP 9.6
NetApp Aggregate Encryption (NAE)	ONTAP 9.6
NetApp Cloud Volumes ONTAP	ONTAP 9.6
NetApp FlexClone®	ONTAP 9.7
In-place conversion of FlexVol to FlexGroup (see “Deploying a FlexGroup Volume on Aggregates with Existing FlexVol Volumes“)	ONTAP 9.7
vStorage APIs for Array Integration (VAAI)	ONTAP 9.7
NDMP	ONTAP 9.7
NFSv4.0 and NFSv4.1 (including parallel NFS, or pNFS)	ONTAP 9.7
FlexGroup volumes as FlexCache origin volumes	ONTAP 9.7

Table 2) General NAS protocol version support.

Supported NAS Protocol Version	Version of ONTAP First Supported
NFSv3	ONTAP 9.0
SMB2.1, SMB3.x	ONTAP 9.1 RC2
NFSv4.x	ONTAP 9.7

Table 3) Unsupported SMB2.x and 3.x features.

Unsupported SMB2.x Features	Unsupported SMB 3.x Features
• SMB Remote Volume Shadow Copy Service (VSS)	• SMB transparent failover • SMB scale-out • VSS for SMB file shares. • SMB directory leasing • SMB direct or remote direct memory access (RDMA) Note: SMB 3.0 encryption is supported.

Note: Remote VSS is not the same as the SMB Previous Versions tab. Remote VSS is application-aware Snapshot functionality and is most commonly used with Hyper-V workloads. FlexGroup volumes have supported the SMB Previous Versions tab since it was introduced.

Behavior of Unsupported SMB Features

Usually, if an SMB feature is unsupported in ONTAP, it simply does not work. With NetApp ONTAP FlexGroup, there are some considerations regarding unsupported SMB features and functionality.

Table 4) How unsupported SMB features behave with FlexGroup volumes.

Feature	Behavior with FlexGroup Volumes
SMBv1.0	Access fails or is denied for any shares accessing with SMB 1.0. This can affect Windows XP, Windows 2003, and office equipment such as scanners or copiers that attempt to connect to the NAS with SMB.
Change notification/SMB watches (before ONTAP 9.3)	• Change notifications are supported as of ONTAP 9.3. Before ONTAP 9.3, the behavior is as follows: • There are no warning or failures. Change notification simply doesn't take place. For more information about change notification with SMB, see this MSDN article. • Lack of change notifications in SMB can affect applications that depend on the immediate appearance of newly created files in Windows folders.
Offloaded data transfer (ODX)	In versions earlier than ONTAP 9.5, failovers occur to the traditional client-side copy. Impact is low; failovers are not as fast. In ONTAP 9.5 and later, ODX works as expected. For more information about ODX, see this TechNet article.
Remote Volume Shadow Copy Service (VSS)	There is no warning; Remote VSS just does not work. Impact should be low because the primary use case for Remote VSS is with Hyper-V, which is not a recommended workload for FlexGroup volumes. For more information about Remote VSS, see this TechNet article.
Continuously available shares	Continuously available shares are not allowed on FlexGroup volumes before ONTAP 9.6. Attempting to set the share property fails. For more information about continuously available shares, see this TechNet article.

6 Maximums and Minimums

This section covers the maximums and minimums that are specific to NetApp ONTAP FlexGroup volumes. Table 5 lists the maximum values and shows whether the maximum is hard-coded/enforced or a recommended/tested value.

Table 5) FlexGroup maximums.

	Value	Value Type
FlexGroup volume size	20PB	Tested/recommended*
File count	400 billion	Tested/recommended*
Cluster node count	24 (12 HA pairs)	Hard-coded/enforced
NetApp FlexVol member volume size	100TB	Hard-coded/enforced
FlexVol member volume file count	2 billion	Hard-coded/enforced
NetApp SnapMirror volume count (member per FlexGroup)	32 (ONTAP 9.4 and earlier) 200 (ONTAP 9.5 and later)	Hard-coded/enforced
SnapMirror volume count (FlexGroup total per cluster)	100 (ONTAP 9.4 and earlier) 6,000 (ONTAP 9.5 and later)	Hard-coded/enforced
File size	16TB	Hard-coded/enforced
FlexVol member constituent count	200	Tested/recommended*
Aggregate size/count	Same as NetApp ONTAP limits	Hard-coded/enforced

Table 6) FlexGroup minimums.

	Value	Value Type
FlexVol member size	100GB	Tested/recommended*
Data aggregate count	1	Hard-coded/enforced
SnapMirror schedule	30 minutes	Tested/recommended*
NetApp Snapshot schedule	30 minutes	Tested/recommended*

*Limits described as tested/recommended are tested limits based on a 10-node cluster. If allowed by the platform, actual limits are not hard-coded and can extend beyond these limits up to 24 nodes. For more information, see the section "FlexGroup Volume Maximums." However, official support for the number of member volumes is 200. If you need to exceed this limit, contact your NetApp sales representative to start the qualification process for more member volumes.

7 Deciding Whether FlexGroup Is the Right Fit

NetApp ONTAP FlexGroup volumes are an ideal fit for many use cases—particularly the ones that are listed in section 4.1, "Ideal Use Cases."

However, not all use cases make sense for FlexGroup volumes. This section provides information to help you decide whether FlexGroup volumes are the right fit for your workloads.

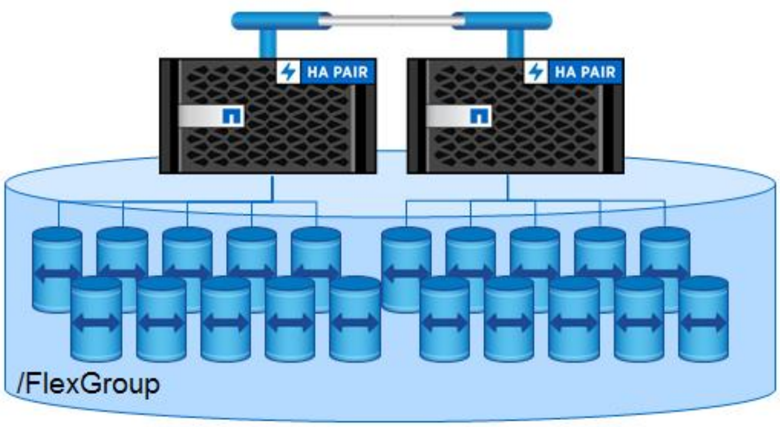
7.1 Scale-Out Performance

FlexGroup volumes distribute their data and load among the multiple constituents that make up the collective FlexGroup volume. This model allows a FlexGroup volume to use more of the resources within each node (CPU, network adapters, disks, and so on) and to use more nodes within a cluster to address a workload.

In addition, the concept ties in nicely with the NetApp ONTAP clustered architecture, which allows the nondisruptive addition of nodes and disks to increase performance without negatively affecting applications. With a FlexGroup volume, you can simply expand the FlexGroup to add more members or use nondisruptive volume move technology to redistribute the member volumes across the new nodes.

A single FlexGroup volume internally comprises multiple separate NetApp FlexVol volumes, which in turn can be stored on any aggregates and can span multiple nodes in your cluster.

Figure 7) FlexGroup volume.



When clients add files and subdirectories to the FlexGroup volume, ONTAP automatically determines the best FlexVol member to use for storing each new file and subdirectory. The FlexGroup volume attempts to organize your data, both for fastest access and for good data and load distribution.

Because of this workload distribution, FlexGroup volumes can handle much more metadata traffic than a FlexVol volume or an infinite volume. Thus, FlexGroup volumes can be useful for workloads that are metadata-intensive or that require a large amount of throughput.

7.2 Feature Compatibility Limitations

FlexGroup volumes in ONTAP 9.1 and later only support some common NAS protocols, such as NFSv3, NFSv4.0 and v4.1, SMB2.x, and SMB3. For details about the support of those protocols, see Table 1, Table 2, and Table 3 in section 5, “Supported Features.” The S3 protocol is supported for use with FlexGroup volumes via public preview in ONTAP 9.7.

SMB 1.0 is not supported for use with FlexGroup volumes. NFSv4.2 is currently not supported for FlexVols nor FlexGroup volumes. FlexGroup volumes do not support SAN access (iSCSI, and so on), and NetApp currently does not recommend that FlexGroup volumes be used for virtual workloads. FlexGroup volumes are missing functionality that is important to virtualization workloads, such as copy offload support/VAAI and NetApp FlexClone technology.

Table 7 provides information for deciding whether FlexGroup volumes are the right fit for an environment by comparing the currently available container types in ONTAP.

Table 7) ONTAP volume family comparison.

	FlexVol Volumes	Infinite Volume	FlexGroup Volumes
Client access protocols (current support)	SAN, NAS—all versions	NAS only: • SMB 1.0 • NFSv3 • NFSv4.0, NFSv4.1 • pNFS	NAS: • SMB2.1/3.x • NFSv3 • NFSv4.0, NFSv4.1 S3 (public preview)
Capacity scaling	• Single FlexVol volume • Can be mounted to other FlexVol volumes • 100TB, 2 billion file limit	• 20PB • 2 billion files	• 20PB* • 400 billion files* *Current tested limits on 10-node cluster; can extend beyond these values
Metadata scaling	Both Infinite Volume and FlexVol volumes are limited to a single node for metadata processing and serial processing of metadata, which does not take full advantage of the node's CPU threads.		FlexGroup volumes can use multiple nodes (and their resources) and multiple aggregates. In addition, FlexGroup can use multiple volume affinities to maximize CPU thread utilization potential.
ONTAP feature compatibility	Compatible with all ONTAP features	Supports many ONTAP features. For FlexGroup information, see TR-4557, NetApp ONTAP FlexGroup Volumes: A Technical Overview and the "Supported Features" section of this document. For details about Infinite Volume, see TR-4037: Introduction to NetApp Infinite Volume.	
Throughput scaling	Limited to: • One node (set of CPU, RAM, network ports, connection limits, and so on) • One aggregate	Both Infinite Volume and FlexGroup volumes can use the resources of an entire cluster in service of I/O, providing much higher throughput than a single FlexVol volume can.	
ONTAP upgrades and reverts	Data stored in any volume family is safely retained during ONTAP version changes, with one exception: If reverting to a release earlier than ONTAP 9.1, FlexGroup volumes cannot be retained.		
GUI compatibility	All volume families have some level of GUI support in NetApp OnCommand products, including: • ONTAP System Manager (formerly OnCommand System Manager) • Active IQ Performance Manager • Active IQ Unified Manager • OnCommand Insight		

7.3 Simplifying Performance

A single FlexGroup volume can consist of multiple FlexVol member volumes, which in turn can reside on any aggregate and on any node in your cluster. As clients drive traffic against that FlexGroup volume, ONTAP automatically breaks that traffic into tasks for different constituent FlexVol volumes to perform. This approach provides for a concurrency of operations that a single FlexVol volume is incapable of handling.

The benefit of this scale-out behavior is a dramatic increase in processing power. A single FlexGroup volume can service much heavier workloads than a single FlexVol volume can at more predictable latencies.

FlexVol Versus FlexGroup: Software Build

In a simple workload benchmark using a software build tool (Git), a Linux kernel was compiled on a single FAS8080 node running ONTAP 9.1 with two aggregates of SAS drives and eight FlexVol member constituents in a FlexGroup volume, versus a single FlexVol volume on the same hardware. The metric being measured was a simple “time to completion” test. In this benchmark, the FlexGroup volume outperformed the FlexVol volume by two to six times across multiple Git operations. In addition, the same Git test was run with a GCC compile on NetApp AFF.

Note: The GCC compile works with a higher file count, thus the differences in completion times.

Figure 8) Git benchmark—Linux compile in FlexGroup versus FlexVol.

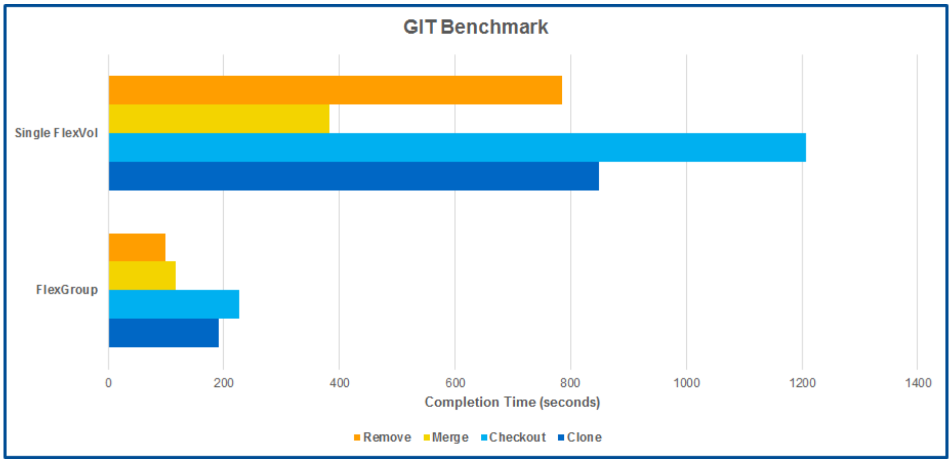
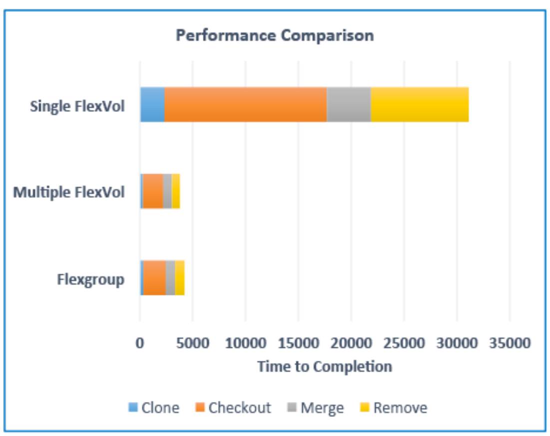


Figure 9) Git benchmark—GCC compile in FlexGroup versus FlexVol.

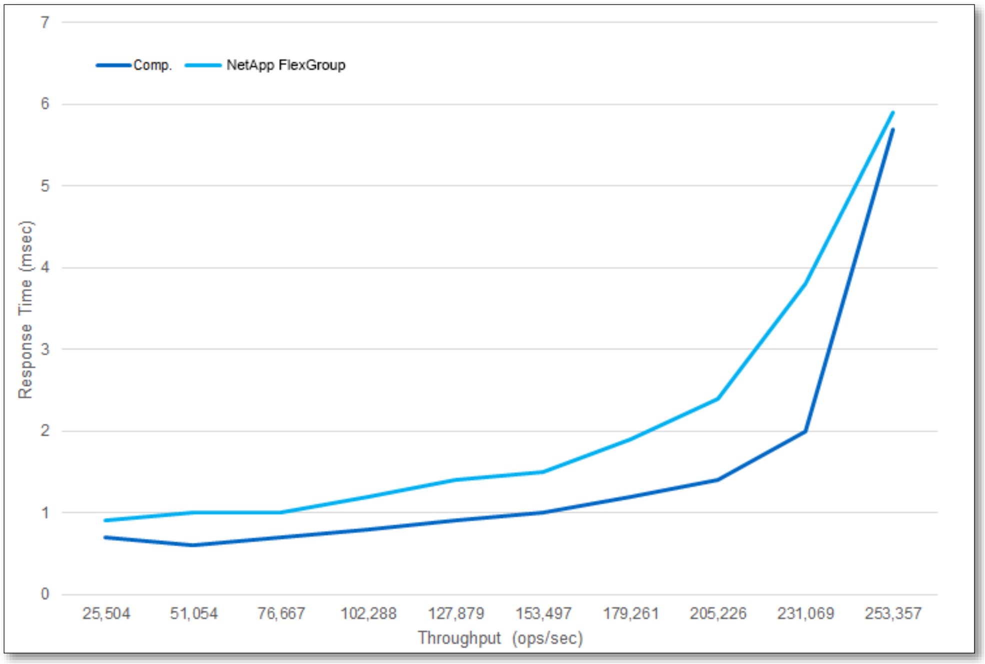


FlexGroup Versus Scale-Out NAS Competitor: Do More with Less

In another benchmark, we compared a FlexGroup volume on a two-node FAS8080 cluster running ONTAP 9.1 using SAS drives against a competitor system using 14 nodes. The competitor system also used some SSDs for metadata caching. This test used a standard NAS workload generation tool to simulate workloads.

In the test, we saw that a single FlexGroup volume with eight member constituents was able to ingest nearly the same number of ops/second at essentially the same latency curve as the competitor’s 14-node cluster.

Figure 10) FlexGroup (two-node cluster) versus competitor (14-node cluster): standard NAS workload.



SPEC SFS 2014_swbuild Submission: FlexGroup Volume, ONTAP 9.2

NetApp also submitted results from the official SPEC SFS 2014_swbuild benchmark test, which allows storage vendors to test their systems against a standardized test that is approved by an independent benchmarking consortium. See the NetApp results of this test here.

See the results for competitor systems here.

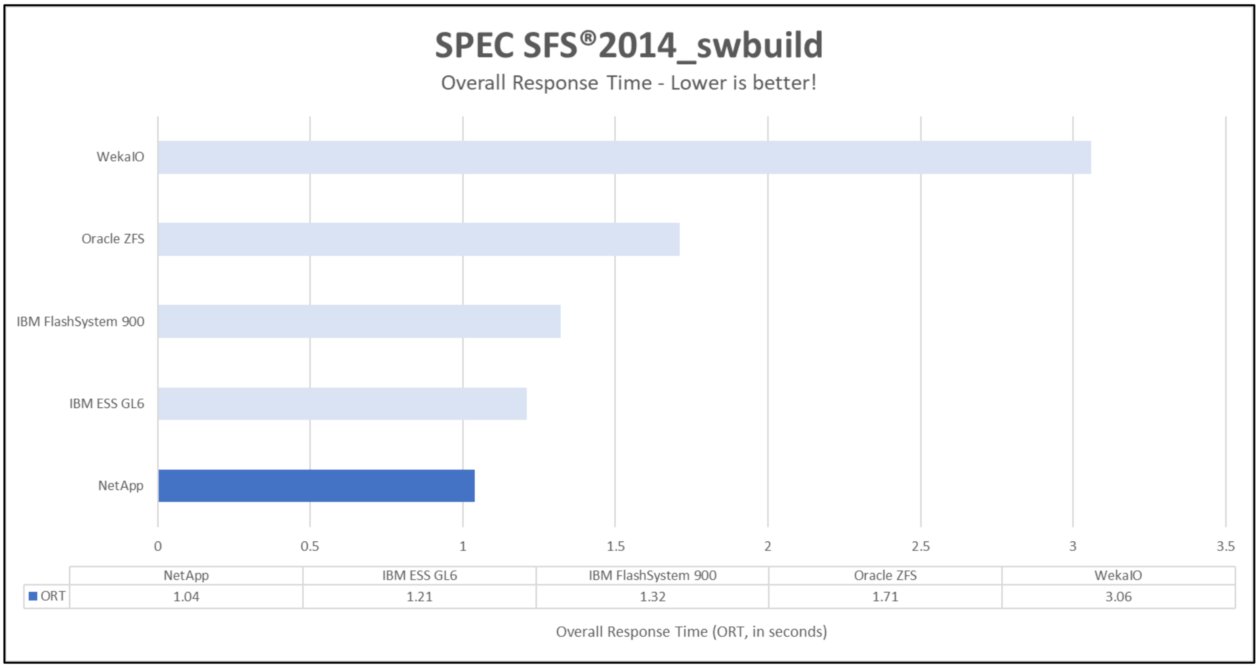
NetApp Results

The benchmark includes a metric known as overall response time (ORT), defined here:

```
The overall response time is a measure of how the system will respond under an average load.
Mathematically, the value is derived by calculating the area under the curve divided by the peak
throughput.
```

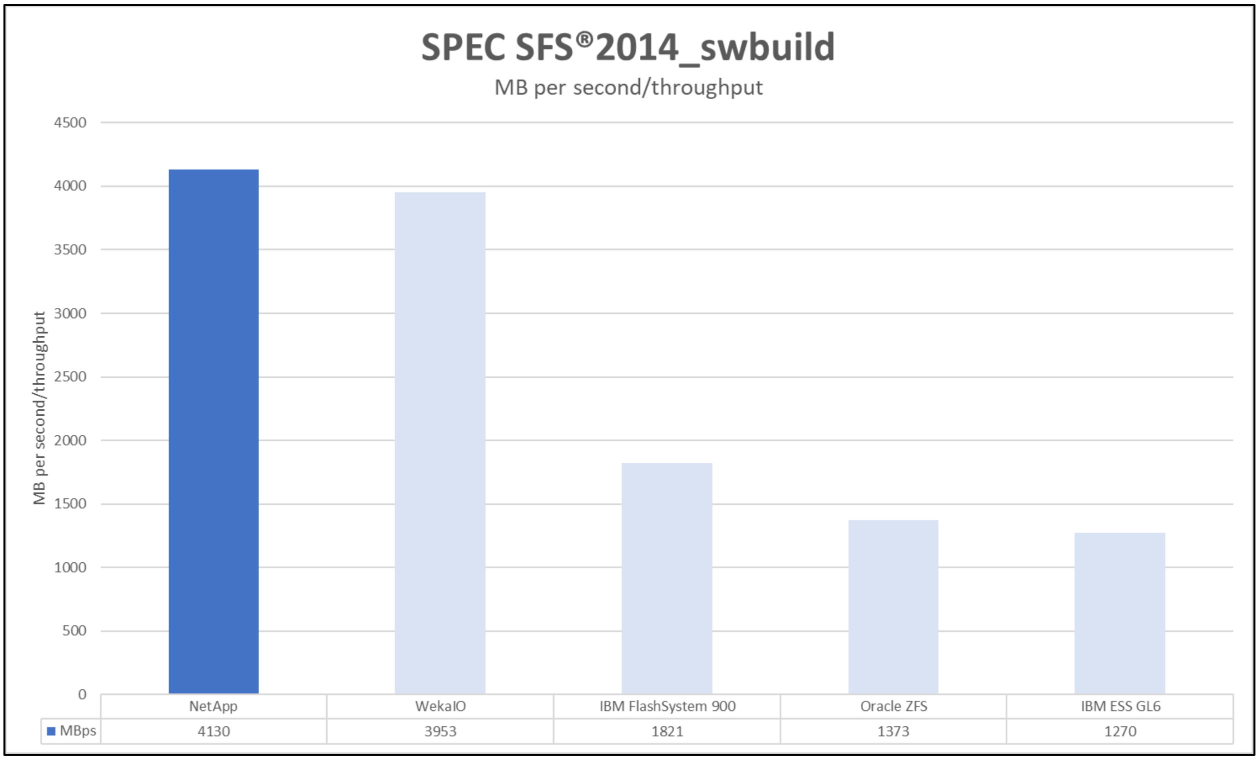
In this test, FlexGroup volumes achieved the lowest ORT ever recorded for a storage system.

Figure 11) Overall response time, SPEC SFS 2014_swbuild submissions.



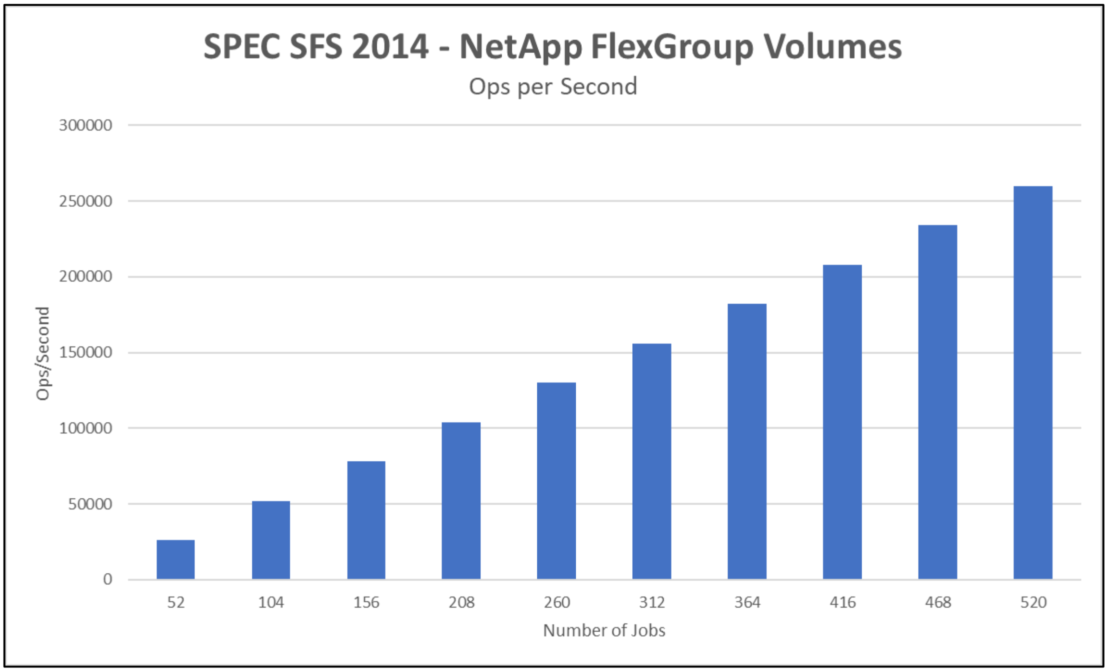
FlexGroup volumes also outperformed other submissions in throughput. In the benchmark, FlexGroup volumes achieved over 4GBps.

Figure 12) Throughput, SPEC SFS 2014_swbuild submissions.



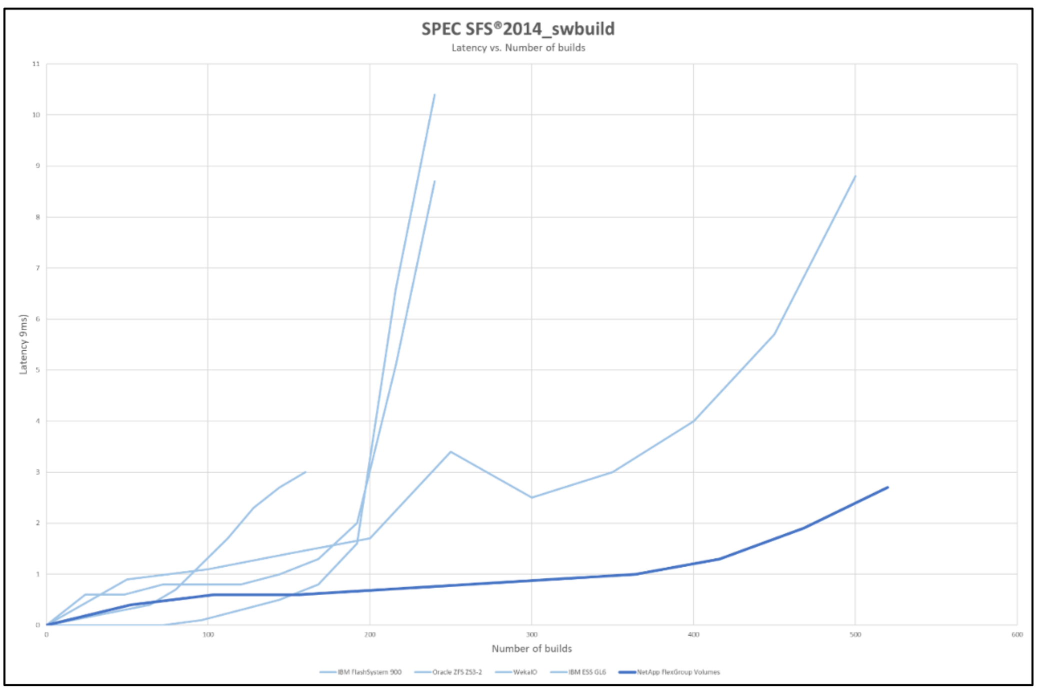
The results of this performance benchmark were achieved with more than 500 concurrent jobs providing 260,000 IOPS.

Figure 13) IOPS, SPEC SFS 2014_swbuild submissions.



If latency is important to your business, FlexGroup volumes also saw the most predictable low latency of all the submissions.

Figure 14) Latency versus number of builds, SPEC SFS 2014_swbuild submissions.



SPEC SFS 2014_swbuild Submission: FlexGroup Volume, ONTAP 9.5

In November of 2018, NetApp submitted a second round of SPEC SFS 2014_swbuild results for publication, this time on an AFF system. NetApp ONTAP using FlexGroup volumes achieved the highest throughput with the most concurrent builds of all other systems by a factor of 2.5 times more than the next highest system. See the following links for the results:

- NetApp four-node AFF A800 with FlexGroup
- NetApp eight-node AFF A800 with FlexGroup

Figure 15, Figure 16, and Figure 17 show the testing results, and comparisons with other submissions.

Figure 15) SPEC SFS 2014_swbuild—concurrent builds.

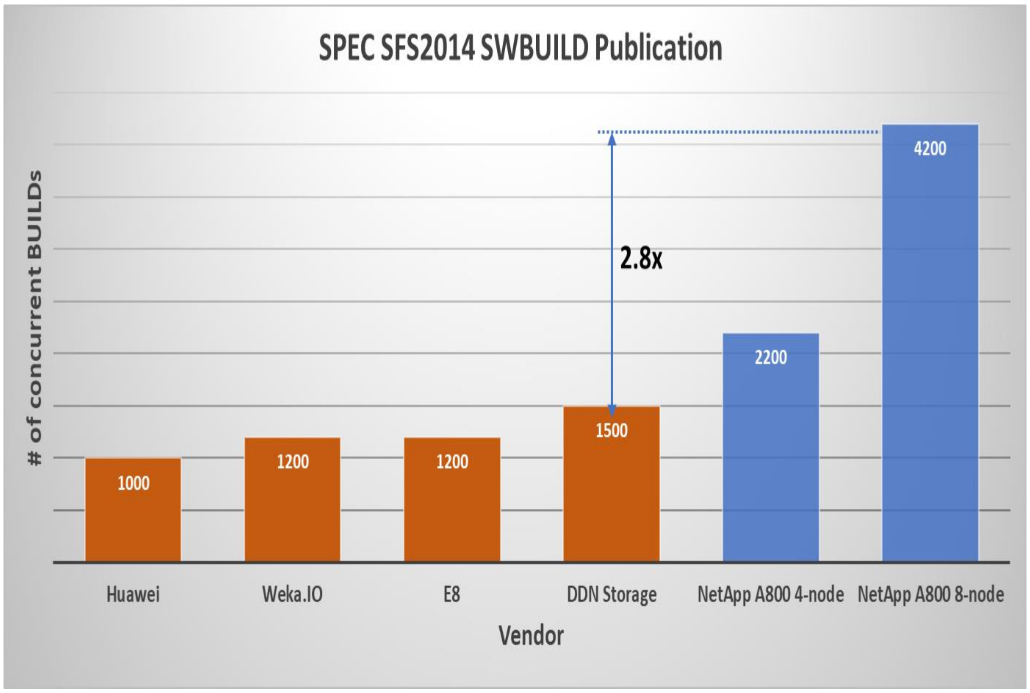


Figure 16) SPEC SFS 2014_swbuild—latency versus build ops/sec.

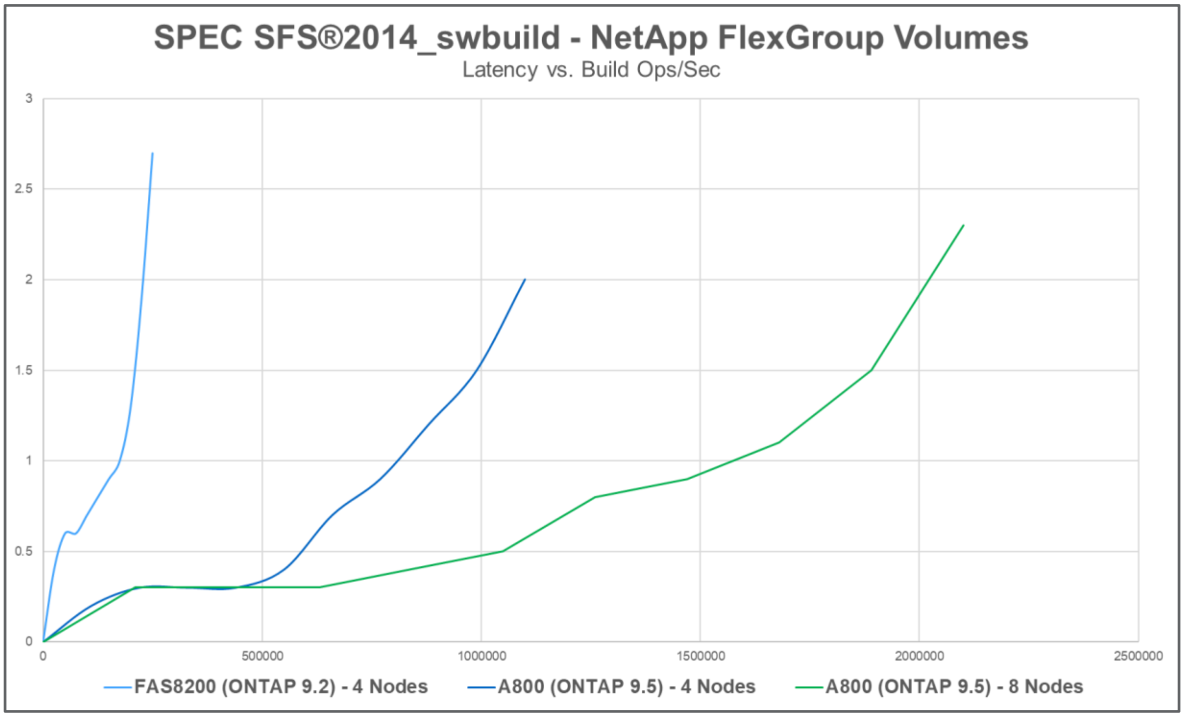
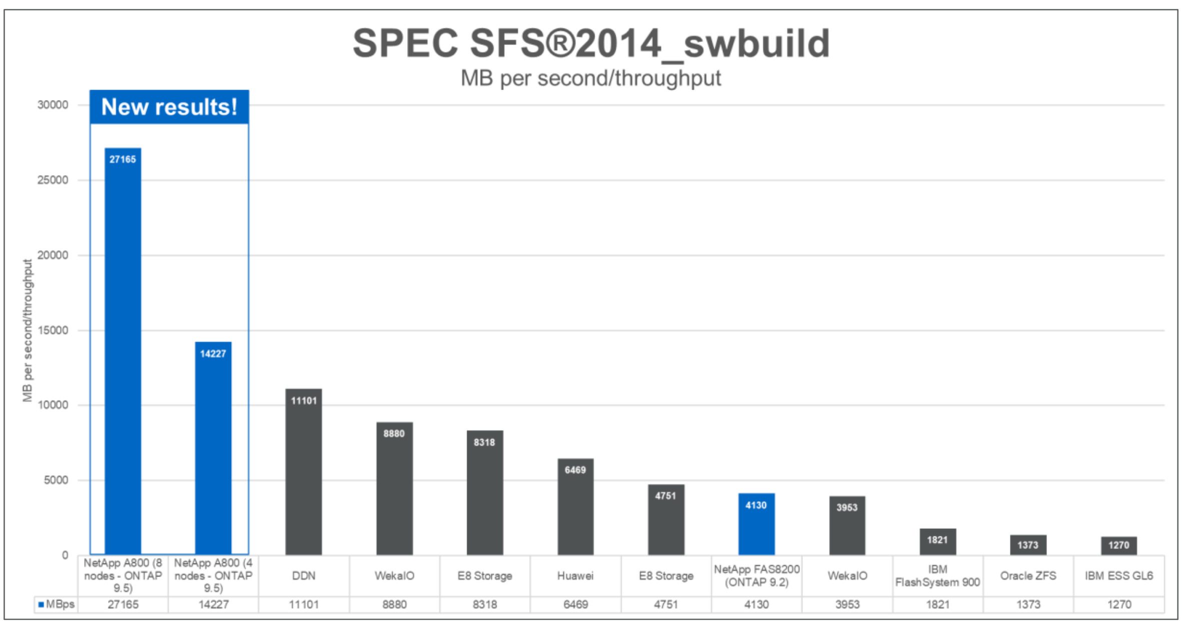


Figure 17) SPEC SFS 2014_swbuild—megabyte per second comparison.



NetApp AFF8080 Testing: ONTAP 9.1

The sample graph in Figure 18 shows latency versus IOPS in a standard NAS workload generator. A single FlexVol volume and a FlexGroup volume with 16-member FlexVol volumes across two nodes are shown. Note how the latency spikes for the FlexVol volume at half the total IOPS that is performed by the FlexGroup volume.

Figure 18) FlexVol versus FlexGroup—standard NAS benchmark test; NFSv3.

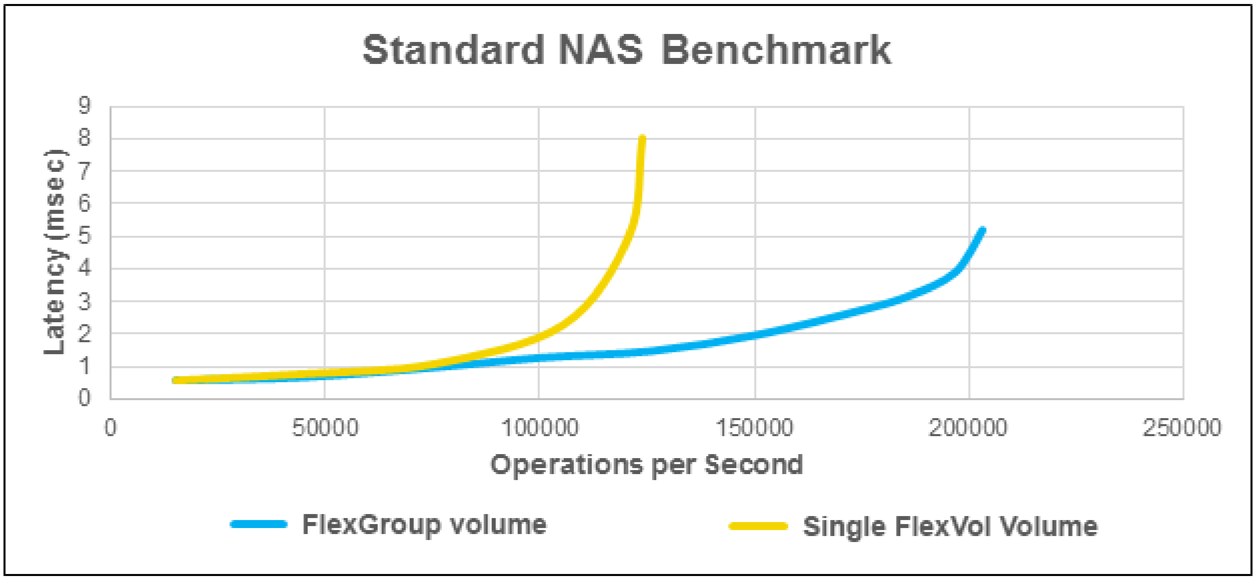


Figure 19 shows a performance comparison between a single FlexVol volume and a single FlexGroup volume. In a simple workload benchmark that used a software build tool (Git), a Linux kernel was compiled with the following hardware kit:

- Four AFF8080 nodes
- A single data aggregate per node, 800GB SSD, 22 data disks

Note: The FlexGroup volume was constructed with eight FlexVol members per node.

The metric was a simple time-to-completion test. In this benchmark, the FlexGroup volume was able to outperform the FlexVol volume by two to six times across multiple Git operations. In addition, the FlexGroup volume was able to push nearly twice the number of gigabytes per second of throughput compared with the single FlexVol volume.

Figure 19) Git benchmark: Linux compile completion time in seconds in FlexGroup volume versus FlexVol volume.

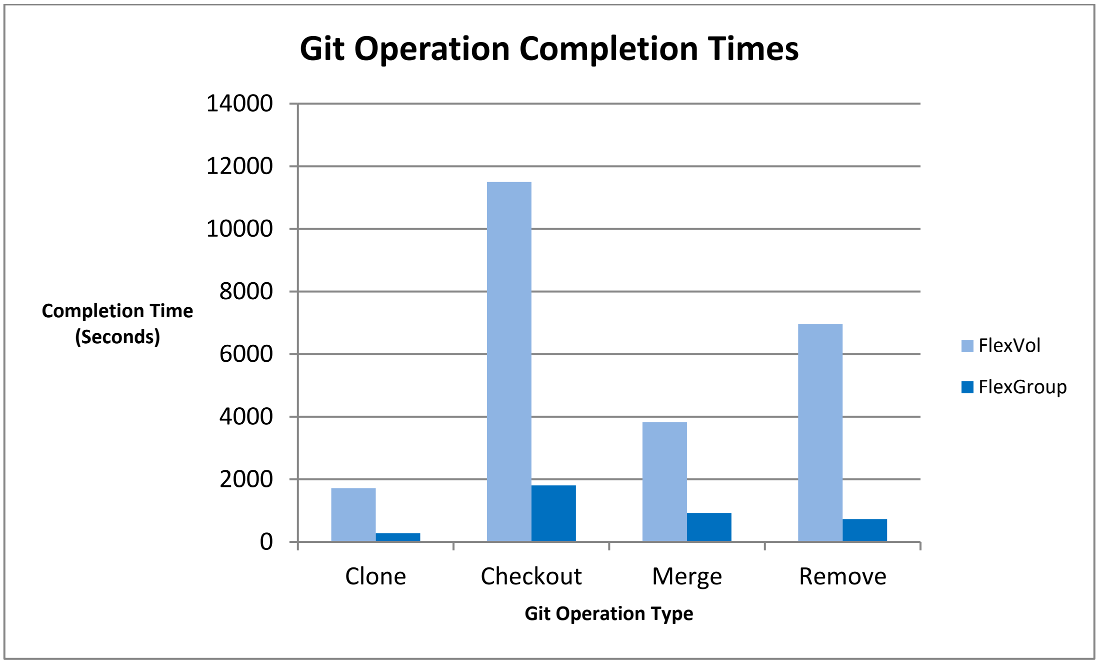
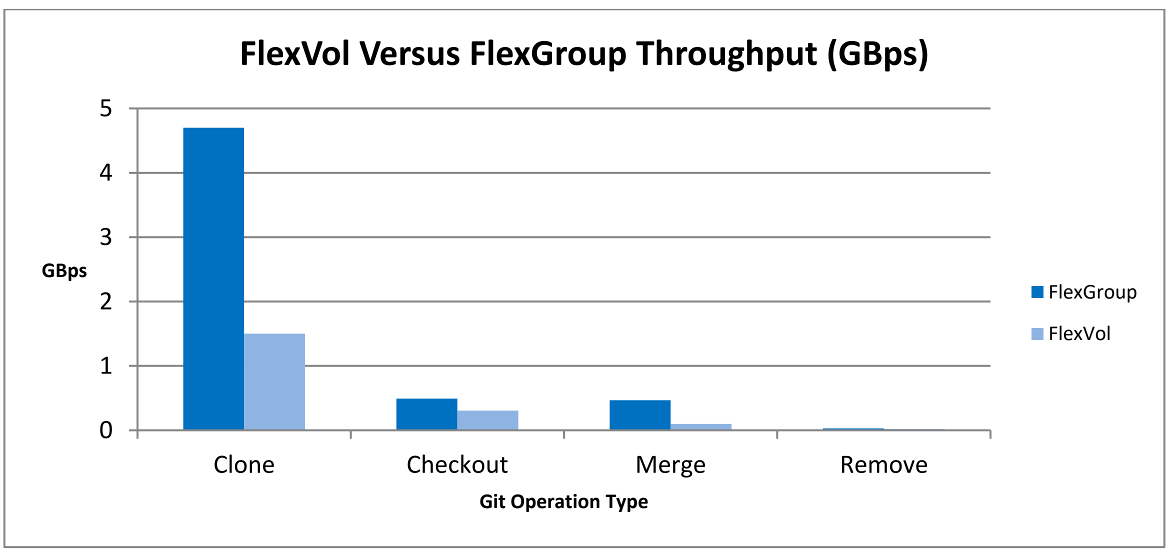


Figure 20 shows the same workloads but measures the overall throughput in gigabytes per second (GBps).

Note: The throughput for a FlexGroup volume is around three times that of a FlexVol volume for the same workload.

Figure 20) Git benchmark: Linux compile; maximum throughput in FlexGroup volume versus FlexVol volume.

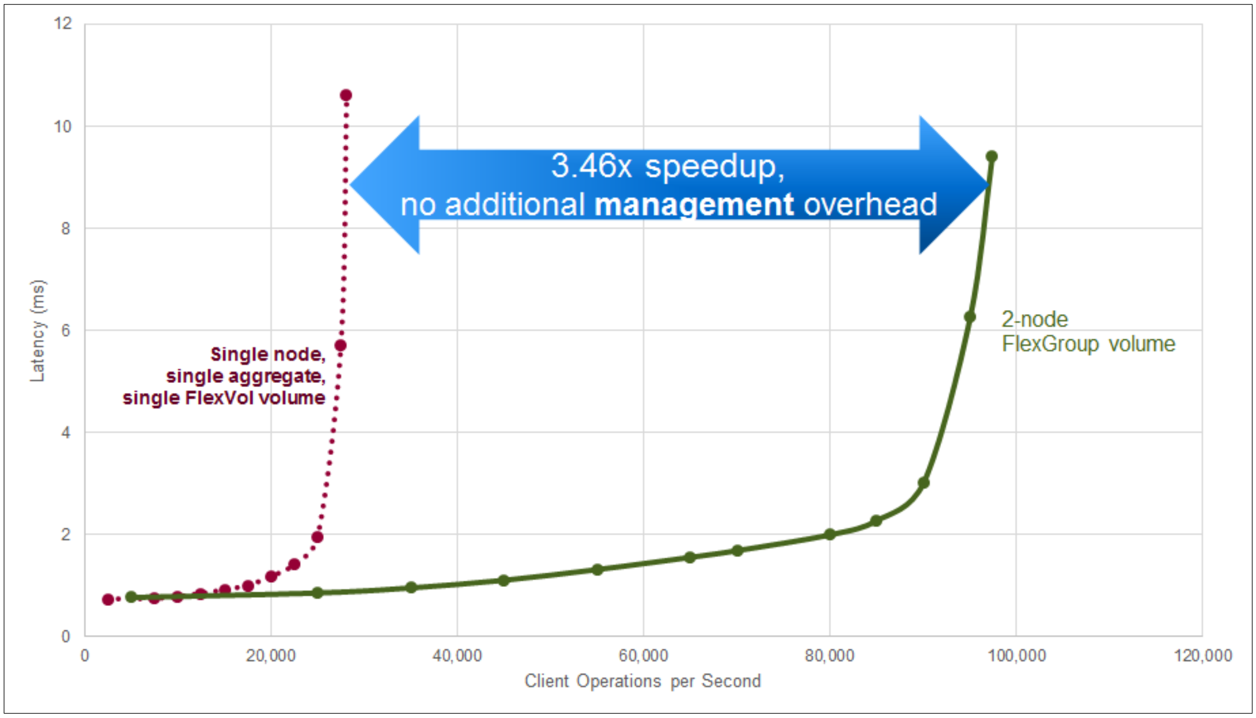


Metadata-intensive workloads, such as those found in common high-performance computing (HPC) applications, benefit particularly from this division because they are often CPU-bound when deployed against a single FlexVol volume.

Standard NAS Benchmark on AFF A300 Testing for SMB2.x: ONTAP 9.3

Figure 21 shows the performance of a standard NAS workload benchmark that used SMB2 on a FlexGroup volume in a two-node cluster versus a single FlexVol volume.

Figure 21) FlexVol volume versus FlexGroup volume: standard NAS benchmark test; SMB2.



AFF A700 Testing

In addition to the four-node AFF8080 tests, the same Git workload was also run on an AFF A700 cluster.

The following configuration was used:

- Two-node AFF A700 cluster
- A single aggregate of 800GB SSDs per node
- FlexVol volume: single node, 100% local
- FlexGroup volume: spans high-availability (HA) pair, eight members per node (16 members total)

The workload was as follows:

- GCC library compile
- Clone operations only (these operations showed the highest maximum throughput for both FlexVol and FlexGroup)
- Four physical servers
- User workloads/threads on the clients that ranged from 4 to 224

Figure 22 compares the maximum achieved throughput (read + write) on Git clone operations on a single FlexVol volume versus a single FlexGroup volume that spanned two nodes.

Note: The maximum throughput of the FlexGroup volume reaches nearly five times the amount of the FlexVol volume without the same degradation of the FlexVol volume as the workload reaches 64 threads.

Figure 22) FlexVol volume versus FlexGroup volume—maximum throughput trends under increasing workload.

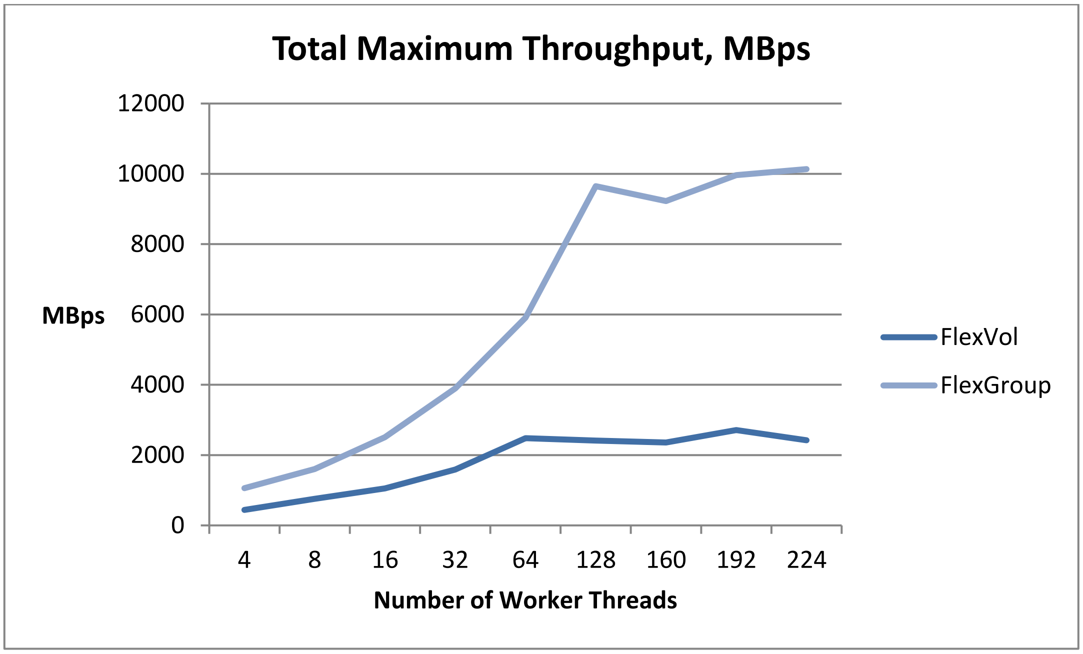


Figure 23 compares a FlexVol volume and a FlexGroup volume in the same configurations. This time, we break down the maximum read and write throughput individually, as well as comparing that against the average throughput for the FlexVol volume and the FlexGroup volume.

Figure 23) FlexVol volume versus FlexGroup volume—maximum throughput trends under increasing workload, detailed.

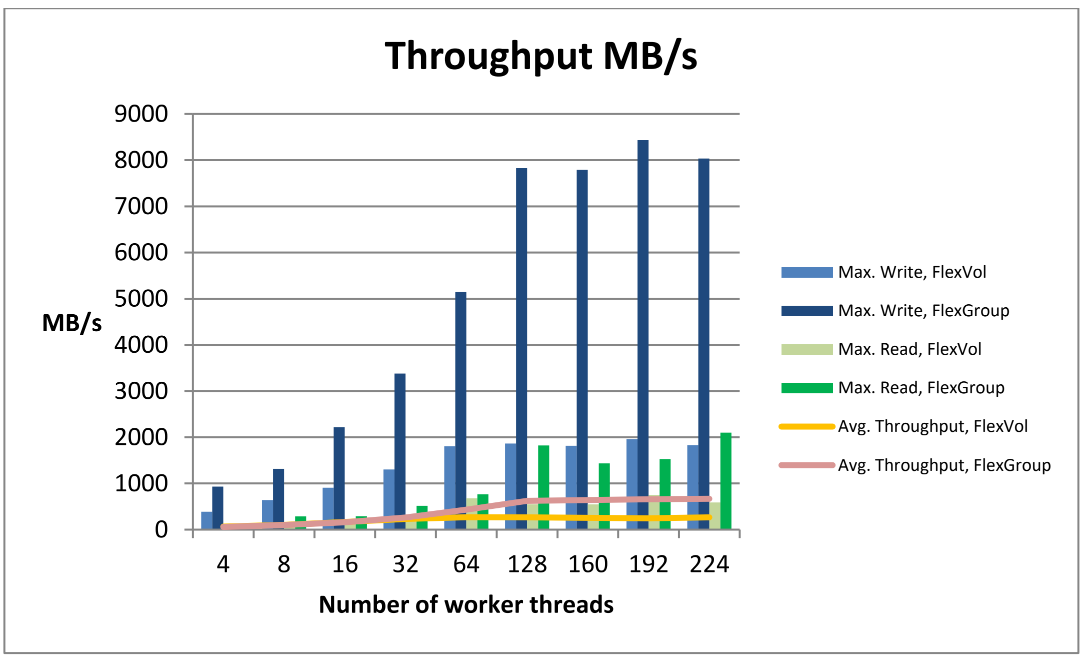
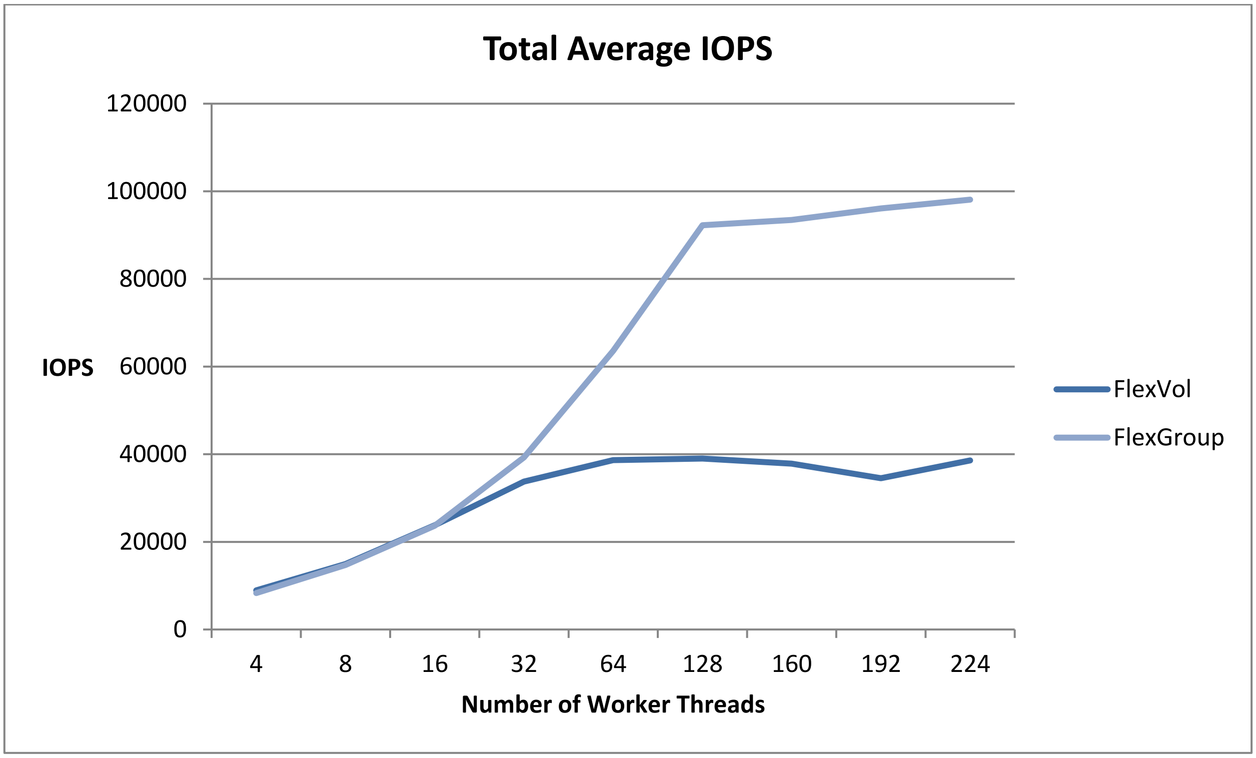


Figure 24 shows the maximum total average IOPS for a FlexGroup volume versus a FlexVol volume on the AFF A700. Again, note the dramatic increase of IOPS for the FlexGroup volume versus the degradation of IOPS at 64 threads for the FlexVol volume.

Figure 24) FlexVol volume versus FlexGroup volume—maximum average total IOPS.



ONTAP 9.4 and 9.5 Performance Testing

For ONTAP versions 9.4 and 9.5, we ran a set of performance tests using standard NAS benchmark suites that simulates both EDA and software build workloads. The goal was to show that ONTAP improves performance with each release.

As discussed earlier in "NetApp AFF8080 Testing – ONTAP 9.1," we ran standard NAS benchmark tests for a four-node AFF8080 cluster running ONTAP 9.1 for software builds. Our tests showed that the total IOPS achieved before latency started to spike was around 150,000–200,000.

The ONTAP 9.4 and 9.5 tests featured the following configurations:

- An AFF A700s cluster
- A FlexGroup volume spanning a single node and two nodes
- 14 NFSv3 clients
- 32 10GB LIFs (16 LIFs per node)
- 32 mount points on each client

The following graphs show that performance in a FlexGroup volume can scale, and that each release provides substantial performance improvements in ONTAP. These improvements can be accomplished with a nondisruptive upgrade.

Figure 25) Standard NAS benchmark (EDA)—ONTAP 9.5: one node versus two nodes (operations/sec).

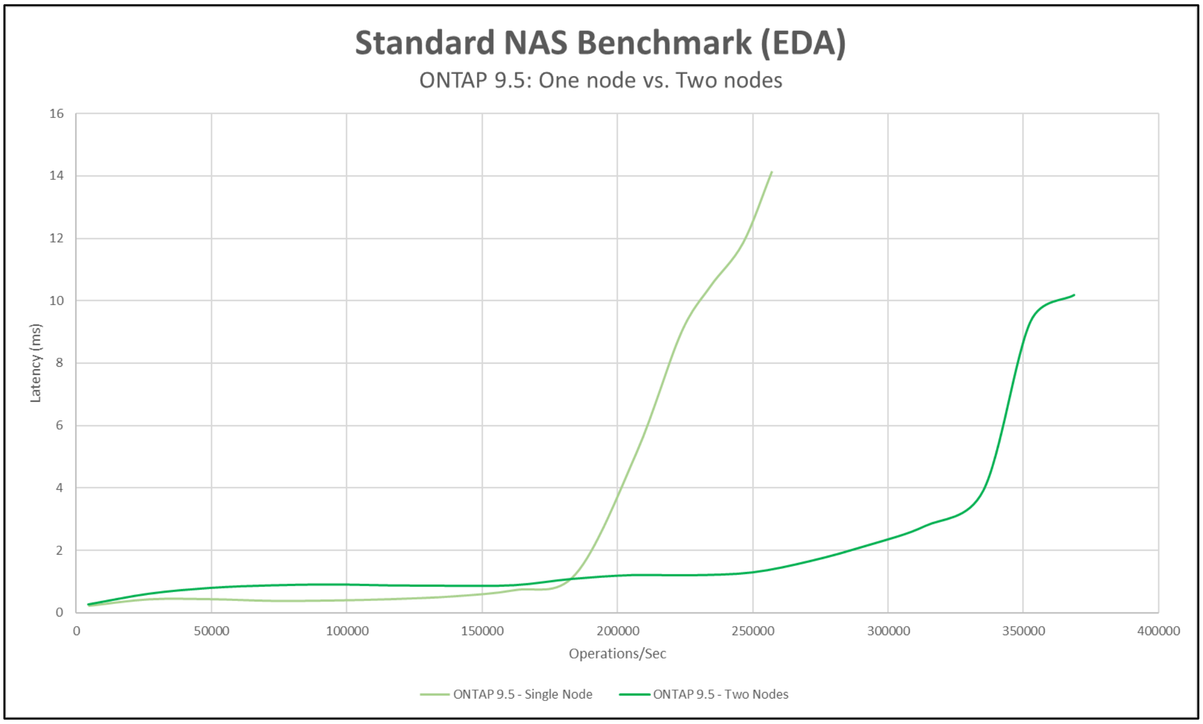


Figure 26) Standard NAS benchmark (EDA)—ONTAP 9.4 versus ONTAP 9.5 (operations/sec).

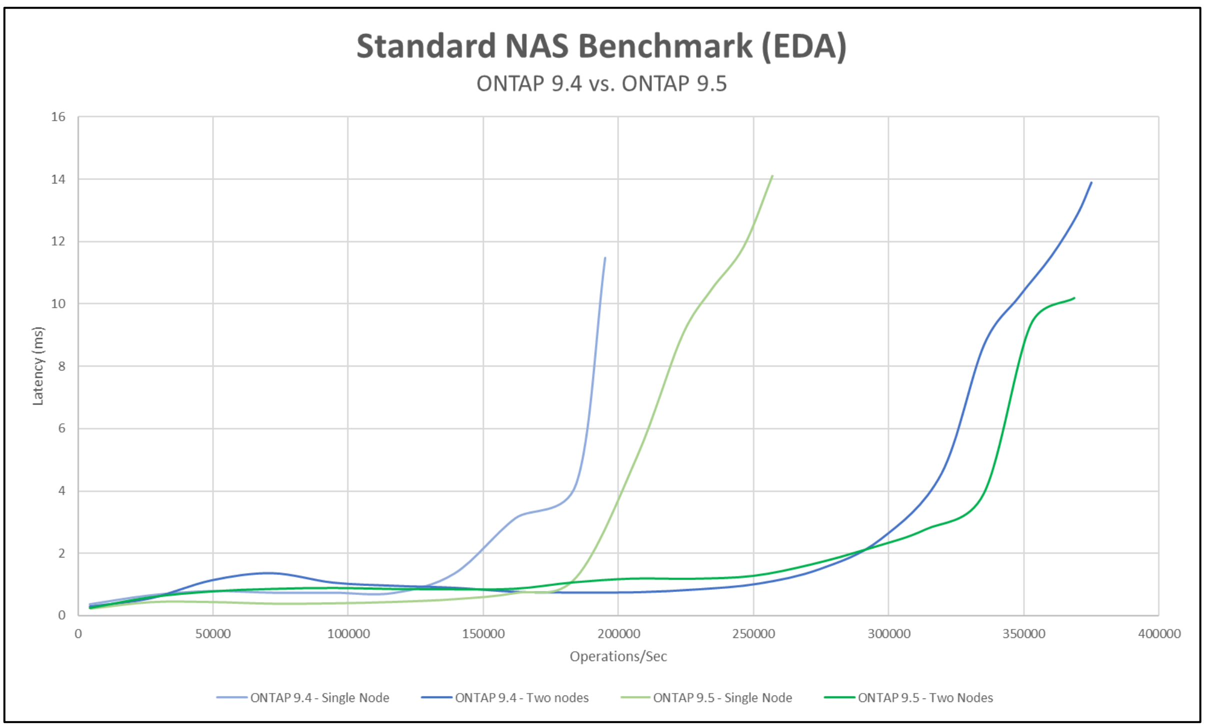


Figure 27) Standard NAS benchmark (EDA)—ONTAP 9.5: one node versus two nodes (MBps).

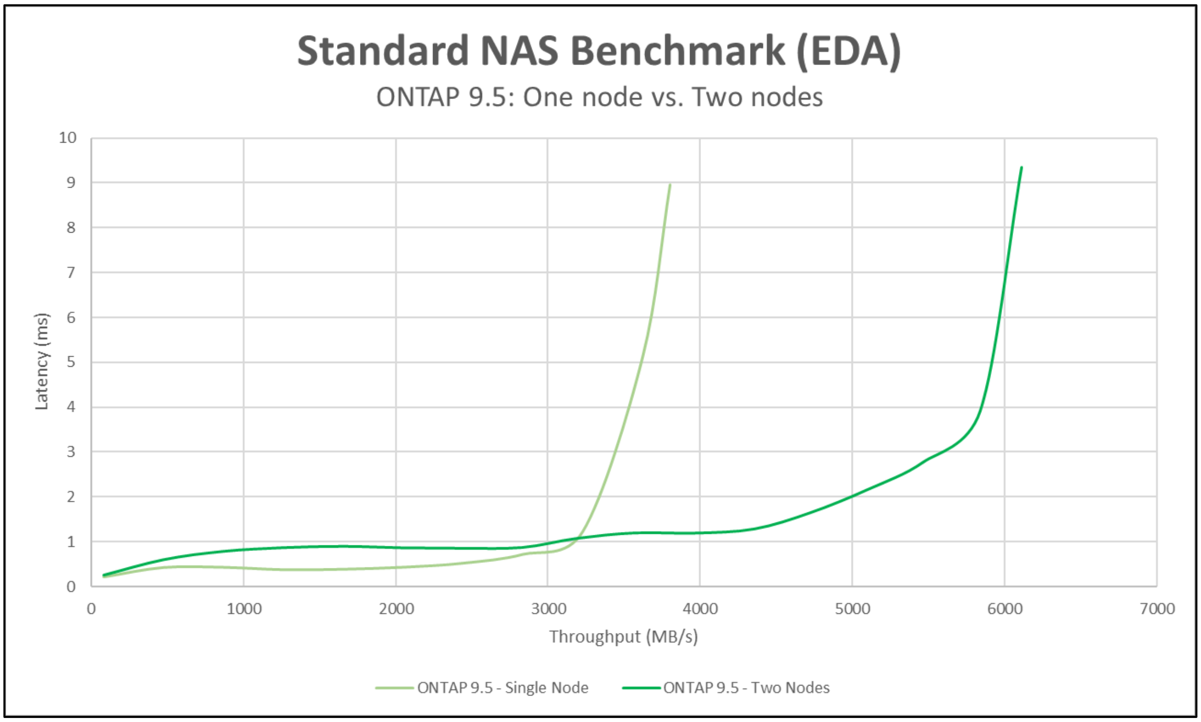


Figure 28) Standard NAS benchmark (EDA)—ONTAP 9.4 versus ONTAP 9.5 (MBps).

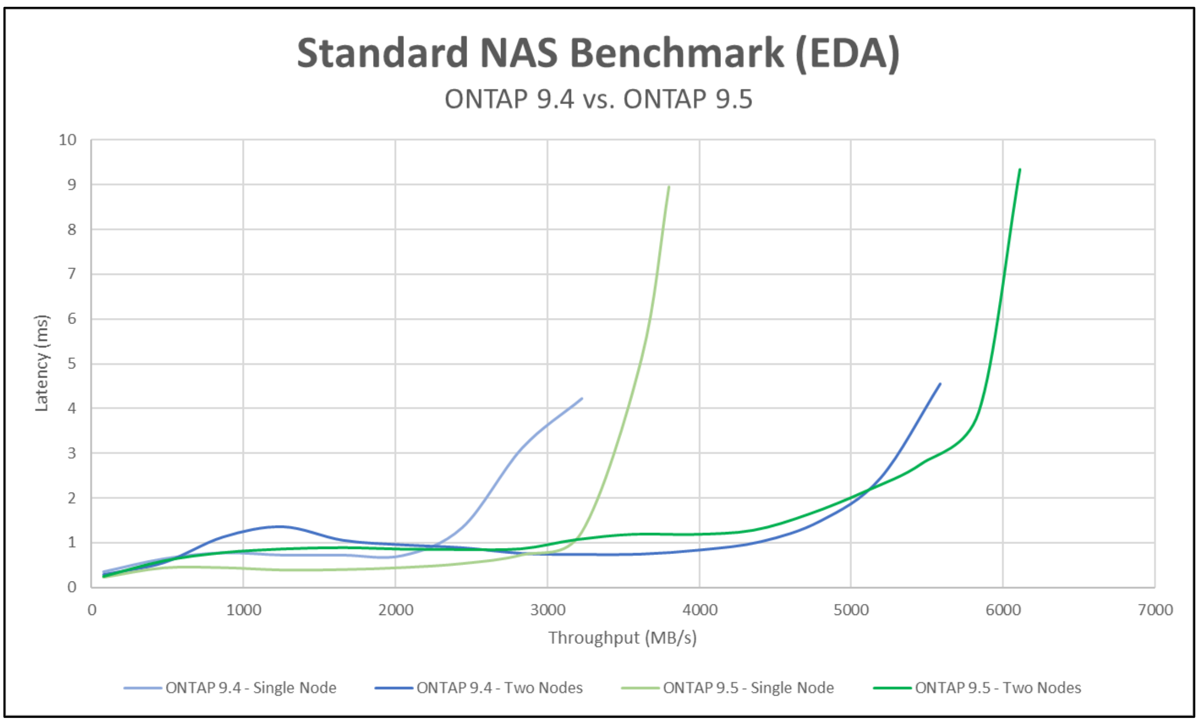


Figure 29 shows the performance for a standard NAS benchmark running a software build workload (such as Git or Perforce). Both types of workloads are ideal for FlexGroup volumes because of the high file ingest rates and need for parallel processing of write metadata.

Figure 29) Standard NAS benchmark (software builds)—ONTAP 9.5 (operations/sec).

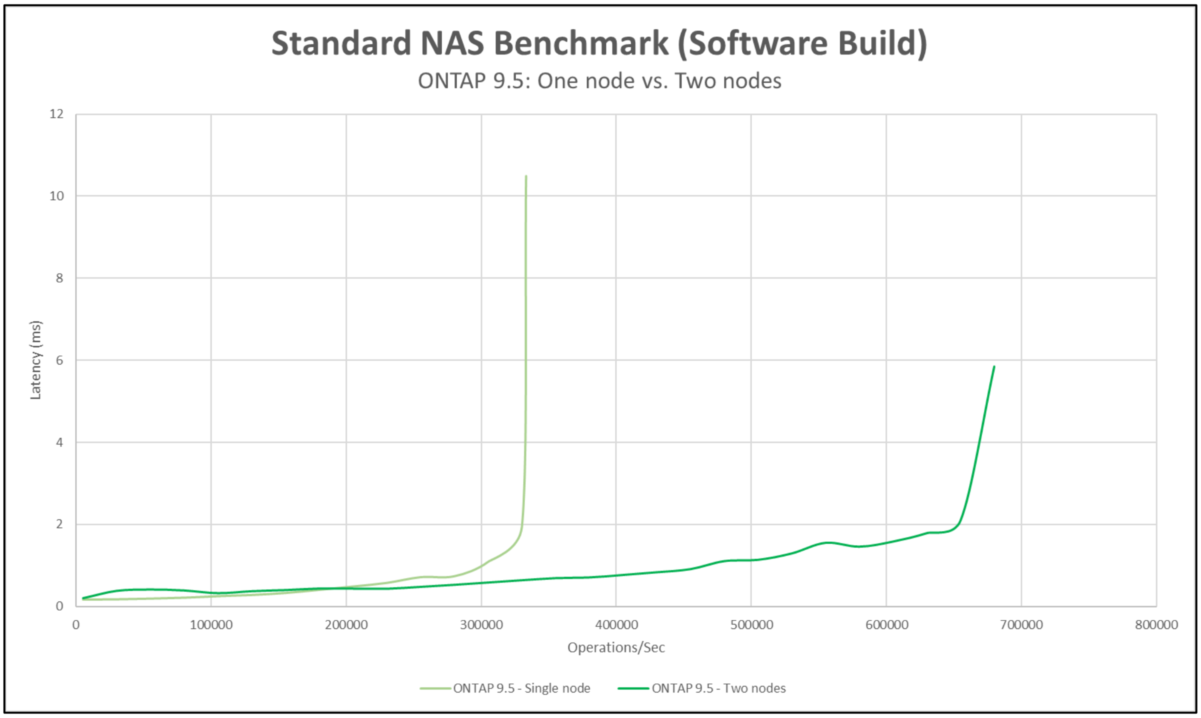


Figure 30) Standard NAS benchmark (software builds)—ONTAP 9.4 versus ONTAP 9.5 (operations/sec).

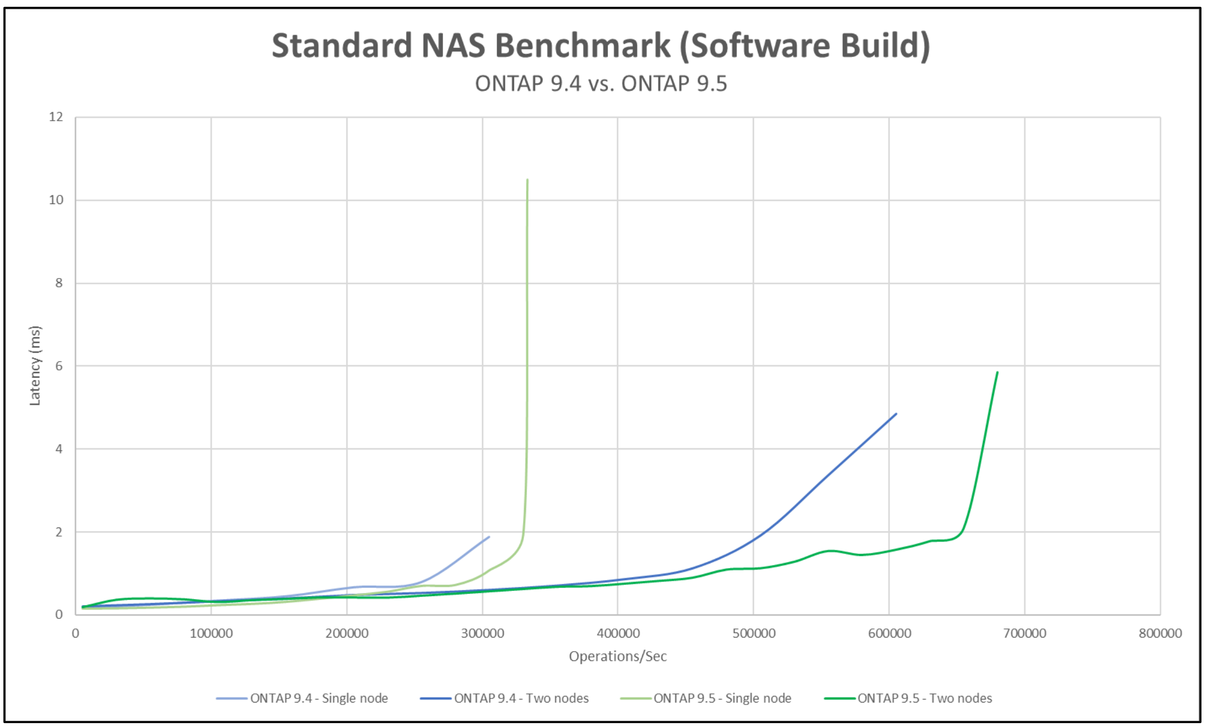


Figure 31) Standard NAS benchmark (software builds)—ONTAP 9.5 (MBps).

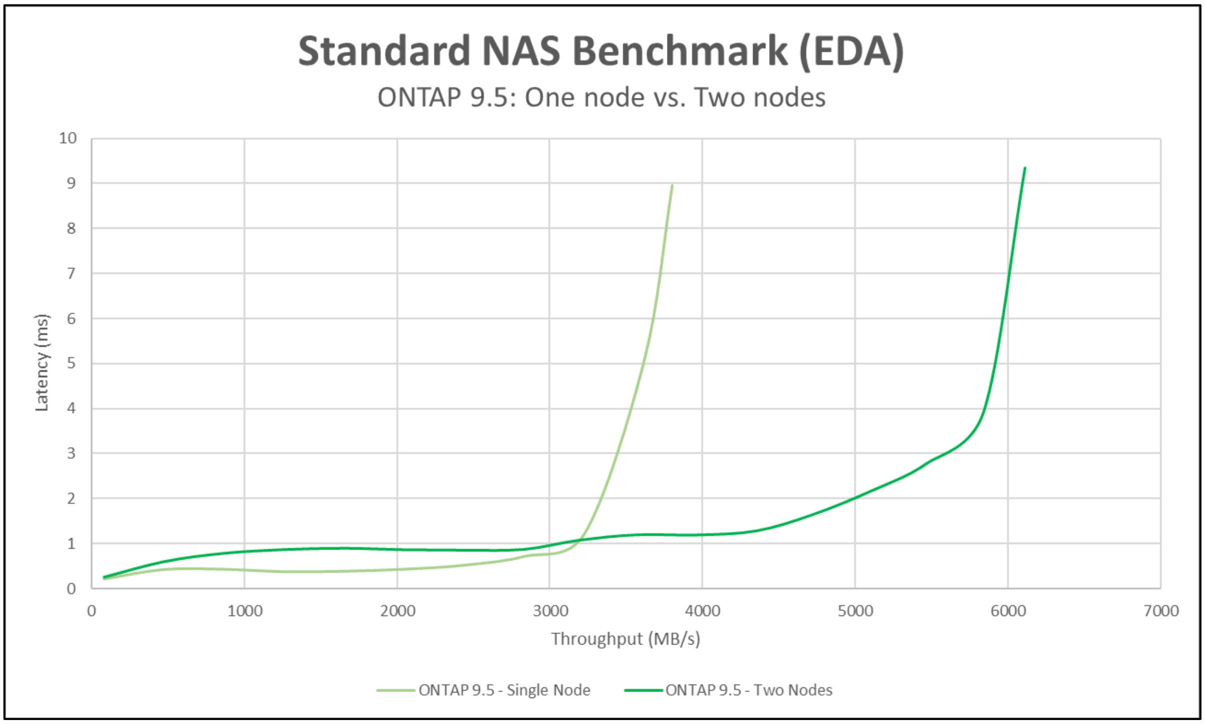
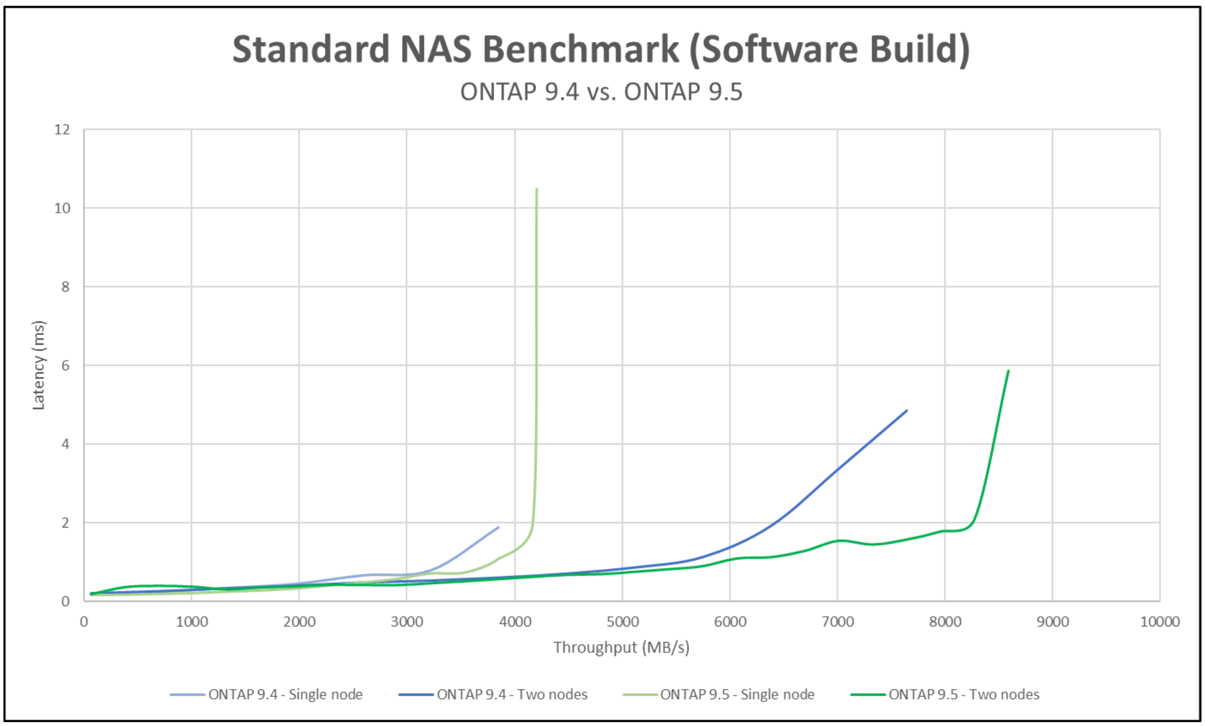


Figure 32) Standard NAS benchmark (software builds)—ONTAP 9.4 versus ONTAP 9.5 (MBps).



FlexGroup Performance with Big Data Workloads

Due to the FlexGroup volume's capacity and ability to scale a single namespace across multiple compute nodes in a cluster, it provides an interesting use case for big data workloads, such as Apache Hadoop, Splunk, and Apache Spark. These applications generally expect only one or two directories to dump large amounts of data and high file counts, requiring high throughput at a low latency. FlexVol volumes were able to accomplish this performance, but not without some tweaks to the application to make it aware of multiple volumes.

Although ONTAP does not natively support Hadoop Distributed File System (HDFS) on NAS, NetApp does offer an NFS connector that can provide NFS connectivity to big data workloads. TR-4570 describes an example of using Apache Spark on FlexGroup volumes.

Also, the NetApp Customer Proof-of-Concept (CPOC) lab conducted some performance testing using the TeraSort benchmark, which is used to test Apache big data workloads. In this testing, a two-node AFF A700 cluster running ONTAP 9.2 was used to push a maximum of 8GBps in and out of the cluster at an average read latency from ~3ms to 5ms and an average write latency from ~4ms to 8ms, while keeping the average CPU utilization around 55% on both nodes. Using a FlexGroup volume with big data workloads allows all available hardware to be used and provides a way to nondisruptively scale the capacity and performance by adding nodes for the workload as needed.

Figure 33) TeraSort benchmark statistics summary on a FlexGroup volume.

Node 1

Sysstat System Utilization

Min	Avg	Max	σ	Cv	Counter
13	42	75	12.4	0.3	cpu
3,447	25,991	54,412	9,913.2	0.4	nfs
752,703	2,179,080	3,572,454	597,657.6	0.3	net_kb_in
828,342	2,417,590	4,107,351	710,786.3	0.3	net_kb_out

Node 2

Sysstat System Utilization

Min	Avg	Max	σ	Cv	Counter
24	56	84	13.3	0.2	cpu
14,403	38,608	68,636	10,024.7	0.3	nfs
254,777	2,216,472	4,283,503	766,327.7	0.4	net_kb_in
340,437	2,392,841	4,439,274	784,810.8	0.3	net_kb_out

As a bonus, big data workloads running on ONTAP FlexGroup volumes have shown a space savings of nearly 50% with storage efficiency features such as inline aggregate deduplication, inline data compaction, and inline compression.

Automatic Workload Adaptation

The FlexGroup volume continually adapts to the current conditions in the cluster, changing behavior constantly to keep usage evenly consumed and to keep dynamic load evenly balanced. Trade-offs are implicit in this continual balancing act. The cost of this automatic balancing is that a FlexGroup volume cannot attain the same theoretical maximum performance that a perfectly balanced and manually organized collection of FlexVol volumes could otherwise attain. However, the FlexGroup volume can reach very close to that maximum, and it requires no foreknowledge of the workload to accomplish its work. In addition, a FlexGroup volume adds a simplicity aspect to large data layouts that a single FlexVol architecture cannot.

FlexGroup volumes perform better—balancing load and usage more smoothly—when faced with a broad variety of workloads and high data-creation rates. Thus, a single FlexGroup volume that performs many different roles can be a more effective use of your cluster's resources than if you use different FlexGroup volumes for different workloads. You can, however, junction multiple FlexVol volumes and FlexGroup volumes together in the same ONTAP SVM.

ONTAP 9.7 introduces new ingest algorithm enhancements to recognize workload types (high metadata ingest versus streaming I/O). If a workload is creating a high number of small files, then the FlexGroup volume will place those files to balance them evenly across volumes while favoring folder locality to increase performance. If the workload is a smaller number of large files, then ONTAP will recognize that difference. Rather than favoring local folder placement (which could result in multiple large files ending up on the same member volume and creating an artificial imbalance of data), ONTAP will instead place files in a more round-robin fashion to ensure even space allocation. This allows for a wider variety of workloads to perform optimally on FlexGroup volumes, preventing space imbalance scenarios and reducing the need for administrator intervention.

Ingest Algorithm Improvements

Every ONTAP release further improves the ingest algorithms for FlexGroup volumes that help ONTAP make better decisions about how new data is placed in FlexGroup volumes. The algorithms also improve the way FlexGroup volumes respond when member volumes approach "nearly full" status.

Therefore, NetApp strongly recommends that you run the latest patched ONTAP version when using FlexGroup volumes. You can download the latest release at NetApp Support for ONTAP 9.

Quality of Service (QoS) Maximums

Starting in ONTAP 9.3, you can apply maximum storage QoS policies to help prevent a FlexGroup volume from acting as a bully workload in ONTAP. Storage QoS can help you manage risks around meeting your performance objectives. You use storage QoS to limit the throughput to workloads and to monitor workload performance. You can reactively limit workloads to address performance problems, and you can proactively limit workloads to prevent performance problems. For more information about storage QoS, see TR-4211: Storage Performance Primer.

How Storage QoS Maximums Work with FlexGroup

With FlexGroup, storage QoS policies are applied to the entire FlexGroup volume. Because a FlexGroup volume contains multiple FlexVol member volumes and can span multiple nodes, the QoS policy gets shared evenly across nodes as clients connect to the storage system. Figure 34 and Figure 35 show how storage QoS gets applied to a FlexGroup volume.

Figure 34) Storage QoS on FlexGroup volumes—single-node connection.

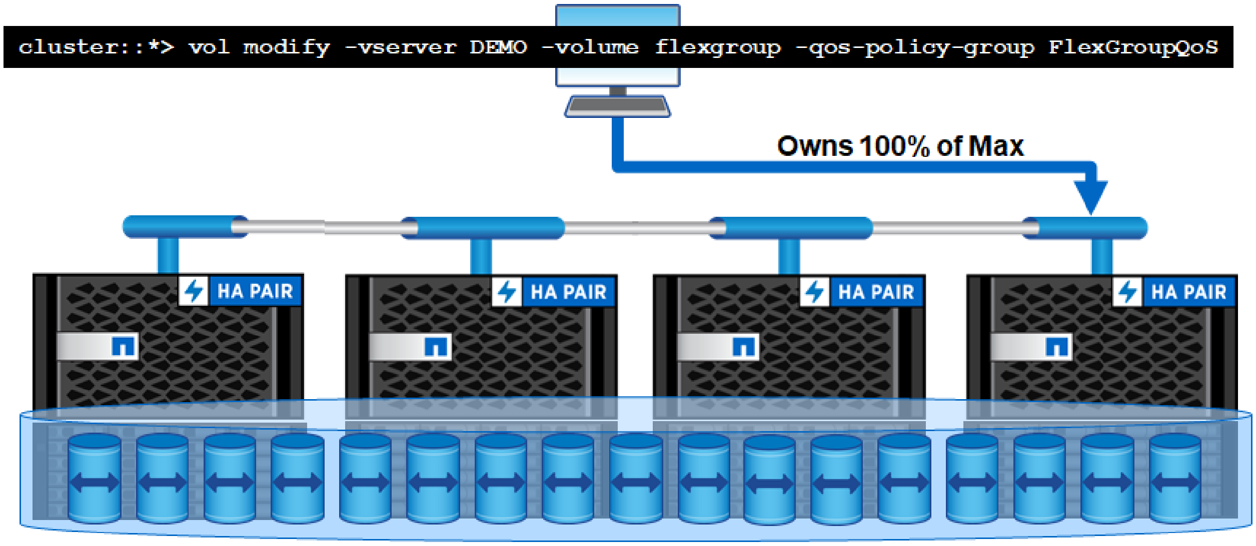
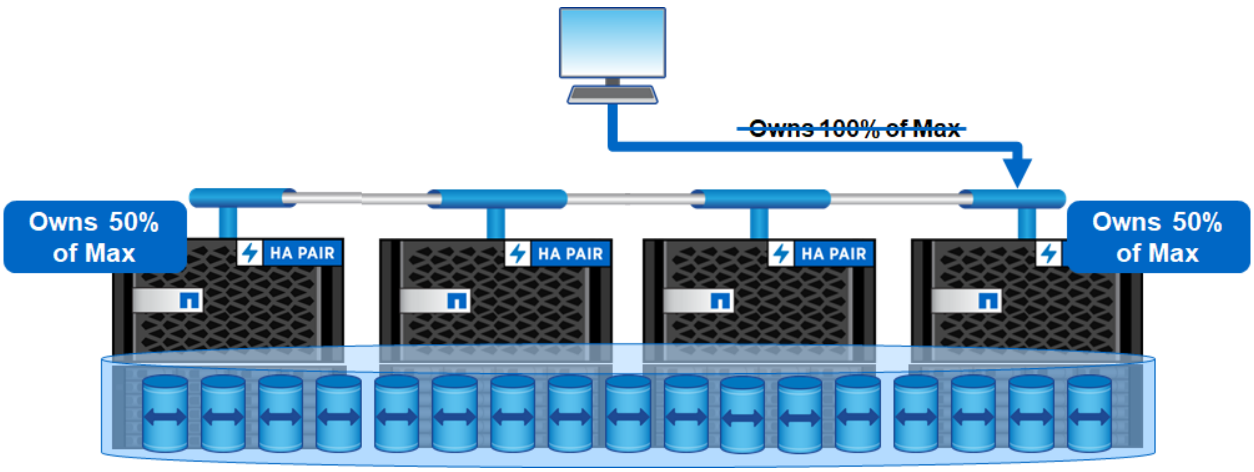


Figure 35) Storage QoS on FlexGroup volumes: multinode connection.



Storage QoS Considerations with FlexGroup Volumes

Storage QoS is applied at the FlexGroup volume level. File-level QoS and nested policies are currently not supported with FlexGroup volumes.

QoS Minimums

ONTAP 9.4 added support to FlexGroup volumes for QoS minimums (also referred to as guarantees or floors), which provide a set threshold of performance that is allocated to a specified object. This feature is supported for AFF systems only. For details on the feature, see TR-4211: Storage Performance Primer.

Adaptive QoS

ONTAP 9.4 also introduced adaptive QoS support for FlexGroup volumes, which allows ONTAP to adjust the IOPS/TB values of a QoS policy as the volume capacity is adjusted. This feature is covered in detail in TR-4211: Storage Performance Primer.

REST APIs

REST API support was introduced in ONTAP 9.6. Rather than navigating a proprietary interface (such as the NetApp Manageability SDK), REST APIs enable you to use a universal standard for accessing and interacting with a cluster.

You can find REST API documentation at `http://[your_cluster_IP_or_name]/docs/api`. This site provides examples and an interactive "try it out" feature that enables you to generate your own REST APIs.

For example, to create a FlexGroup volume, you can use the `POST` REST API under `/storage/volumes`. What makes a FlexGroup a FlexGroup (and not a FlexVol) in this call are one or a combination of the following values:

- **Aggregates.** If you specify more than one, then the REST API creates a FlexGroup volume. This is the same behavior as `-aggr-list` in the CLI.
- **constituents_per_aggregate.** Specifies the number of times to iterate over the aggregates listed with `aggregates.name` or `aggregates.uuid` when a FlexGroup volume is created or expanded. If a volume is being created on a single aggregate, the system creates a flexible volume if the `constituents_per_aggregate` field is not specified; it creates a FlexGroup volume if this field is specified. If a volume is being created on multiple aggregates, the system will always create a FlexGroup volume. This is the same behavior as `-aggr-list-multiplier` in the CLI.
- **Style.** If you specify `style` as "flexgroup" and don't set the `constituents_per_aggregate` value or more than one aggregate, ONTAP automatically provisions a FlexGroup volume of four members per aggregate. This is the same behavior as `-auto-provision-as` in the CLI.

In the REST API documentation, the "try it out" functionality helps guide you as you try to create the correct REST API strings. When you make a mistake, the interface delivers error messages and a list of error codes. Also, a job string URL is given if the REST API command is correct but the job fails for another reason (such as creating a FlexGroup volume that has members that are too small) You can access the job string can be accessed through the browser by using http://[your_cluster_IP_or_name]/api/cluster/jobs/job_uuid].

This is what a failure message might look like:


```
{
  "uuid": "b5b04f0b-82ea-11e9-b3aa-00a098696eda",
  "description": "POST /api/storage/volumes/b5b02a66-82ea-11e9-b3aa-00a098696eda",
  "state": "failure",
  "message": "Unable to set parameter \"-min-autosize\" to specified value because it is too
small. It must be at least 160MB (167772160B).",
  "code": 13107359,
  "start_time": "2019-05-30T10:53:39-04:00",
  "end_time": "2019-05-30T10:53:39-04:00",
  "_links": {
    "self": {
      "href": "/api/cluster/jobs/b5b04f0b-82ea-11e9-b3aa-00a098696eda"
    }
  }
}
```


And here is what a "success" job would look like:


```
{
  "uuid": "ac2155d1-82ec-11e9-b3aa-00a098696eda",
  "description": "POST /api/storage/volumes/ac2131c5-82ec-11e9-b3aa-00a098696eda",
  "state": "success",
  "message": "success",
  "code": 0,
  "start_time": "2019-05-30T11:07:42-04:00",
  "end_time": "2019-05-30T11:07:46-04:00",
  "_links": {
```



```
    "self": {
      "href": "/api/cluster/jobs/ac2155d1-82ec-11e9-b3aa-00a098696eda"
    }
  }
}
```

For a sample REST API string that creates a FlexGroup volume, see the Command Examples section of this document.

NetApp MetroCluster

ONTAP 9.6 introduced support for FlexGroup volumes on MetroCluster deployments (FC and IP).

MetroCluster software is a solution that combines array-based clustering with synchronous replication to deliver continuous availability and zero data loss at the lowest cost. There are no stated limitations or caveats for FlexGroup volumes with MetroCluster.

For more information about MetroCluster, see TR-4705: NetApp MetroCluster Solution Design and Architecture.

NetApp Cloud Volumes ONTAP

ONTAP 9.6 introduced official support for Cloud Volumes ONTAP—an ONTAP solution running in the cloud. You can now deploy a FlexGroup volume using Cloud Volumes ONTAP, but only by using the CLI or ONTAP System Manager. Currently, you cannot use Cloud Manager to create FlexGroup volumes.

FlexGroup volumes running in Cloud Volumes ONTAP can use the same feature sets available in the ONTAP version deployed to the Cloud Volumes ONTAP instance. Some common use cases seen for Cloud Volumes ONTAP and FlexGroup include:

- Data lake for analytics
- EDA repositories for use with Amazon Elastic Compute Cloud (Amazon EC2) instances
- Data backup and archive for use with on-premises SnapMirror

Although FlexGroup volumes are able to support multiple petabytes in a single namespace for on-premises deployments, Cloud Volumes ONTAP instances max out at 368TB per instance and FlexGroup volumes cannot span more than one instance. Also, creating a FlexGroup volume currently requires use of System Manager or the CLI. For more information about Cloud Volumes ONTAP, see Cloud Volumes ONTAP Enterprise Data Management Solution.

7.4 Workloads and Behaviors

In an optimal FlexGroup volume, all constituents have roughly the same amount of data and load, and it can maintain that state while using a high frequency of local placement for best performance. A less optimal FlexGroup volume might have some constituents that hold more or less data than their peers, or that are receiving much more or much less traffic.

Also undesirable is a FlexGroup volume that appears to be perfectly balanced in usage and load, but that has had to resort to remote placement frequently to maintain that state. This situation can occur when FlexGroup member volumes get closer to being 100% used. Remote placement encourages data distribution, at the cost of giving up some performance (roughly 5% to 10%). However, the performance loss caused by remote placement is more than made up for by the concurrency gains of a multimember FlexGroup volume versus a single FlexVol volume.

Workloads determine the degree to which a FlexGroup volume behaves optimally. A FlexGroup volume responds well to some types of workloads, but it might struggle to accommodate others. Understanding what works well with a FlexGroup volume can help you determine what traffic should be sent there for optimal results.

Optimal Workloads

In general terms, a FlexGroup volume works optimally when it is under heavy ingest load—that is, when there is a high rate of creating files and directories. ONTAP makes its placement decisions as new files and directories are created, so the more often this action occurs, the more frequently ONTAP has an opportunity to correct existing imbalances in load and usage.

- **FlexGroup volumes work best with numerous small subdirectories.** This means dozens to hundreds of files per directory, because they allow the FlexGroup volume to place new child subdirectories remotely while keeping individual files local to their parent directories for best performance.
- **A FlexGroup volume responds well to heavy concurrent traffic.** Bringing more workloads—especially traffic from multiple clients that are doing different things at the same time—to bear against a FlexGroup volume simultaneously can improve its overall performance. In other words, don't expect to push a FlexGroup volume to its limit with only a few clients.
- **A FlexGroup volume works best when there is plenty of free space.** When constituents begin to fill up, the FlexGroup volume begins to employ remote placement more frequently so that no one constituent becomes full before its peers do. This increased usage of remote placement comes with a metadata performance penalty.
- **FlexGroup volumes work best with high rates of write metadata operations.** ONTAP FlexVol volumes already process read and write I/O in parallel, and metadata read operations (such as `GETATTR`). However, ONTAP processes write metadata (such as `SETATTR` and `CREATE`) serially, which can create bottlenecks on normal FlexVol volumes. FlexGroup volumes provide a parallel processing option for these types of workloads, which results in performance that is two to six times better for these types of workloads.

Performance and Capacity Considerations

For best performance, keep plenty of free space on the FlexGroup volume (at least 50GB on each constituent, although more is better) when it's under heavy load. For more information about space discrepancy and the concept of tolerance in a FlexGroup volume, see TR-4557, NetApp ONTAP FlexGroup Volumes: A Technical Overview. To help prevent situations in which a member volume runs out of space, enable volume autogrow (available in ONTAP 9.3 and later). Additionally, take advantage of elastic sizing in ONTAP 9.6 and later.

Free space for FlexVol member or constituent volumes can be monitored at the **admin privilege** level with the following command:

```
cluster::> vol show -vserver SVM -volume [flexgroupname__]* -is-constituent true -fields
available,percent-used
```

Note: You can also monitor free space by using GUI utilities such as NetApp Active IQ Unified Manager, and/or by configuring ONTAP to generate alerts.

Good Workloads

Even if a workload does not conform to the preceding parameters, odds are good that a FlexGroup volume can accommodate it with ease. Remember that "Optimal Workloads" describes situations that can help a FlexGroup volume perform optimally, but even a suboptimal one provides good throughput, scaling, and load distribution for most use cases.

Nonideal Workloads

A few activities can make a FlexGroup volume work harder to maintain its balance of load and usage among constituents. Most of these activities relate to large files in one way or another. Although these workloads are able to use FlexGroup volumes, you should strive to understand the file size issues before implementing.

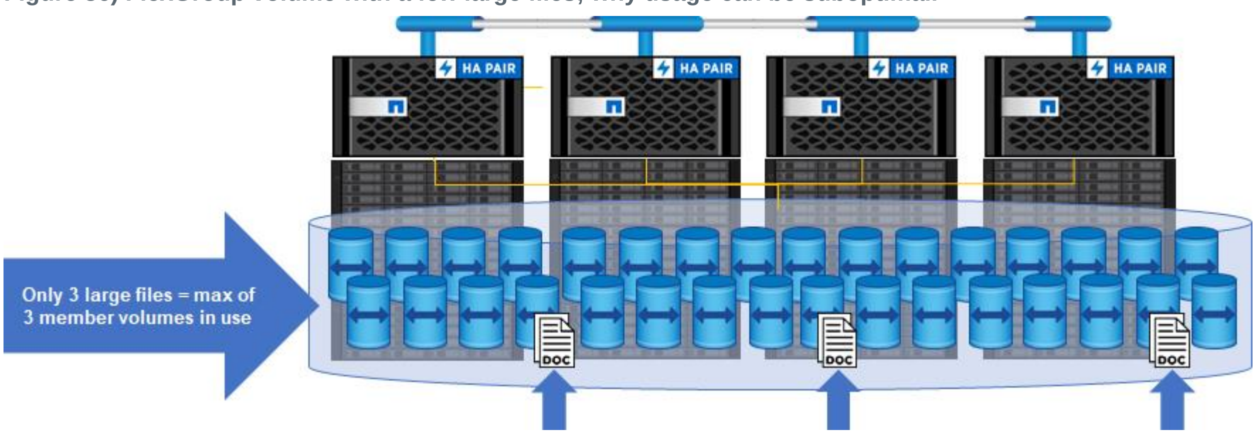
One of the key challenges of large file workloads is that storage systems are not aware of how large a file will become over time. Clients often do not have this information either; instead, a file starts out as a small inode in the storage system, and then data is written to it until the file creation/write is completed. As a result, if a client writes five files at one time and all files will eventually grow to be large, ONTAP still might place all those files locally to a parent folder, so five large files might end up in the same member volume. Generally speaking, workloads like this balance themselves out over time and the FlexGroup volume maintain even distribution. However, at times, an administrator might notice that the FlexGroup member volume distribution is uneven. The benchmark for concern should not be "My member volumes are uneven," but rather, "My FlexGroup volume is not performing as well as I expect."

Large files are marginally more difficult for the FlexGroup volume to process than small files are, primarily because using large files typically means using fewer of them overall. As previously mentioned, the FlexGroup volume performs best when new files and directories are being created frequently. If the working set consists of many large files (that are roughly the same size), the FlexGroup volume should not have trouble maintaining usage and load distribution among constituents.

Note: For more information about what the term "large file" means for a FlexGroup volume, see "What Are Large Files?" earlier in this document.

Large files also have the property of holding a great deal of information. Reading or writing that much information can take a long time. If the workload concentrates on only a few of those large files (say, reading or writing a large single-file database), then all that traffic is handled exclusively by the constituents that host those files. Because other constituents are not participating in the workload at the time, this situation can result in suboptimal usage of the FlexGroup volume. In general, expect roughly the same performance for large file/streaming workloads in a FlexGroup volume as you would see in a FlexVol volume.

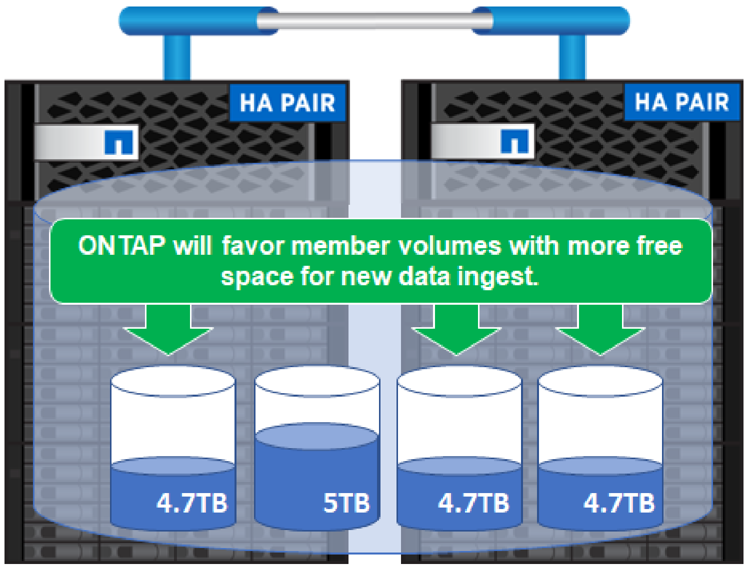
Figure 36) FlexGroup volume with a few large files; why usage can be suboptimal.



Another concern with large files is that a single file can consume enough space on the constituent to substantially affect the balance of usage among constituents. Sometimes a few files grow to a size that is orders of magnitude above the average file size. The result is that some constituents (the ones that happen to hold the aberrantly large files) end up with much more data than their peers have. In response, the FlexGroup volume begins to divert other new content creations onto the underused constituents. As a result, a subset of constituents can end up servicing most of the traffic. This problem is not typically severe; it simply represents suboptimal behavior. ONTAP 9.7 and later versions make substantial strides in handling the placement of these types of files and workloads so that they are better balanced across member volumes.

Best Practice 1: Large File Size Considerations
Before sizing a FlexGroup volume, perform an analysis to determine the largest possible file size in a workload. Then, the member volume sizes should reflect those large file sizes, so that a large file cannot consume more than 1% to 5% of a FlexGroup member volume. Following this best practice helps avoid “out of space” concerns. Also, running ONTAP 9.6 and later can help avoid “out of space” concerns by way of the elastic sizing functionality. Running the latest patched version of ONTAP is always a good practice for FlexGroup volumes.

Figure 37) FlexGroup volume with a few large files; unbalanced ingest.



One other concern relates to running with the FlexGroup volume continually very close to full. As the FlexGroup volume becomes full, ONTAP becomes proactive in placing new content on those constituents that still have free space. If the working set consists primarily of small files, this behavior is adequate to prevent clients from receiving `Volume Full` errors until the collective FlexGroup volume is indeed full. However, before ONTAP 9.6, when large files were in the workload, those files continued to grow until they completely filled their constituent/member volume, resulting in `Volume Full` errors (`ENOSPC`) even if other constituents/members still had free space. Beginning with ONTAP 9.6, elastic sizing provides a way for ONTAP to borrow space from less-full member volumes and allow file writes to complete in member volumes.

Best Practice 2: Volume Autogrow
If large files are in the workload, NetApp recommends keeping more free space on the FlexGroup member volumes and using volume autogrow functionality in ONTAP 9.3 and later or elastic sizing in ONTAP 9.6 and later.

Best Practices When Using Large Files with FlexGroup Volumes

FlexGroup volumes operate best when dealing with lots of smaller files. However, they can also be effective when storing larger files, so long as the FlexGroup volume is configured for that workload up front. When you’re sizing a FlexGroup volume for large files, it’s important to consider what a large file is, and what the largest and average file sizes in a workload are.

File sizes need to be factored in when you design a FlexGroup volume, so that member volumes are sized appropriately. ONTAP 9.6 and later versions make this process unnecessary with the addition of elastic sizing. But, in general, you can apply the following best practices for large file workloads:

- For large file sizes, consider deploying larger member volumes, with fewer members per FlexGroup volume. See Initial Volume Size Considerations for details.
- Enable volume autogrow on the FlexGroup volume to avoid running out of space in a member volume that contains large files. Or use ONTAP 9.6 and later to take advantage of elastic sizing. (Using volume autogrow disables elastic sizing for a FlexGroup volume.)
- Enable quota enforcement to limit the capacity in qtrees or by user (ONTAP 9.5 and later).
- Before deploying, use the NetApp XCP Migration Tool to scan the file system and analyze the file sizes to understand average file size, largest file size, and so on.
- Aim for sizing a FlexGroup volume so that the largest file size does not exceed 1% to 5% of the member volume's capacity. For example, if a FlexGroup with eight member volumes is 8TB in size, then the member volumes would be 1TB each. Therefore, file sizes should ideally not exceed 10GB to 50GB, or the member volume sizes should be made larger to accommodate files larger than 50GB.

Performance Expectations: Read-Heavy Workloads

Performance in a FlexGroup volume can greatly exceed that of a FlexVol volume or competitor systems for specific workloads—mainly write-metadata-heavy workloads (high `CREATE` and `SETATTR` calls) that ingest many small files. However, other workloads, such as file streams or read-heavy workloads, don't see the same extreme performance gains over FlexVol volumes that the ingest-heavy workloads see. Sometimes (especially with all local traffic), a set of FlexVol volumes might perform slightly better than a FlexGroup volume for random and sequential read workloads. However, the complexity involved with creating and managing multiple FlexVol volumes versus a single FlexGroup volume might outweigh the slight performance gains.

For read-heavy workloads, using FlexGroup volumes has benefits over using single FlexVol volumes, such as the following:

- Scaling across multiple CPUs and nodes for reads to multiple files
- Single namespace for a large-capacity bucket

When deciding whether to use a FlexGroup volume, consider support for specific features. See the earlier section on what is and is not currently supported with FlexGroup volumes.

8 Initial FlexGroup Design Considerations

This section covers initial NetApp ONTAP FlexGroup volume design considerations. In presenting this information, NetApp assumes that no previous FlexGroup volumes have been created on the cluster. NetApp also assumes that you have experience with and knowledge about managing ONTAP through the CLI and the GUI and that you have administrator-level access to the storage system.

8.1 Cluster Considerations

An ONTAP cluster that uses only NAS functionality (CIFS/SMB and NFS) can expand to up to 24 nodes (12 HA pairs). Each HA pair is a homogenous system (that is, two NetApp AFF nodes, two FAS8080 nodes, and so on), but the cluster itself can contain mixed system types. For example, a 10-node cluster could have a mixture of four AFF nodes, four NetApp FAS spinning disk systems, and two hybrid nodes for storage tiering functionality.

A FlexGroup volume can potentially span an entire 24-node cluster. However, keep the following considerations in mind.

- **FlexGroup volumes should span only hardware systems that are identical.**
  Because hardware systems can vary greatly in terms of CPU, RAM, and overall performance capabilities, the use of only homogenous systems helps promote predictable performance across the FlexGroup volume.

- **FlexGroup volumes should span only disk types that are identical.**
  Like hardware systems, disk type performance can vary greatly. For best results, make sure that the aggregates that are used are either all SSD, all spinning, or all hybrid.
- **FlexGroup volumes can span portions of a cluster.**
  A FlexGroup volume can be configured to span any combination of nodes in the cluster, from a single node, an HA pair, to across all 24 nodes. The FlexGroup volume does not have to be configured to span the entire cluster. However, doing so can take advantage of all the hardware resources that are available.

8.2 Volume Name Considerations

Before ONTAP 9.6, it was not possible to rename a FlexGroup volume. ONTAP 9.6 introduced support for volume renaming. FlexGroup volume names are meant only for identification of the volumes within the cluster by storage administrators. Client-facing names for volumes are exposed by way of CIFS/SMB shares and volume junction paths (export paths) for NFS, not by how you name a volume in the cluster. For example, a volume could be named `vol1` but exported to clients as `/accounting`. Junction paths and SMB shares can be changed at any time in a FlexGroup volume, but doing so causes a disruption for clients, because they must reconnect to the new share or export path through a remount. Volume renames, however, do not cause any client-side disruption.

In ONTAP 9.6, volume renames behave identically to volume renames for FlexVol volumes.

8.3 Failure Domains

A failure domain is an entity that, if failure occurs, can negatively affect workloads. For example, in an ONTAP cluster, if an HA pair fails (a rare occurrence), the volumes on those nodes become unavailable because there is nowhere for them to fail over to. As a result, the HA pair is considered a failure domain in the cluster. However, a single node can fail in a cluster without disruption, because its partner can take it over. In this situation, a single node would not be considered a failure domain.

FlexGroup volumes can span multiple nodes and HA pairs, and thus, multiple failure domains. However, even if a FlexGroup volume spans an entire 10-node cluster, the failure domain is still the HA pair. If you lose access to members in a FlexGroup volume (such as in the rare instance of failure of the HA pair), write access is disabled until all those members are repaired/reintroduced into the FlexGroup volume. The more HA pairs a FlexGroup volume spans, the higher the probability for failure is, because you are now spanning more failure domains.

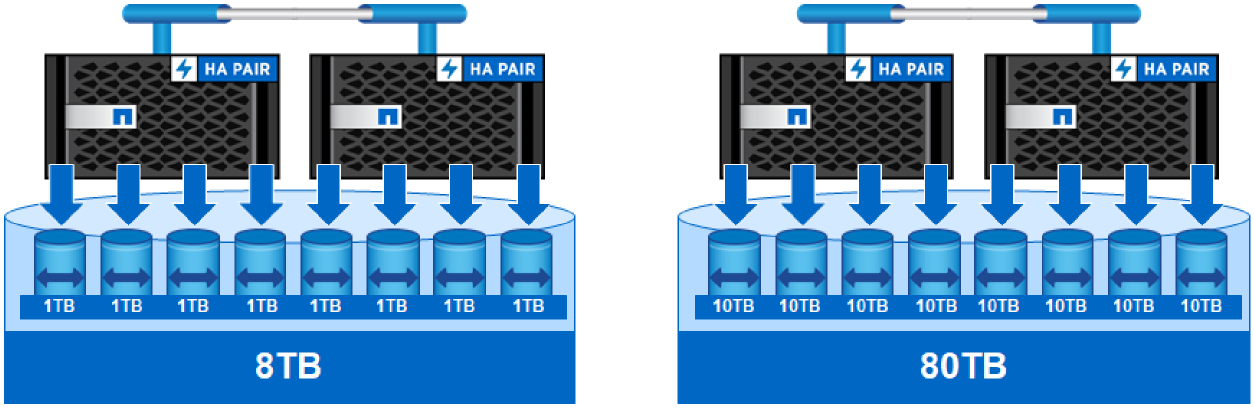
However, errors within a failure domain (such as RAID errors, losing a disk, multipath configuration errors, and metadata inconsistencies) are handled in ONTAP and do not negatively affect the FlexGroup volume.

Therefore, when planning deployment, consider how many nodes to span in a FlexGroup volume and what SLAs are acceptable, and weigh those considerations against the capacity required and performance needed.

8.4 Initial Volume Size Considerations

A common deployment issue is undersizing a FlexGroup volume's member volume capacity. FlexGroup volumes can be created at almost any capacity, but it's important to remember that several FlexVol member volumes make up the total size of the FlexGroup volume. By default, each node has a set number of member volumes (for example, 16 members in ONTAP 9.4 and later for high-end platforms with two aggregates), so the FlexGroup capacity is broken up into smaller FlexVol chunks in the form of (total FlexGroup size / number of member volumes in the FlexGroup). For example, in an 80TB FlexGroup volume with 8 member volumes, each member volume is 10TB in size. These member volume sizes are intended to be inconsequential in most workload cases, but it's important to know what the file sizes of the workload are to help plan the capacity accordingly. For example, if you know your workload has 50GB files, then you probably won't want to create a FlexGroup volume that has 100GB member volumes.

Figure 38) FlexGroup volumes—member sizes versus FlexGroup volume capacity.



Initial Space Consumption on New FlexGroup Volumes

Each FlexGroup member volume sets aside a small amount of space (around 50MB) for internal use. When a member volume is sized to the minimum of 100GB, the used space is around 0.05%, which is negligible to ONTAP. However, used space still shows up in the output of empty FlexGroup volumes, so this is something to keep in mind as a nonissue when deploying a FlexGroup volume.

For example:

```
cluster::*> vol show -vserver DEMO -volume fgautogrow* -fields used
vserver volume           used
------- ---------------- -------
DEMO    fgautogrow__0006 57.48MB
DEMO    fgautogrow__0008 57.48MB
DEMO    fgautogrow__0001 57.50MB
DEMO    fgautogrow__0004 57.50MB
DEMO    fgautogrow__0005 57.52MB
DEMO    fgautogrow__0007 57.52MB
DEMO    fgautogrow__0002 57.57MB
DEMO    fgautogrow__0003 57.57MB
DEMO    fgautogrow         460MB
```

Aggregate Free Space Considerations

When you create a FlexGroup volume, it's ideal for the aggregate (or aggregates) that the FlexGroup is deployed on to have the following characteristics:

- A roughly even amount of free space across multiple aggregates (especially important when using thin provisioning)
- Roughly 3% free space available for aggregate metadata after creation of the FlexGroup.volume (ONTAP 9.5 and earlier)

ONTAP 9.5 and earlier versions prevent a FlexGroup volume from filling an aggregate past 97% when using thick provisioning/space guarantees. Attempts fail with the error `request size is greater than maximum size`. It's possible to overcommit an aggregate by using thin provisioning, but if one aggregate has more space than the other, you run the risk of affecting performance or running out of space in members on one aggregate before the other aggregate runs out of space. ONTAP 9.6 and later versions no longer check for deduplication metadata.

Why Is This Important?

Available member volume size affects how often files are ingested locally or remotely in a FlexGroup volume, which in turn can affect performance and capacity distribution in the FlexGroup volume. File sizes are also important to consider when you're designing an initial FlexGroup volume, because large files can fill up individual members faster, causing more remote allocation, or even causing member volumes to run out of space prematurely. For example, if your FlexGroup volume has member volumes that are 100GB in size, and your files are a mix of files that are 10GB in size with many smaller files, you might run into performance issues because the larger files create an imbalance that affects the smaller files. You might also run out of space in a member volume prematurely, which results in the entire FlexGroup volume reporting that it is out of space. If possible, size your FlexGroup volumes to larger capacities (and use thin provisioning), or use fewer member volumes to allow larger capacities per member volume.

Best Practice 3: Member Volume Size Recommendations
NetApp recommends sizing a member volume such that the largest file does not exceed 1% to 5% of the member volume's capacity. Avoid creating FlexGroup volumes with fewer than two members per node.

Before ONTAP 9.6, FlexGroup volumes did not support shrinking of the volume footprint. Even with volume shrink support, do not size the volumes too large at the initial creation. If you size them too large, your administration options might be limited when you need to grow capacity later and you have to add new member volumes that are about 100TB each.

Best Practice 4: Preferred ONTAP Release
Each new ONTAP release improves how a FlexGroup volume places data as it is ingested. When you factor in the rapid pace of feature innovation for FlexGroup volumes, the best ONTAP release to use with FlexGroup volumes is the latest patched ONTAP release.

8.5 Capacity Considerations

Although FlexGroup allows massive capacity and file count possibilities, the FlexGroup volume itself is mostly limited to the physical maximums of the underlying hardware. The current maximums are only tested maximums; the theoretical maximums could go a bit higher, but the official supported member volume count in a FlexGroup volume currently stands at 200. If you need more member volumes in a FlexGroup volume, contact your NetApp sales representative or email flexgroups-info@netapp.com.

Also, there are node-specific aggregate size limitations that allow only a set number of 100TB FlexVol volumes. Be sure to review your hardware's physical capacity limitations for more information.

For example, the FAS8080 EX allows 400TB aggregates before ONTAP 9.2TB and 800TB aggregates after ONTAP 9.2, which means that we would see a maximum of four 100TB volumes allowed per aggregate or eight 100TB volumes per aggregate, depending on the ONTAP version being used. However, NetApp recommends not reaching the 100TB limit for member volumes, because doing so would make it impossible to expand member volumes further in the future. Instead, aim to leave a cushion of no less than 10% to 20% of the total maximum FlexVol member space to enable emergency space allocation.

These numbers are raw capacities, before features such as NetApp Snapshot reserve, NetApp WAFL® reserve, and storage efficiencies are factored in. For more information, see the storage limits on the NetApp Support site. To correctly size your FlexGroup solution, use the proper sizing tools, such as the System Performance Modeler (requires a NetApp login).

8.6 Elastic Sizing

Files written to a FlexGroup volume live in individual member volumes. They do not stripe across member volumes, so if a file is written and grows over time, or a large file is written to a FlexGroup volume, that write might fail because of lack of space in a member volume.

There are a few reasons why a member volume might fill up.

- You try to write a single file that exceeds the available space of a member volume. For example, a 10GB file is written to a member volume with 9GB available.

- If a file is appended over time, it eventually fills up a member volume—for example, if a database resides in a member volume.
- Snapshot copies eat into the active file system space available.

FlexGroup volumes do a good job of allocating space across member volumes, but if a workload anomaly occurs, it can have a negative effect. (For example, your volume is composed of 4,000 files but then you zip some up and create a giant single file).

One solution is to grow volumes (either manually or by using volume autogrow) or delete data. However, administrators often don't see the issue until it's too late and "out of space" errors have occurred.

For example, a FlexGroup volume can be hundreds of terabytes in size, but the underlying member volumes and their free capacities are what determine the space available for individual files. If a 200TB FlexGroup volume has 20TB remaining (10% of the volume), the amount of space available for a single file to write is not 20TB; instead, it is closer to 20TB/[number of member volumes in a FlexGroup], provided all member volumes in the FlexGroup volume have evenly distributed capacities.

In a two-node cluster, a FlexGroup volume that spans both nodes is likely to have 16 member volumes. That means if 20TB are available in a FlexGroup volume, the member volumes would have 1.25TB available. Before ONTAP 9.6, any single file that exceeded 1.25TB could not write to a FlexGroup volume without volume autogrow enabled.

Starting in ONTAP 9.6, the elastic sizing feature helps avoid "out of space" errors in this scenario. This feature is enabled by default and does not require administrator configuration or intervention.

Elastic Sizing: An Airbag for Your Data

One of our FlexGroup developers refers to elastic sizing as an "airbag": It's not designed to stop you from getting into an accident, but it does help soften the landing when it happens. In other words, it's not going to prevent you from writing large files or running out of space, but it is going to provide a way for those writes to complete.

Here's how it works:

1. When a file is written to ONTAP, the system has no idea how large that file will become. The client doesn't know. The application usually doesn't know. All that's known is "hey, I want to write a file."
2. When a FlexGroup volume receives a write request, it is placed in the best available member based on various factors, such as free capacity, inode count, time since last file creation, member volume performance (introduced in ONTAP 9.6), and so on.
3. When a file is placed, since ONTAP doesn't know how large a file will become, it also doesn't know if the file is going to grow to a size that's larger than the available space. So, the write is allowed as long as we have space to allow it.
4. If/when the member volume runs out of space, right before ONTAP sends an "out of space" error to the client, it queries the other member volumes in the FlexGroup volume to see if there's any available space to borrow. If there is, ONTAP adds 1% of the volume's total capacity (in a range of 10MB to 10GB) to the volume that is full (while taking the same amount from another member volume in the same FlexGroup volume) and then the file write continues.
5. During the time ONTAP is looking for space to borrow, that file write is paused. This appears to the client as a performance issue. But the overall goal isn't to finish the write fast—it's to allow the write to finish at all. Usually, a member volume is large enough to provide the 10GB increment (1% of 1TB is 10GB), which is often more than enough to allow a file creation to complete. In smaller member volumes, the effect on performance could be greater, because the system needs to query to borrow space more often.
6. The capacity borrowing maintains the overall size of the FlexGroup volume—for example, if your FlexGroup volume is 40TB in size, it remains 40TB.

Figure 39) File write behavior before elastic sizing.

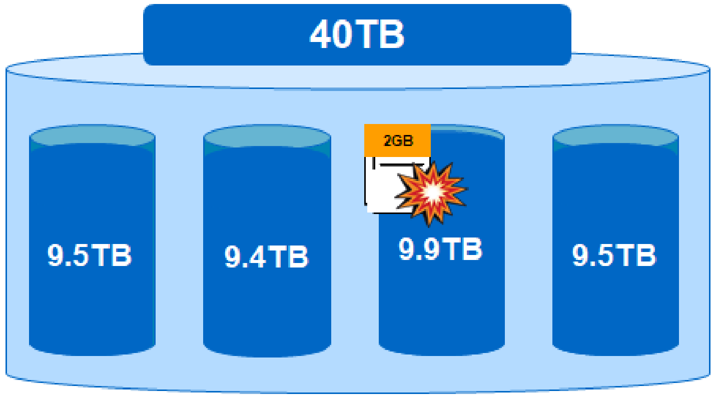
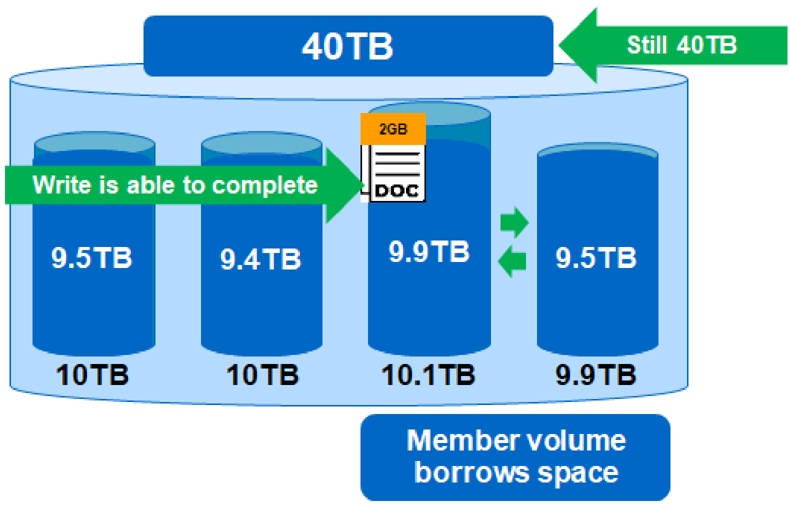


Figure 40) File write behavior after elastic sizing.



After files are deleted or volumes are grown and space is available in that member volume again, ONTAP re-adjusts the member volumes back to their original sizes to maintain an evenness in space.

Ultimately, elastic sizing helps remove the administrator overhead of managing space and worrying so much about the initial sizing and deployment of a FlexGroup volume. You can spend less time thinking about how many member volumes you need, what size they should be, and so on.

When you combine elastic sizing in ONTAP 9.6 and later with features like autogrow/shrink, ONTAP can manage your capacity and help avoid emergency space issues.

When to Use Volume Autogrow Versus Elastic Sizing

When volume autogrow is enabled on a FlexGroup volume, elastic sizing is effectively disabled for that volume. The two features are essentially redundant. But there are some differences in how they work and when you'd want to use one over the other.

- **Volume autogrow** should be used when the total capacity of the FlexGroup volume should be grown to accommodate new data being written to it.
- **Elastic sizing** should be used when individual files grow over time, when large files are present, when file sizes in the workload vary greatly, or when the total size of the FlexGroup volume should not be allowed to grow past the specified capacity.

Note: Elastic sizing is disabled only when volume autogrow is enabled. When autogrow is disabled, elastic sizing is reenabled.

Performance Impact of Elastic Sizing

Each time a file write must pause for ONTAP to find more space in the FlexGroup volume, client latency occurs. The amount of latency seen for a write operation to a file depends on the number of times the write has to pause to find more space. For example, if a member volume has just 10GB available, but a 100GB file is being written, then elastic sizing will cause the write to pause a number of times to allow the write to complete. That number is determined by the member volume total size, which can be anywhere from 10MB to 10GB.

The following example shows a test in which a file was copied to a FlexGroup volume. In the first test, the FlexGroup constituent wasn't large enough to hold the file, so elastic sizing was used. The 6.7GB file took around 2 minutes to copy:

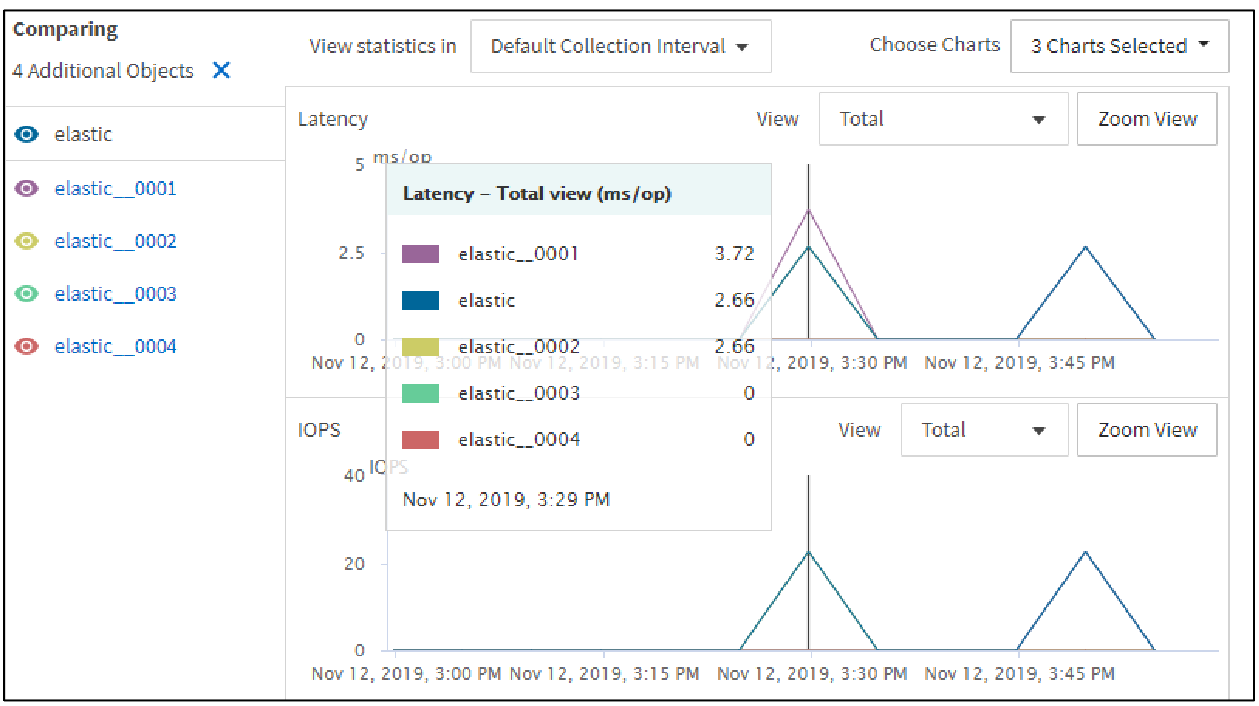
```
[root@centos7 /]# time cp Windows.iso /elastic/
real     1m52.950s
user     0m0.028s
sys      1m8.652s
```

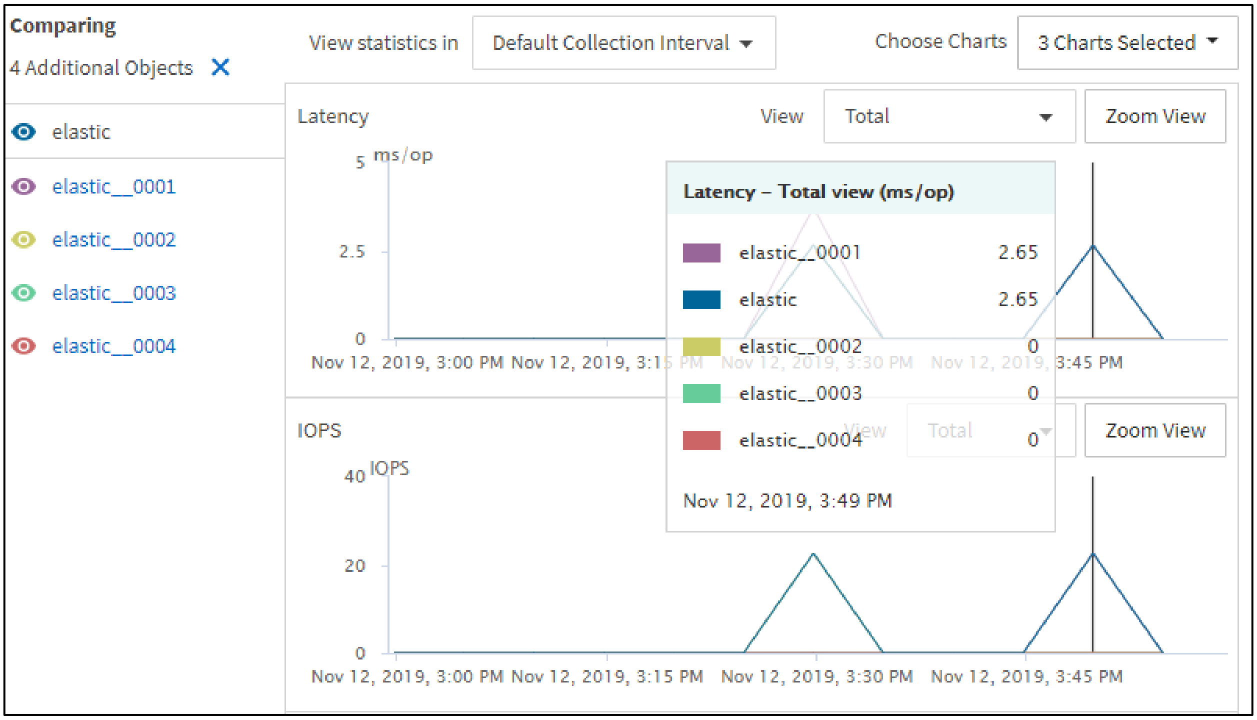
When the FlexGroup constituent volume was large enough to avoid elastic sizing, the same copy took 15 seconds less:

```
[root@centos7 /]# time cp Windows.iso /elastic/
real     1m37.233s
user     0m0.052s
sys      0m54.443s
```

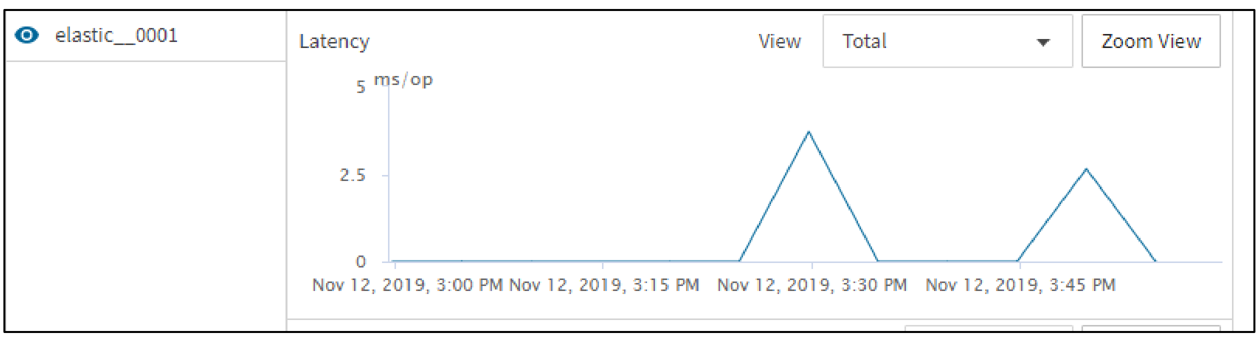
That shows there can be a real latency effect with elastic sizing.

The following graphs illustrate the latency hit on the constituent volume:





The constituent volume 0001 has about .5ms more latency when elastic sizing is in effect:



If you suspect that elastic sizing might be causing performance issues, you can do one of the following:

- Open a support case to confirm symptoms and logging.
- Grow the FlexGroup volume to make sure that there is enough space to remove elastic sizing from the equation.

Ultimately, it's crucial to keep enough free space in a FlexGroup volume to avoid needing elastic sizing. Elastic sizing is intended to be a safeguard for workloads, rather than being a way to constantly resize volumes.

8.7 Aggregate Layout Considerations

An aggregate is a collection of physical disks that are laid out into RAID groups and provide the back-end storage repositories for virtual entities such as FlexVol and FlexGroup volumes. Each aggregate is owned by a specific node and is reassigned during storage failover events.

Starting in ONTAP 9, aggregates have dedicated NVRAM partitions for consistency points to avoid scenarios in which slower or degraded aggregates cause issues on the entire node. These consistency points are also known as per-aggregate consistency points and allow mixing of disk shelf types on the same nodes for more flexibility in the design of the storage system.

Best Practice 5: Aggregate Usage with FlexGroup
For consistent performance across the entire FlexGroup volume when using FlexGroup, make sure that the FlexGroup volume spans only aggregates with the same disk type and RAID group configurations.

Table 8 shows the best practices that NetApp recommends for aggregate layout when you use FlexGroup volumes. Keep in mind that these practices are not hard requirements. The one-aggregate-per-node recommendation for AFF systems originates from disk cost with NetApp RAID-TEC and no Advanced Disk Partitioning (ADP), because you don't want to use up expensive SSD space just for parity. However, with ADP, partitions are spread across data disks, so two aggregates per node on AFF systems are better, because there are more available volume affinities per node with more aggregates present.

Table 8) Best practices for aggregate layout with FlexGroup volumes.

Spinning Disk or Hybrid Aggregates	AFF
Two aggregates per node	One aggregate per node (without ADP) Two aggregates per node (with ADP)

Note: For consistent performance, aggregates should have the same number of drives and RAID groups across the FlexGroup volume.

For more information about aggregate layouts when dealing with existing FlexVol volumes, see "Failure Domains" in this document.

Drive Failures

When a drive failure occurs, a new drive is automatically transitioned into the aggregate, through disk copy or building from parity. This can affect performance while the drive is being added. In a FlexGroup volume, where data can land on any aggregate the FlexGroup resides on at any time, inconsistent performance is possible. This is also true of drives that are affecting performance as they start to fail. Be proactive with hardware management and drive replacement whenever possible.

8.8 Security and Access Control List Style Considerations

With the FlexGroup feature in ONTAP, you can access the same data through NFS and SMB/CIFS by using multiprotocol NAS access. The same general guidance that applies to a FlexVol volume applies to a FlexGroup volume. That guidance is covered in the product documentation in the CIFS, NFS, and Multiprotocol Express Guides and the CIFS and NFS Reference Guides, which can be found with the product documentation for the specific ONTAP version being used.

In general, for multiprotocol access, you need the following:

- Valid users (Windows and UNIX)
- Valid name-mapping rules or 1:1 name mappings through local files and/or servers such as LDAP or NIS
- Volume security style (NTFS, UNIX, or mixed)
- A default UNIX user (pcuser, created by default)

When a volume is created, a security style is chosen. If you create the volume without specifying a security style, the volume inherits the security style of the SVM root volume. The security style determines the style of access control list (ACL) that is used for a NAS volume and affects how users are authenticated and mapped into the SVM. When a FlexGroup volume has a security style selected, all member volumes have the same security style settings.

Basic Volume Security Style Guidance

The following is some general guidance on selecting a security style for volumes:

- In UNIX security style, Windows users must map to valid UNIX users.
- In NTFS security style, Windows users must map to valid UNIX users, and UNIX users must map to valid Windows users to authenticate. Authorization (permissions) are handled by the Windows client after the initial authentication.
- Neither UNIX nor NTFS security style allows users from the opposite protocol to change permissions.
- A mixed security style allows permissions to be changed from any type of client. However, it has an underlying “effective” security style of NTFS or UNIX, based on the last client type to change ACLs.
- A mixed security style does not retain ACLs if the security style is changed. If the environment is not maintained properly and user mappings are not correct, this limitation can result in access issues.
- If granularity of ACLs in a FlexGroup volume is desired, consider deploying qtrees in the FlexGroup volume, which are available starting in ONTAP 9.3. Qtrees allow you to set security styles per logical directory in ONTAP. If you want other home directory features such as NetApp FPolicy, antivirus, native file auditing, and quota enforcement, then use the most recent patched release of ONTAP 9.5 or later.

Best Practice 6: Volume Security Style: Mixed Security Style Guidance
NetApp recommends a mixed security style only if clients need to be able to change permissions from both styles of clients. Otherwise, it’s best to select either NTFS or UNIX as the security style, even in multiprotocol NAS environments.

More information about user mapping, name service best practices, and so on, can be found in the product documentation. You can also find more information in TR-4073: Secure Unified Authentication, TR-4067: NFS Best Practice and Implementation Guide, and TR-4379: Name Services Best Practices Guide.

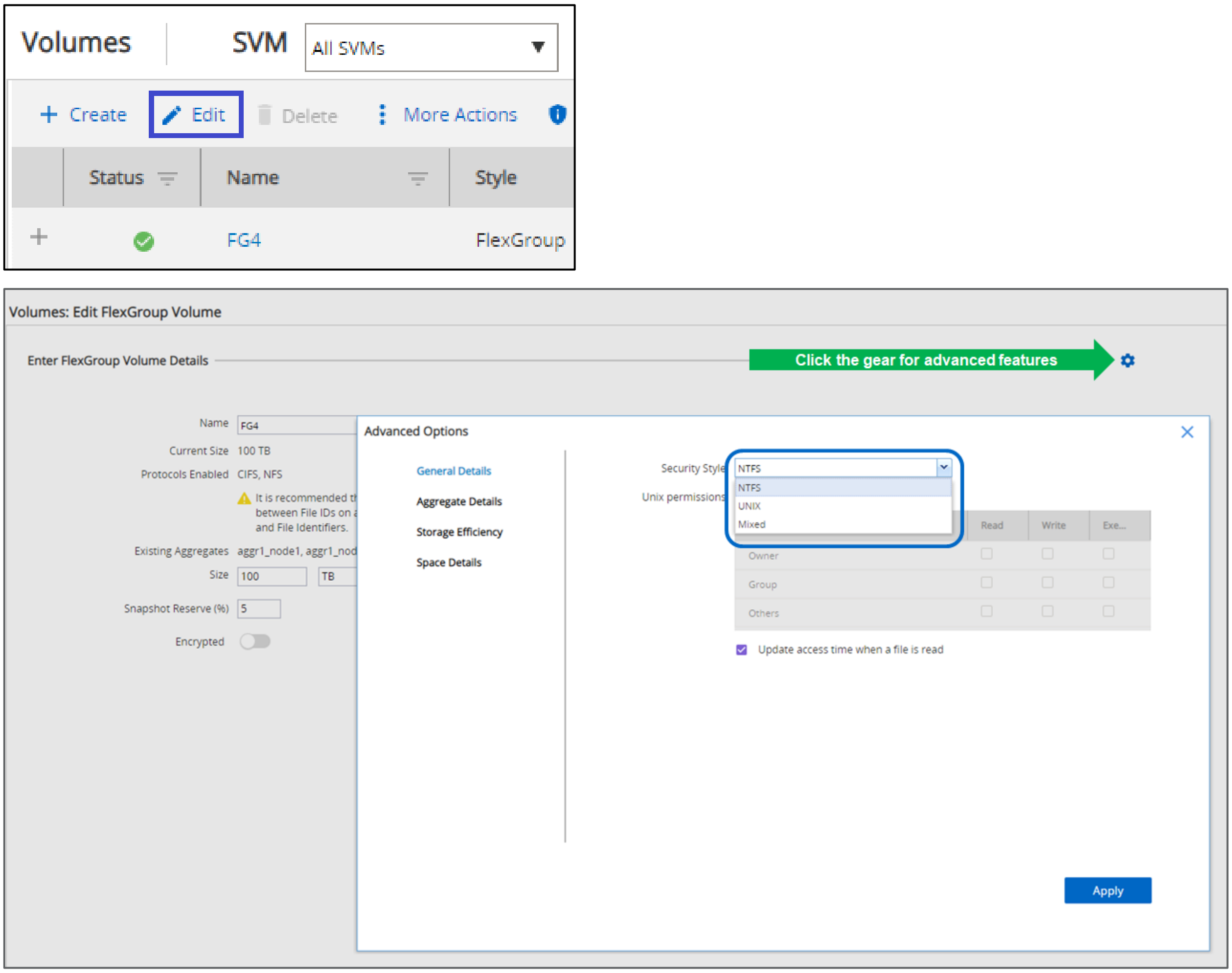
Changing the Security Style of a FlexGroup Volume

FlexGroup volumes are intended to be managed similarly to FlexVol volumes. Changing the security style of volumes is included in that methodology. Volume security styles can be changed live, with no need for clients to remount. However, the subsequent change in ACL styles means that access permissions might become unpredictable. For the best possible results, NetApp recommends changing security styles in a maintenance window on production datasets.

To change the security style of a FlexGroup volume, do one of the following:

- Use `volume modify` from the command line.
- Use the Edit button or Advanced Features when initially creating the FlexGroup volume in NetApp ONTAP System Manager.

Figure 41) Modifying FlexGroup volume security styles in ONTAP System Manager.



Workarounds for Lack of NFSv4.x ACL Support: ONTAP 9.6 and earlier

Sometimes, storage administrators might want the extra, more granular security offered by NFSv4.x ACLs in addition to the performance characteristics and large-capacity footprint of FlexGroup volumes. Before ONTAP 9.7, FlexGroup volumes did not support NFSv4.x. If you require NFSv4.x ACLs, upgrade to ONTAP 9.7.

If you want to keep using ONTAP 9.6 or earlier, here are some possible workarounds, depending on the use case:

- **If the need for NFSv4.x ACLs is due to the NFSv3 group limitation (16 GIDs per user)**, then the NFS server options `-auth-sys-extended-groups enable` and `-extended-groups-limit [1-1024]` can be used to increase the number of supported GIDs per user. See TR-4067 for information about using this feature. NetApp highly recommends the use of a name service, such as LDAP, with this feature. See TR-4073 and TR-4668 for information about setting up name services.
- **If the need for NFSv4.x ACLs is due to the need for more granular permissions**, then consider adding a CIFS server to the SVM and using NTFS-style ACLs on the FlexGroup volume. NTFS-style ACLs provide the same level and granularity as NFSv4.x ACLs, with the added benefit of having an easy-to-manage GUI option. When you're using CIFS/SMB with NFS (also known as multiprotocol NAS), using a name service like LDAP to handle user authentication and mapping makes

management easier. See TR-4073 and TR-4668 for information about setting up name services to apply multiprotocol NAS access.

- **If it's not possible to use a name service server like LDAP and/or set up CIFS in the environment**, then NFSv4.x ACL-like functionality is not currently available with FlexGroup volumes. If you need NFSv4.x ACLs, send email to flexgroups-info@netapp.com with the subject "NFSv4.x ACL support" and describe your use case.

Using NFSv4.x ACLs with NFSv3 clients

In ONTAP, it is possible to leverage the benefits of NFSv4.x's granular ACL support even if your clients are only using NFSv3. You can set NFSv4.x ACLs on files and folders through a client that has mounted an export using NFSv4.x. Clients accessing from NFSv3 will honor those ACLs without needing to use NFSv4.x mounts.

8.9 Networking Considerations

When you use CIFS/SMB or NFS, each mount point is made over a single TCP connection to a single IP address. In ONTAP, these IP addresses are attached to data LIFs, which are virtual network interfaces in an SVM.

The IP addresses can live on a single hardware Ethernet port or multiple hardware Ethernet ports that participate in a Link Aggregation Control Protocol (LACP) or another trunked configuration. However, in ONTAP, these ports always reside on a single node, which means that they are sharing that node's CPU, PCI bus, and so on. To help alleviate this situation, ONTAP allows TCP connections to be made to any node in the cluster, after which ONTAP redirects that request to the appropriate node through the cluster back-end network. This approach helps distribute network connections and load appropriately across hardware systems.

Best Practice 7: Network Design with FlexGroup
FlexGroup networking best practices are similar to FlexVol networking best practices. When you design a NAS solution in ONTAP, consider the following networking best practices regardless of the volume style: • Create at least one data LIF per node, per SVM to confirm a path to each node. • Present multiple IP addresses to clients behind a single fully qualified domain name (FQDN) by using some form of DNS load balancing. For DNS load balancing details, see TR-4523. • When possible, use LACP ports to host data LIFs for throughput and failover considerations. • When you mount clients, spread the TCP connections across cluster nodes evenly. • For clients that do frequent mounts and unmounts, consider using on-box DNS to help balance the load. (If clients are not mounted and unmounted frequently, on-box DNS doesn't help much.) • If the workload is that of a "mount storm" (that is, hundreds or thousands of clients mounting at the same time), use off-box DNS load balancing and/or consider using NetApp FlexCache volumes. • Follow the latest general networking best practices that are listed in TR-4191.

LACP Considerations

There are valid reasons for choosing to use an LACP port on client-facing networks. A common and appropriate use case is to offer resilient connections for clients that connect to the file server over the SMB 1.0 protocol. Because the SMB 1.0 protocol is stateful and maintains session information at higher levels of the OSI stack, LACP offers protection when file servers are in an HA configuration. Later implementation of the SMB protocol can deliver resilient network connections without the need to set up LACP ports. For more information, see TR-4100: Nondisruptive Operations with SMB File Shares.

LACP can provide benefits to throughput and resiliency, but you should consider the complexity of maintaining LACP environments when you are deciding. Even if LACP is involved, multiple data LIFs should still be used.

DNS Load-Balancing Considerations

DNS load balancing (both off-box and on-box) provides a method to spread network connections across nodes and ports in a cluster. FlexGroup volumes do not change the overall thinking behind DNS load balancing; storage administrators should still spread network connections across a cluster evenly, regardless of what the NAS container is. However, because of the design of FlexGroup volumes, remote cluster traffic is a certainty when you provision a FlexGroup volume across nodes, so network connection/data locality considerations are nullified in those configurations. Therefore, DNS load balancing fits in a bit better with a FlexGroup volume, because data locality is no longer a factor. Ultimately, the decision to use a DNS load-balancing method comes down to the storage and network administrators' goals. For more information about DNS load balancing, see TR-4523: DNS Load Balancing in ONTAP.

Best Practice 8: DNS Load Balancing
When possible, use some form of DNS load balancing with FlexGroup volumes on nodes that contain FlexGroup member volumes.

On-Box DNS or Off-Box DNS?

ONTAP provides a method to service DNS queries through an on-box DNS server. This method factors in a node's CPU and throughput to help determine which available data LIF is the best one to service NAS access requests.

- Off-box DNS is configured by way of the DNS administrator creating multiple "A" name records with the same name on an external DNS server that provides round-robin access to data LIFs.
- For workloads that create mount-storm scenarios, the ONTAP on-box DNS server cannot keep up and balance properly, so it's preferable to use off-box DNS.

NetApp recommends as a best practice creating at least one data LIF per node per SVM. However, it might be prudent to create multiple data LIFs per node per SVM and to mask the IP addresses behind a DNS alias through DNS load balancing. Then you should create multiple mount points to multiple IP addresses on each client to allow more potential throughput for the cluster and the FlexGroup volume.

Network Connection Concurrency: NFSv3

In addition to the preceding considerations, it's worth noting that ONTAP has a limit of 128 concurrent operations per TCP connection for NFSv3 operations. This limit means that for every IP address, the system can handle only up to 128 concurrent operations. Therefore, it's possible that an NFSv3 client would not be able to push the storage system hard enough to reach the full potential of the FlexGroup technology. Clients can be configured to control the number of concurrent operations (by using RPC slot tables) that are sent through NFSv3, which can help avoid hard-to-track performance issues.

Identifying Potential Issues with RPC Slot Tables

Many modern NFSv3 clients use dynamic values for RPC slot tables, which means that the client will send as many concurrent operations on a single TCP thread as possible—up to 65,336. However, ONTAP allows only 128 concurrent operations per TCP connection, so if a client sends more than 128, ONTAP will enact a form of flow control on NFSv3 operations to prevent rogue clients from overrunning storage systems by blocking the NFS operation (exec contexts in ONTAP) until resources free up. This flow control may manifest as performance issues that cause extra latency and slower job completion times that may not have a readily apparent reason from the general storage system statistics. These issues can appear to be network related, which can send storage administrators down the wrong troubleshooting path.

To investigate whether RPC slot tables might be involved, use the ONTAP performance counter. You can check whether the number of exec contexts blocked by the connection being overrun is incrementing.

To gather those statistics, run the following command:

```
statistics start -object cid -instance cid
```

Then, review the statistics over a period of time to see if they are incrementing.

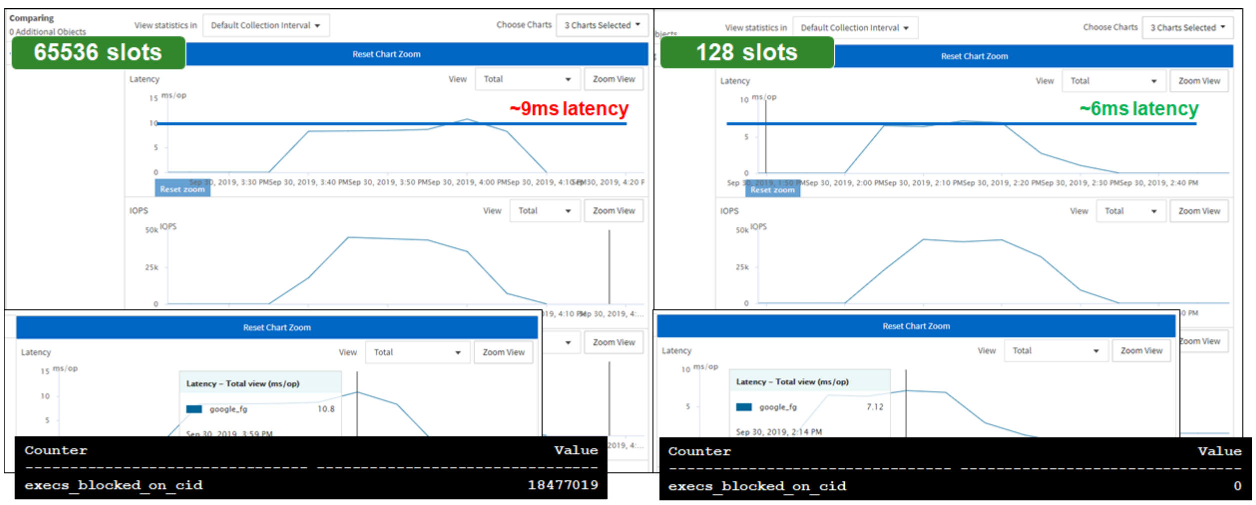
```
statistics show -object cid -instance cid -counter execs_blocked_on_cid
```

Example of RPC Slot Table Impact on Performance

In the following example, a script was run to create 18 million files across 180,000 subdirectories. This load generation was done from three clients to the same NFS mount. The goal was to generate enough NFS operations with clients that had the default RPC slot table settings to cause ONTAP to enter a flow-control scenario. Then the same scripts were run again on the same clients—but with the RPC slot tables set to 128.

The result was that the default slot tables (65,536) generated 18 million `execs_blocked_on_cid` events and added 3ms of latency to the workload versus the run with the lower RPC slot table setting (128).

Figure 42) Impact of RPC slot tables on NFSv3 performance.



Although 3ms might not seem like a lot of latency, it can add up over millions of operations, considerably slowing down job completion.

Resolving Issues with RPC Slot Tables

ONTAP cannot control the number of slot tables a client sends per TCP connection for NFSv3 operations. Therefore, clients must be configured to limit the maximum slot tables sent through NFS to 128. This setting will vary depending on the client OS version. Contact the client vendor for more information.

It is possible to get more performance out of a client’s NFS connectivity by connecting more mount points to different IP addresses in the cluster on the same client, but that approach can create complexity. For example, rather than mounting a volume at `SVM:/volumename`, multiple mount points on the same client across different folders and IP addresses in the volume could be created.

For example:

```
LIF1:/volumename/folder1
LIF2:/volumename/folder2
LIF3:/volumename/folder3
```


Another possible option is to use the nconnect option available for some Linux distributions that can perform multiplexing of NFSv3 over the same TCP connection. This option provides more available concurrent sessions and better overall performance.

Does the RPC Slot Table Limit Affect Other NAS Protocols?

RPC slot table limits affect only NFSv3 traffic.

- SMB clients use different connection methodologies for concurrency, such as SMB multichannel, SMB multiplex, and SMB credits. The SMB connection methodology depends on client/server configuration and protocol version. For example, SMB 1.0 uses SMB multiplex (mpx), whereas SMB2.x uses SMB credits.
- NFSv4.x clients do not use RPC slot tables—instead, they use state IDs and session tables to control the flow of concurrent traffic from clients.

Border Gateway Protocol (BGP): ONTAP 9.5 and Later

Starting in version 9.5, ONTAP supports BGP to provide a more modern networking stack for your storage system. BGP support provides layer 3 (L3) routing, improved load-balancing intelligence, and virtual IPs (VIPs) for more efficient port utilization.

FlexGroup volumes need no configuration changes to use this new networking element.

8.10 FlexVol Member Volume Layout Considerations

A FlexVol volume is provisioned from the available storage in an aggregate. FlexVol volumes are flexible and can be increased or decreased dynamically without affecting or disrupting the environment. A single aggregate can contain many FlexVol volumes. A FlexVol volume is not tied to any specific set of disks in the aggregate and exists across all the disks in the aggregate. However, files themselves are not striped; they are allocated to individual FlexVol member volumes.

FlexVol volumes are the building blocks to a FlexGroup volume. Each FlexGroup volume contains several member FlexVol volumes to provide concurrent performance and to expand the capacity of the volume past the usual 100TB limits of single FlexVol volumes.

When designing a FlexGroup volume, consider the following for the underlying FlexVol member volumes:

- When you use automated FlexGroup creation methods such as `volume create -auto-provision-as flexgroup` (introduced in ONTAP 9.2), `flexgroup deploy`, or ONTAP System Manager, the default number of member FlexVol volumes in a FlexGroup volume depends on the platform and ONTAP release. It can range from 8 per node (4 per aggregate) or 16 per node (8 per aggregate).

  Here is an example of `volume create` with the `-auto-provision-as flexgroup` option:

```
volume create -vserver vs0 -volume fg -auto-provision-as flexgroup -size 200TB
```

  Note: For almost all use cases, NetApp recommends that you let ONTAP determine the member volume count per node. ONTAP provides a member count that provides optimal performance through volume affinities and CPU slots. If needed, the number of member volume counts can be fewer or greater than eight, depending on the desired configuration. For more information, see the section Volume Affinity and CPU Saturation.

- Currently, when two aggregates of spinning disks are on a node, the automated FlexGroup creation methods create four member volumes per aggregate.
- When a single SSD aggregate is on a node, the automated FlexGroup creation methods create eight member volumes per aggregate.
- If a node with spinning disk does not contain two aggregates, the automated FlexGroup creation method might fail in some earlier ONTAP versions. If this happens, continue with manual creation.

- Automated FlexGroup creation methods currently do not consider CPU, RAM, or other factors when deploying a FlexGroup volume. Instead, they follow a hard-coded methodology.
- FlexVol member volumes are deployed in even capacities, regardless of how the FlexGroup volume was created. For example, if an eight-member, 800TB FlexGroup volume was created, each member is deployed with 100TB. If a larger or smaller quantity of FlexVol member volumes is required at the time of deployment, use the `volume create` command with the `-aggr-list` and `-aggr-list-multiplier` options to customize the number of member volumes deployed per aggregate. For an example, see "Directory Size Considerations."
- When growing or shrinking (shrink is supported starting in ONTAP 9.6) a FlexGroup volume, use the `volume size` command at the FlexGroup level; avoid resizing individual FlexVol member volumes.
- Consider disabling the space guarantee (thin provisioning) on the FlexGroup volume to allow the member volumes to be overprovisioned.

When Would I Need to Manually Create a FlexGroup Volume?

In most cases, letting ONTAP choose the member volumes by using the automated commands or GUI operations is a good option when creating a FlexGroup volume. In other words, don't worry about CPU count/volume affinity best practices for FlexGroup creation—let ONTAP do that for you.

However, in some use cases, manual creation might be needed.

Concern Regarding Overprovisioning Volume Counts

In ONTAP, FlexVol volumes are currently limited to a maximum of 1,000 per node. Because a FlexGroup volume is composed of FlexVol volumes, those limits also apply to FlexGroup volumes. If you have FlexGroup volumes with many member volumes or you want to create many FlexGroup volumes in a cluster, then you would need to consider the overall volume limits per node. You might also need to manually create the FlexGroup volumes to modify the default volume counts or aggregate placement to keep total FlexVol volume numbers below the node limits.

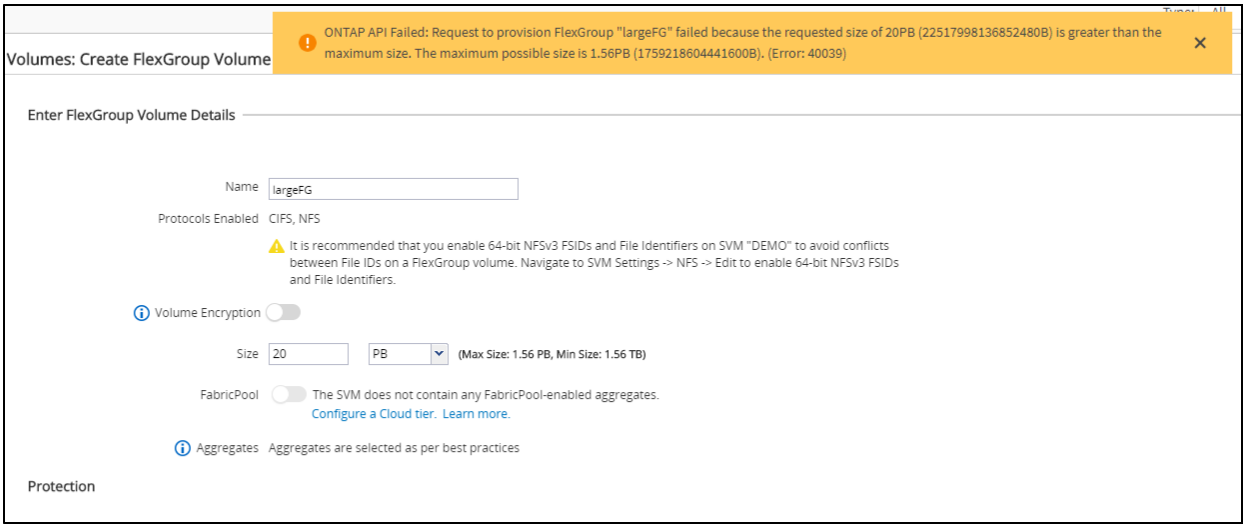
Large Files, Limited Capacity

If you have a workload with larger files and you want to comply with best practices for large files, you might need to adjust the member volume count to create fewer members at larger individual capacities—particularly if your cluster is already limited in the amount of capacity that's available. For example, if you want to create a 16TB FlexGroup volume across 4 nodes, that would create a minimum of 32 member volumes that are 500GB in size each. If your average file size is 250GB, then 500GB member volumes won't be large enough to distribute the data effectively. Manually creating a FlexGroup volume with fewer, larger member volumes will work better for those use cases.

A Need for a Large Amount of Capacity or High File Counts

FlexVol volumes are limited to 100TB in size and can contain up to 2 billion files. If you have a 2-node cluster and let ONTAP create a FlexGroup volume, you will get at most 16 member volumes in a single FlexGroup volume in some cases, because it is code-limited to best practices of 8 per node. In the following example, the 2-node cluster can only create a FlexGroup volume with a maximum of 1.56PB of capacity (8 members per node; 16 members total; 100TB per member volume).

Figure 43) Error when creating a FlexGroup volume beyond the allowed maximum in System Manager.



If you desire a larger FlexGroup volume, you would need to create the FlexGroup manually to allow a higher number of member volumes by using the `-aggr-list-multiplier` option. For a 20PB FlexGroup volume, you would need 200 member volumes.

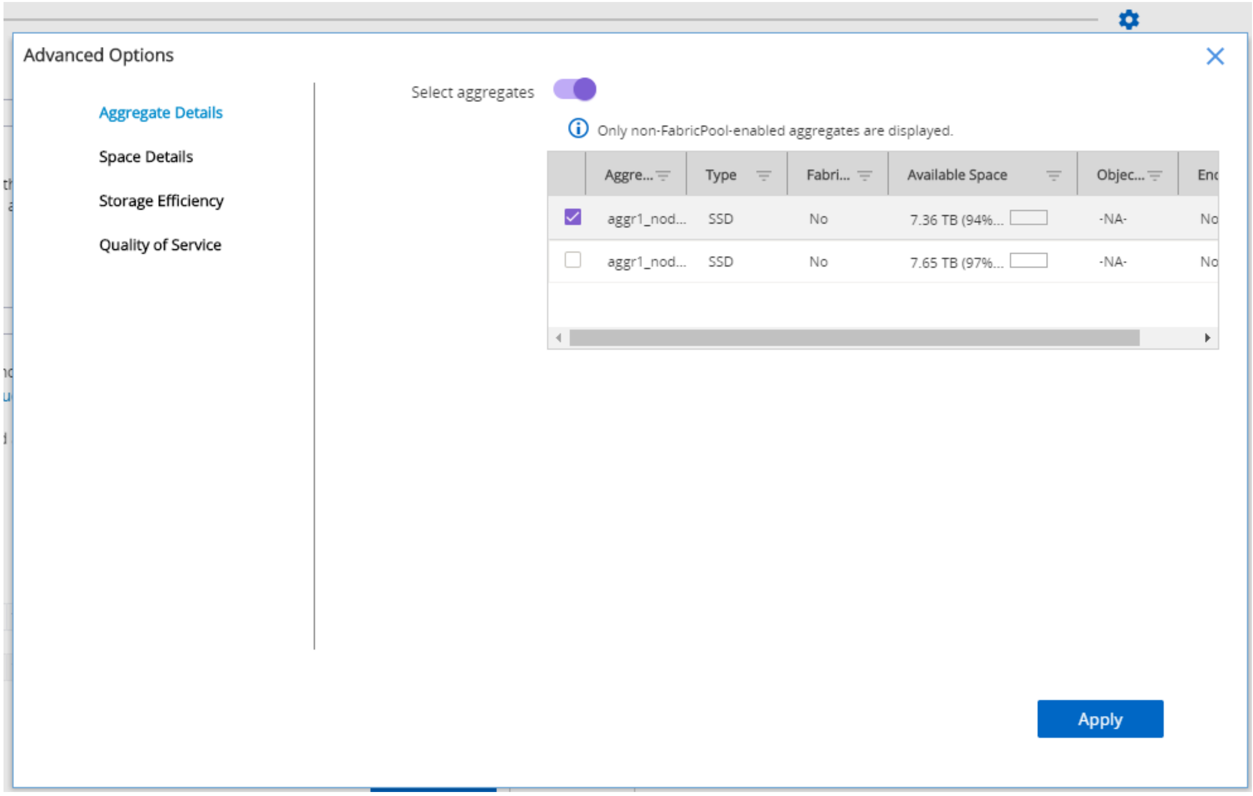
Similar considerations should be made if the file count needs to exceed the maximum files allowed. In the 16-member FlexGroup example, a maximum of 32 billion files would be allowed. If more files are needed, add more member volumes.

For an example of how to create a FlexGroup volume from the CLI and specify the number of members, see “Command Examples” later in this document.

Avoiding the Cluster Network

A less common scenario is the desire to avoid the cluster network by creating a FlexGroup volume across a single node. In this use case, you can use the classic view of System Manager to create the FlexGroup volume by selecting the Advanced Options gear in the volume creation window and toggle the Select Aggregates button in the Aggregate Details menu. Then, choose which aggregate you want to provision the FlexGroup volume across. Currently, the ONTAP 9.7 System Manager doesn’t let you manually select aggregates, so use the CLI in ONTAP 9.7.

Figure 44) Selecting a single aggregate in the classic System Manager (before ONTAP 9.7).



Maximum Number of FlexGroup Volumes in a Cluster

A FlexGroup volume can consist of a single FlexVol member volume or hundreds of FlexVol member volumes. The maximum number of FlexVol member volumes is physically constrained only by the total volume count in a cluster. As a result, a FlexGroup volume could theoretically have up to 24,000 member volumes in a 24-node cluster.

The total number of FlexGroup volumes is similarly constrained by the total volume count in a cluster. Each FlexGroup volume's member volumes are part of the volume count, so the number of FlexGroup volumes allowed in a cluster depends on the number of member volumes.

For example, a two-node cluster would have 2,000 volumes to work with. As a result, you could have one of the following configurations (although others are possible):

- 10 FlexGroup volumes with 200 member volumes
- 20 FlexGroup volumes with 100 member volumes
- 40 FlexGroup volumes with 50 member volumes
- 200 FlexGroup volumes with 10 member volumes

Keep in mind that the existence of other FlexVol volumes in the cluster (including SVM root volumes) affects the total number of available member volumes. FlexVol member volume limits also can be constrained by the FlexGroup volumes participating in NetApp SnapMirror relationships. For the latest details on those limitations, see TR-4678.

Also, as a best practice, limit a FlexGroup volume to 8 to 16 member volumes per node (depending on the system platform and volume affinities available). ONTAP deploys the FlexGroup volume with the recommended member volume count if you use the automated deployment methods in ONTAP System Manager or the `flexgroup deploy/-auto-provision-as` options in the CLI. For more information, see "FlexVol Member Volume Layout Considerations" earlier in this document.

Do I Need a Large Number of Member Volumes?

Usually, you do not need to exceed the best practice volume count for a FlexGroup volume. However, if you need more capacity or higher file counts, you can increase the number of member volumes at initial deployment, or you can do so later by using the `volume expand` command. For more information about when you might need to stray from ONTAP best practices for member volume counts, see the section above.

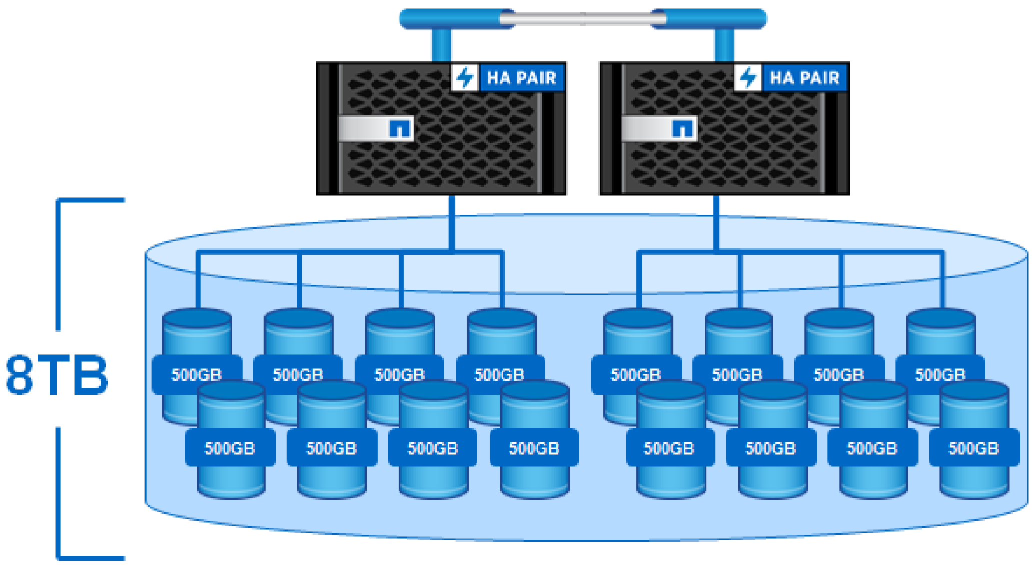
Member Count Considerations for Large and Small Files

FlexGroup volumes work best in a high-file-count environment of many small files. However, they also work well with larger files. As mentioned in "What Are Large Files?" earlier in this document, large files should be considered in terms of percentage of the total space allocated to a member volume.

When larger files are present in a workload, the initial deployment size of a FlexGroup should be kept in mind. By default, a FlexGroup deploys eight-member volumes per node, so any capacity footprint that is defined at the FlexGroup level effectively gets divided into [total space/n number of member volumes].

For example, if an 8TB FlexGroup is deployed across two nodes in a cluster, and the member count is 16, then each member volume is about 500GB in size.

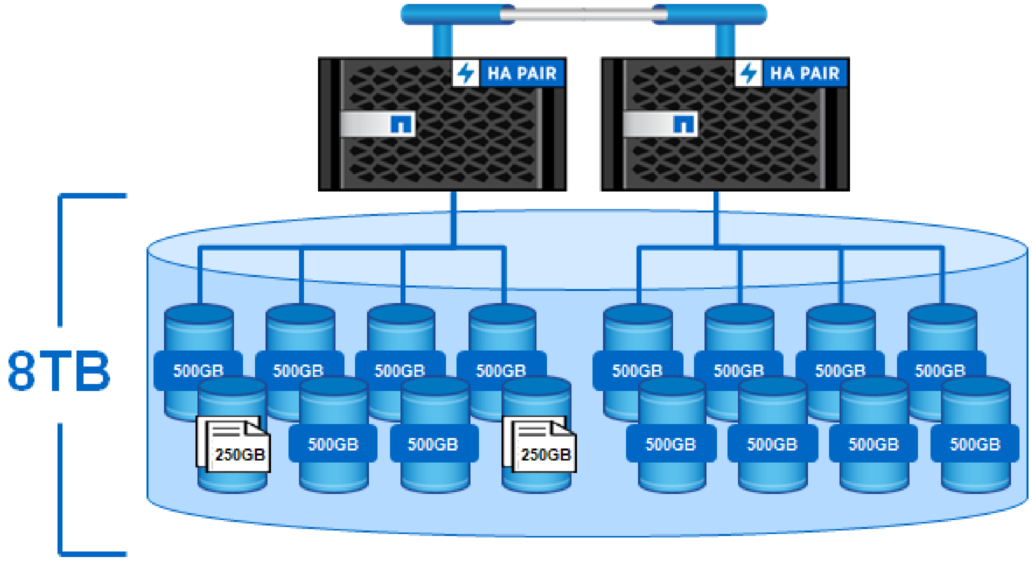
Figure 45) How capacity is divided among member volumes in a FlexGroup volume.



In many workloads, the distribution shown in Figure 45 would work well. However, if larger files in a workload would potentially fill in member volumes' large chunks of capacity used, then performance or even accessibility could be affected. Recall that if a member volume fills before other member volumes, the FlexGroup volume could report an "out of space" error in ONTAP versions earlier than 9.6.

For example, if some files in a workload are 250GB, then each time a file is written to the FlexGroup volume, 50% of the total capacity of a member volume is filled.

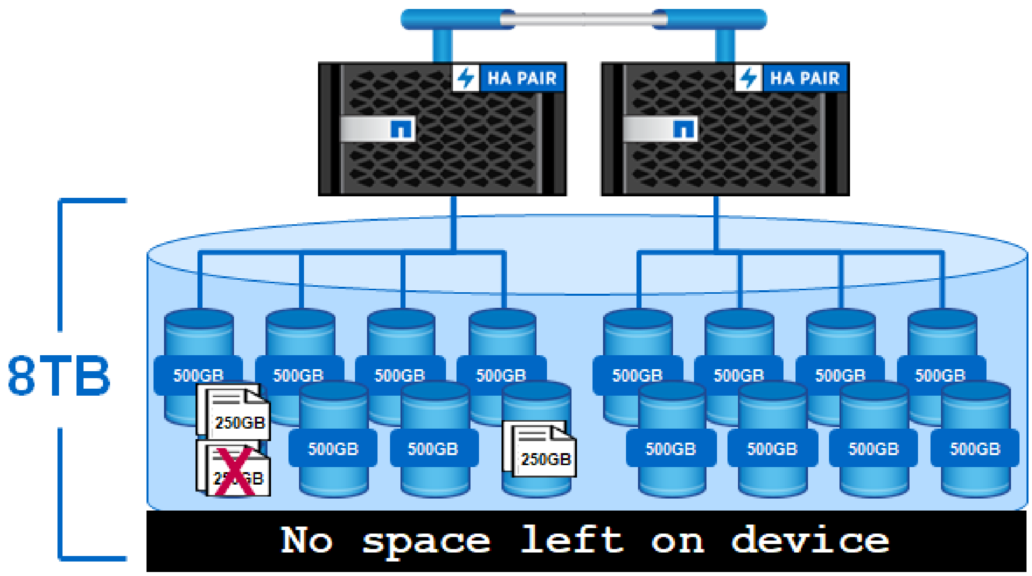
Figure 46) Effect of larger files in a FlexGroup member volume.



If a second 250GB file attempts to write to a member volume that is not able to honor the write, ONTAP reports back to the client that the volume is out of space, even if there are other member volumes with enough space to honor the write. Subsequent retries might find a viable member volume, but the space error remains intermittent. Remember, files in a FlexGroup volume do not stripe; they always write to a single FlexVol member volume. Therefore, there must be enough space in a single member volume to honor the write.

Note: ONTAP 9.6 and later versions help alleviate this scenario with the inclusion of elastic sizing.

Figure 47) Effect of larger files in a FlexGroup member volume; no space left on device.



Features like volume autogrow (available starting in ONTAP 9.3) and elastic sizing (starting in ONTAP 9.6) can help offset the impact of this type of workload. Eventually, when physical space is exhausted, more storage must be added. Also, adding nodes or disks to a FlexGroup volume is nondisruptive, easy, and fast.

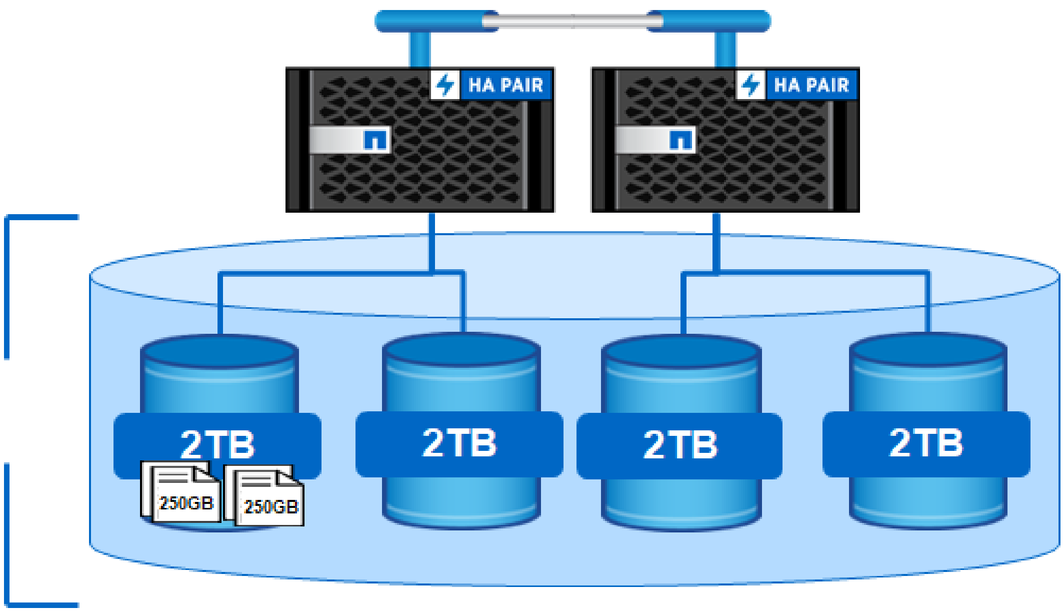
A better approach to sizing a FlexGroup volume is to analyze your workload and average file sizes before deploying a new FlexGroup volume or before allowing new workloads to access existing FlexGroup volumes. NetApp offers the XCP Migration Tool, which can quickly analyze files and report on sizes. For more information about XCP, see "Migrating to NetApp FlexGroup," later in this document.

After you have a good idea of what size files are going to land in a FlexGroup volume, you can make design decisions about how the volume should be sized at initial deployment.

Options include, but are not limited to the following:

- **Leave the member volume count at the defaults and grow the FlexGroup volume.**
  Size the total FlexGroup volume to a value large enough to accommodate member volume sizes that can handle the workload. In our example, the FlexGroup volume is 80TB, which would give 16-member volumes at 5TB per volume. However, this approach requires more physical capacity.
- **Manually reduce the member volume count and leave the FlexGroup capacity as is.**
  Rather than accept the default values from the automated commands, you can use the CLI to create a FlexGroup volume that is identical in total capacity but contains fewer (but larger) member volumes. In our example, reducing the member volume count to two per node in an 8TB FlexGroup would provide member volume sizes of 2TB each. This would reduce the number of volume affinities available (and could reduce the performance of the FlexGroup volume), but it would allow larger files to be placed.

Figure 48) Fewer, larger member volumes.



Use volume autogrow whenever possible to avoid "out of space" issues; ideally, use ONTAP 9.6 or later to apply the benefits of elastic sizing. Remember that volume autogrow and elastic sizing currently operate independently of one another—if volume autogrow is enabled, elastic sizing will not function on the FlexGroup volume. For decision making, see the section in this document on when to use elastic sizing versus volume autogrow.

In most cases, the preferred way to create a FlexGroup volume is to let ONTAP create it and/or set the member volume count through the automated methods (GUI or `-auto-provision-as` command). For guidance on what situations require manual FlexGroup volume creation, see the section above.

For an example of how to create a FlexGroup volume from the CLI and specify the number of members, see "Command Examples" later in this document.

Volume Affinity and CPU Saturation

To support concurrent processing, ONTAP assesses its available hardware at startup and divides its aggregates and volumes into separate classes called affinities. In general terms, volumes that belong to one affinity can be serviced in parallel with volumes that are in other affinities. In contrast, two volumes that are in the same affinity often have to take turns waiting for scheduling time (serial processing) on the node's CPU.

A node's affinities are viewed with the advanced privilege nodeshell command `waffinity_stats -g`.

In ONTAP 9.3 and earlier, a node had up to eight affinities available (four per aggregate).

```
cluster::> set -privilege advanced

cluster::*> node run * waffinity_stats -g

     Waffinity configured with:
                                      # AGGR affinities :    2
        # AGGR_VBN_RANGE affinities / AGGR_VBN affinity :    4
                       # VOL affinities / AGGR affinity :    4
          # VOL_VBN_RANGE affinities / VOL_VBN affinity :    4
             # STRIPE affinities / STRIPEGROUP affinity :    9
                # STRIPEGROUP affinities / VOL affinity :    1
                      # total AGGR_VBN_RANGE affinities :    8
                                 # total VOL affinities :    8
                       # total VOL_VBN_RANGE affinities :   32
                              # total STRIPE affinities :   72
                                     # total affinities : 149
                                              # threads :   19
```

This example NetApp FAS8080 EX node is reporting that it can support fully concurrent operations on eight separate volumes simultaneously. It also says that to reach that maximum potential, it would work best with at least two separate aggregates hosting four constituents each. Therefore, when you're building a new FlexGroup volume that is served by this node, ideally that new FlexGroup volume would include eight constituents on this node, evenly distributed across two local aggregates. If two such nodes are in the cluster, then a well-formed FlexGroup volume would consist of four aggregates (two per node) and 16 constituents (four per aggregate).

Starting in ONTAP 9.4, the number of available affinities increased to eight per aggregate (two aggregates, 16 per node) for high-end platforms like the AFF A700 and AFF A800:

```
cluster::*> node run * waffinity_stats -g

Waffinity configured with:
                                 # AGGR affinities :    2
   # AGGR_VBN_RANGE affinities / AGGR_VBN affinity :    8
                  # VOL affinities / AGGR affinity :    8
     # VOL_VBN_RANGE affinities / VOL_VBN affinity :    4
        # STRIPE affinities / STRIPEGROUP affinity :    3
           # STRIPEGROUP affinities / VOL affinity :    3
                 # total AGGR_VBN_RANGE affinities :   16
                            # total VOL affinities :   16
                  # total VOL_VBN_RANGE affinities :   64
                         # total STRIPE affinities :  144
                                # total affinities :  325
                                         # threads :   18
                                          # pinned :    0
                                # leaf sched pools :   18
                                     # sched pools :   21
```

However, storage administrators usually don't need to worry about volume affinities, because ONTAP deploys a FlexGroup volume according to best practices for most use cases. For guidance on when you might need to manually create a FlexGroup volume, see the section above.

Best Practice 9: Member Volume Count Recommendations
When you use the automated deployment methods (CLI or GUI), ONTAP provisions the appropriate number of member volumes in a FlexGroup volume. The best practice for member volume counts depends on platform, aggregate count, and ONTAP version, because the member volume counts are tied to volume affinity availability. Generally, let ONTAP decide on member volume counts for most use cases. ONTAP 9.4 and later versions provide up to 16 affinities per node (eight per aggregate) for higher-end systems. The best practice for volume count per node in a FlexGroup volume is to either use the automated methods or to match the member volume count per node to the number of VOL affinities listed in the `waffinity stats` command listed earlier.

To simplify the experience, the `vol create -auto-provision-as flexgroup` command (introduced in ONTAP 9.2), the `flexgroup deploy` command, and the ONTAP System Manager GUI handle this setup for the storage administrator.

Considerations When Deleting FlexGroup Volumes

In Data ONTAP 8.3, a feature called the volume recovery queue was added to help prevent accidental deletion of volumes by maintaining a recovery queue of deleted volumes for 12 hours. Although the volume is no longer accessible from clients and is hidden from administrator view at **admin privilege** levels, the space is still allocated, and the remnants of the volume remain in case an emergency recovery is needed. FlexGroup volumes use this recovery queue too, so space is not freed up until the recovery queue expires or is manually purged.

You can see deleted volumes from the command line by specifying `-type DEL` at the **diag privilege** level. (Neither volumes nor the recovery queue can be seen from the GUI.).

```
cluster::*> volume show -vserver DEMO -type DEL
Vserver   Volume       Aggregate    State      Type       Size  Available Used%
--------- ------------ ------------ ---------- ---- ---------- ---------- -----
DEMO      flexgroup__0001_2321
                       aggr1_node1  offline    DEL         5TB          -     -
DEMO      flexgroup__0002_2322
                       aggr1_node2  offline    DEL         5TB          -     -
```

Deleted volumes can also be seen with the `volume recovery-queue` command, also with **diag privileges**:

```
cluster::*> volume recovery-queue show
Vserver   Volume      Deletion Request Time     Retention Hours
--------- ----------- ------------------------  ---------------
DEMO      flexgroup__0001_2321
                      Tue May 01 17:14:14 2018               12
DEMO      flexgroup__0002_2322
                      Tue May 01 17:14:13 2018               12
2 entries were displayed.
```

To purge the volumes from the recovery queue manually, run the following commands:

```
cluster::*> volume recovery-queue purge -vserver DEMO -volume flexgroup__0001_2321
Queued private job: 4660

cluster::*> volume recovery-queue purge -vserver DEMO -volume flexgroup__0002_2322
Queued private job: 4661
```

To bypass the recovery queue when deleting a volume, use the `-force true flag` from the CLI with the `volume delete` command at the **advanced privilege** level. ONTAP System Manager does not support forced deletions.

Qtrees

ONTAP 9.3 introduced support in FlexGroup volumes for logical directories called qtrees. Qtrees allow a storage administrator to create folders from the ONTAP GUI or CLI to provide logical separation of data within a large bucket. Qtrees also provide flexibility in data management by enabling unique export policies, unique security styles, and granular statistics. Individual qtrees currently cannot be replicated through SnapMirror—all replication takes place at the volume level.

Qtrees are useful for home directory workloads, because folders can be named to reflect the user names of users accessing data, and dynamic shares can be created to provide access based on a username.

The following bullets give more information regarding qtrees in FlexGroup volumes.

- Qtrees are distributed across a FlexGroup volume the same way as in a normal folder.
- Qtrees are created and managed the same way as a FlexVol qtree is managed.
- A maximum of 4,995 qtrees is supported per FlexGroup volume. Quota monitoring and enforcement (enforcement in ONTAP 9.5 and later) can be applied at the qtree or user level.

Note: ONTAP 9.5 added quota enforcement support.

For more information about tree quotas and enforcement, see the section in this document on managing quotas.

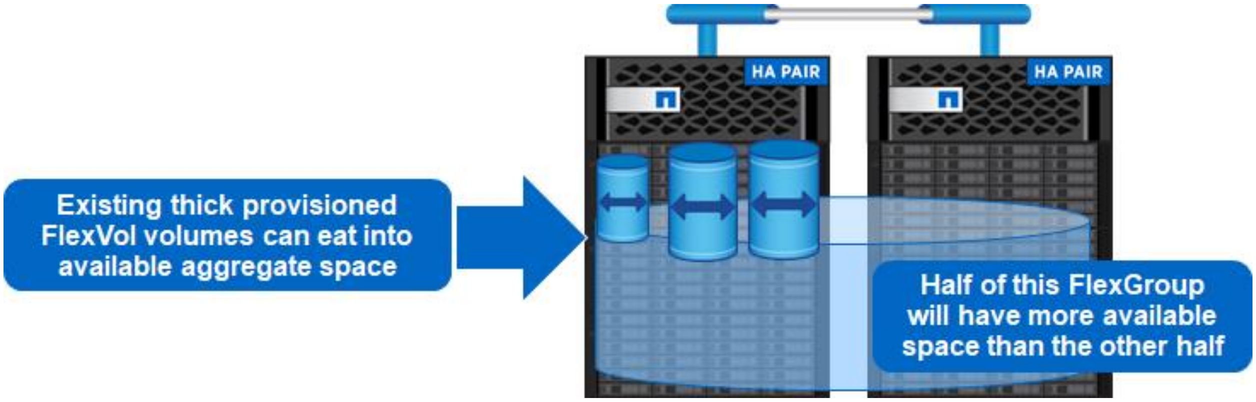
Deploying a FlexGroup Volume on Aggregates with Existing FlexVol Volumes

Because a FlexGroup volume can span multiple aggregates in a cluster and can coexist in the same SVM as normal FlexVol volumes, it is possible that a FlexGroup volume might have to share an aggregate with preexisting FlexVol volumes. Therefore, it's necessary to consider the factors described in this section when you're deploying a FlexGroup volume.

Consider the Capacity Footprint of the Existing FlexVol Volumes

A FlexGroup volume can span multiple aggregates that might not have the same number of FlexVol volumes on them. Therefore, the aggregates might have disparate free space that can affect the ingest distribution of a FlexGroup volume that has space guarantees disabled, because the existing FlexVol volume capacity might eat into the FlexGroup volume's capacity. For example, if aggr1 on node1 has four FlexVol volumes at 1TB each and aggr2 on node2 has two FlexVol volumes at 1TB each, then node1's aggregate would have 2TB less space than node2. If you deploy a FlexGroup volume that spans both nodes and is overprovisioned to fill both aggregates, then node1's member volumes already have "space used" in their capacity reports, which would cause node2's members to absorb most of the ingest of data until the capacity used is even across all member volumes. For more information, see "Effect of Overprovisioning or Thin Provisioning in a FlexGroup Volume." Also, in ONTAP versions earlier than 9.6, keep in mind that when thick-provisioned volumes are present in any aggregate that the FlexGroup attempts to span, the aggregates must have at least 3% free space available. For more information, see "Aggregate Layout Considerations" earlier in this document.

Figure 49) How FlexVol capacity can affect FlexGroup load distribution.



Note: This is an issue only if the FlexGroup volume is thin provisioned. Space-guaranteed FlexGroup volumes would not have other volumes eating into the space footprint. However, space-guaranteed FlexGroup volumes might not be created as large as desired if other volumes in the system prevent the space from being allocated.

Consider the Performance Impact of the Existing FlexVol Volumes

When you deploy a FlexGroup volume, it's also important to consider the amount of work the existing FlexVol volumes are doing. If a set of FlexVol volumes on one node is being hit heavily at given times, that can negatively affect the performance of a FlexGroup volume that spans the same nodes and aggregates as the existing FlexVol volumes. This is similar to the impact that can be seen with FlexVol volumes, but because a FlexGroup volume can span multiple nodes, the performance impact might appear to be intermittent from the client perspective, depending on which node the data I/O is occurring on.

Consider the Volume Count Limits

ONTAP places a volume count of 1,000 per node in a cluster. FlexGroup volumes can contain multiple FlexVol volumes that count against this limit. If you have existing FlexVol volumes, be sure to verify that adding FlexGroup volumes to the mix won't affect the volume count limits.

Best Practice 10: Deploying FlexGroup Volumes with Existing FlexVol Volumes in Place
Before deploying a FlexGroup volume, be sure to use the performance headroom features in NetApp Active IQ Performance Manager and ONTAP System Manager to review which nodes are being more heavily used. If there is an imbalance, use nondisruptive volume moves to migrate "hot" volumes to other less-utilized nodes to achieve as balanced a workload across nodes as possible. Also, be sure to evaluate the free space on the aggregates to be used with the FlexGroup volume and make sure that the available space is roughly equivalent. If the effect of volume count limit is a potential factor, create the FlexGroup volumes across nodes that have room to add more new volumes, or use nondisruptive volume moves to relocate volumes and balance out volume counts.

8.11 General NAS and High-File-Count Considerations

This section covers general NAS and high-file-count environment considerations.

Inode Count Considerations

An inode in ONTAP is a pointer to any file or folder within the file system, including Snapshot copies. Each FlexVol volume has a finite number of inodes and has an absolute maximum of 2,040,109,451. The default or maximum number of inodes on a FlexVol volume depends on the volume size and has a ratio of one inode to 32KB of capacity. Inodes can be increased after a FlexVol volume has been created and can be reduced starting in Data ONTAP 8.0.

When a volume inode count reaches 21,251,126, it remains at that default value, regardless of the size of the FlexVol volume. This feature mitigates potential performance issues, but it should be considered when you design a new FlexGroup volume. The FlexGroup volume can handle up to 400 billion files and 200 FlexVol member volumes, but the default inode count for 200 FlexVol members in a FlexGroup volume is 4,250,225,200. This count is based on the following formula:

```
200 member volumes * 21,251,126 default inodes per member = 4,250,225,200 total default inodes
```

If the FlexGroup volume requires more inodes than what is presented as a default value, the inodes must be increased by using the `volume modify -files` command.

Best Practice 11: Inode Count in a FlexGroup Volume (Before ONTAP 9.3)
When deciding where to place files, ONTAP versions earlier than 9.3 did not consider inode counts in the ingest calculations for data that is written into a FlexGroup volume. Thus, a member FlexVol volume could run out of inodes before other members run out of inodes, which would result in an overall “out of inodes” error for the entire FlexGroup volume. ONTAP 9.3 introduced inode count consideration for ingest of files to help avoid member volumes running out of inodes prematurely. If you’re running ONTAP 9.2 or earlier, NetApp strongly recommends increasing the default inode count in the FlexGroup volume before using it in production. The recommended value depends on workload, but do not set the value to the maximum at the start. If you do, you won’t have room to increase later without adding member volumes. If possible, upgrade to ONTAP 9.3 or later to take advantage of the new ingest calculations for high-file-count environments.

Table 9 shows a sample of FlexVol sizes, inode defaults, and maximums.

Table 9) Inode defaults and maximums according to FlexVol size.

FlexVol Size	Default Inode Count	Maximum Inode Count
20MB*	566	4,855
1GB*	31,122	249,030
100GB*	3,112,959	24,903,679
1TB	21,251,126	255,013,682
10TB	21,251,126	2,040,109,451
100TB	21,251,126	2,040,109,451

*FlexGroup member volumes should not be any smaller than 100GB in size.

When you use a FlexGroup volume, the total default inode count depends on both the total size of the FlexVol members and the number of FlexVol members in the FlexGroup volume.

Table 10 shows various examples of FlexGroup configurations and the resulting default inode counts.

Table 10) Inode defaults resulting from FlexGroup member sizes and member volume counts.

Member Volume Size	Member Volume Count	Default Inode Count (FlexGroup)
100GB	8	24,903,672
100GB	16	49,807,344
1TB	8	170,009,008
1TB	16	340,018,016
100TB	8	170,009,008
100TB	16	340,018,016

High File Counts, Low Capacity Needs

ONTAP allocates a default inode and maximum inode count based on volume capacity. In Table 9, member volumes smaller than 10TB will not be able to achieve the two billion inodes available to a FlexVol volume. To get two billion inodes per member volume, the member volume capacity would need to be 10TB or greater. In a FlexGroup volume with eight member volumes, that would support 16 billion files, but would also provision 80TB of storage.

This can present a challenge to high-file-count environments, because file sizes might be small and won't require that much total capacity. For example, if all files in a workload are 288 bytes each in size, 16 billion files would use up only about 4.6TB. To achieve a maximum of 16 billion files, you'd need at least eight member volumes that are 10TB each, which would take up 80TB. 4.6TB is around 6% of the total capacity needed to contain 16 billion files in ONTAP in an 80TB FlexGroup volume. In these cases, you would need to disable space guarantees to avoid reserving about 75TB of unused space.

When deploying high file counts that will use up little capacity, there are two main options for deploying the FlexGroup volume.

- **Deploy the FlexGroup volume with 10TB or greater member volumes with thin provisioning.** Thin provisioning a volume simply means that you are telling ONTAP a volume will be a certain size, but that the size will not be guaranteed in the file system. This provides flexibility in the file system to limit storage allocation to physical space. However, other volumes in the aggregate can affect the free capacity, so it's important to monitor available aggregate space when using thin provisioning.
- **Manually create the FlexGroup volume with more member volumes than the default.** If you want to keep space guarantees for volumes, another option for high-file-count/small capacity environments is to create more member volumes in a FlexGroup volume.

  Because inode counts are limited per FlexVol member volume according to capacity, adding more smaller member volumes can provide for higher file counts at the same capacity. The following table shows some possible configurations. For more information about manual creation of FlexGroup volumes, see "Command Examples" later in this document.

Table 11) High-file-count/small capacity footprint examples—increasing member volume counts.

Total FlexGroup Size	Member Volume Count (Size)	Maximum Inode Count (Entire FlexGroup)
80TB (no space guarantee)	8 (10TB)	16,320,875,608
64TB (space guarantee enabled)	32 (2TB)	16,320,875,608
64TB (space guarantee enabled)	64 (1TB)	16,320,875,608

Planning for Inode Counts in ONTAP

With utilities like the NetApp [XCP Migration Tool](#) (using the scan feature), you can evaluate your file count usage and other file statistics to help you make informed decisions about how to size your inode counts in your new FlexGroup volume. For more information about using XCP to scan files, contact ng-xcp-support@netapp.com.

Viewing Inodes and Available Inodes

In ONTAP, you can view inode counts per volume by using the following command with advanced privileges:

```
cluster::*> volume show -volume flexgroup -fields files,files-used
vserver volume    files     files-used
------- --------- --------- ----------
SVM     flexgroup 170009008 823
```

You can also use the classic `df -i` command. To show all member volumes, use an asterisk with the volume name in **diag privilege**:

```
cluster ::*> df -i Tech_ONTAP*
Filesystem              iused      ifree  %iused  Mounted on                 Vserver
/vol/Tech_ONTAP/        10193  169998815      0%  /techontap                 DEMO
/vol/Tech_ONTAP__0001/    923   21250203      0%  /techontap                 DEMO
/vol/Tech_ONTAP__0002/   4177   21246949      0%  ---                        DEMO
/vol/Tech_ONTAP__0003/    878   21250248      0%  ---                        DEMO
/vol/Tech_ONTAP__0004/    848   21250278      0%  ---                        DEMO
/vol/Tech_ONTAP__0005/    750   21250376      0%  ---                        DEMO
/vol/Tech_ONTAP__0006/    972   21250154      0%  ---                        DEMO
/vol/Tech_ONTAP__0007/    879   21250247      0%  ---                        DEMO
/vol/Tech_ONTAP__0008/    766   21250360      0%  ---                        DEMO
```

Tips for Managing Slow Directory Listings in High-File-Count Environments

Some workflows in high-file-count environments include running `find`, `ls`, or some other read metadata-heavy operation on an existing dataset. Generally, this process is inefficient and can take a long time to complete. If it's necessary to run these operations, there are a few things you can try to help speed things along.

Generally speaking, the issue with these types of operations are client, protocol, or network related. The storage rarely is the bottleneck for read metadata slowness. ONTAP is able to multithread read metadata operations. With `ls` operations, `getattr` requests are sent one at a time, in serial, which means for millions of `getattr` operations, there might be millions of network requests to the storage. Each network request will incur n milliseconds of latency, which adds up over time. So, there are a few ways to speed these up:

- **Send more `getattr` requests at a time.** By itself, `ls` can't send requests in parallel. But with utilities like the XCP Migration Tool, it is possible to send multiple threads across the network to greatly speed up `ls` operations. Using XCP scan can help with speed, depending on what the `ls` output is being used for later. For example, if you need the user permissions/owners of the files, using `ls` by itself might be a better fit. But for sheer listing of file names, XCP scan is preferable.
- **Add more network hardware (for example, 100GB instead of 10GB) to reduce round-trip time (RTT).** With larger network pipes, more traffic can be pushed over the network, thus reducing load and potentially reducing overall RTT. With million of operations, even shaving off a millisecond of latency can add up to a large amount of time saved for workloads.
- **Run `ls` without unnecessary options, such as highlighting/colors.** When running `ls`, the default behavior is to add sorting, colors, and highlighting for readability. These add work for the operation, so it might make sense to run `ls` with the `–f` option to avoid those potentially unnecessary features.

- **Cache `getattr` operations on the client more aggressively.** Client-side caching of attributes can help reduce the network traffic for operations, as well as bringing the attributes local to the client for operations. Clients manage NFS caches differently, but in general, avoid setting `noac` on NFS mounts for high-file-count environments. Also, keep `actimeo` to a level no less than 30 seconds.
- **Create FlexCache volumes.** NetApp FlexCache volumes are able to create instant caches for read-heavy workloads. Creating FlexCache volumes for workloads that do a lot of read metadata operations, such as `ls`, can have the following benefits:
  - For local clusters, it can help offload the read metadata operations from the origin volume to the cache volumes, and, as a result, frees the origin volume up for regular reads and writes.
  - FlexCache volumes can reside on any node in a cluster, so creating FlexCache volumes makes the use of cluster nodes more efficient by allowing multiple nodes to participate in these operations, in addition to moving the read metadata operations away from the origin node.
  - For remote clusters across a WAN, FlexCache volumes can provide localized NFS caches to help reduce WAN latency, which can greatly improve performance for read-metadata-heavy workloads.

  When using FlexCache volumes to help read metadata workloads, be sure to disable `fastreaddir` on the nodes that use FlexCache.

```
cluster::> node run "priv set diag; flexgroup set fastreaddir=false persist
```

Note: For this to take effect, a reboot/storage failover is required.

Starting in ONTAP 9.7, FlexGroup volumes can be origins for FlexCache volumes. For more information about FlexCache volumes, see TR-4743: FlexCache in NetApp ONTAP.

Effect of Being Out of Inodes

When a volume runs out of inodes, no more files can be created in that volume until the number of inodes is increased or existing inodes are freed.

When a volume runs out of inodes, the cluster triggers an EMS event (`callhome.no.inodes`), and a NetApp AutoSupport® message is triggered.

```
Message Name: callhome.no.inodes
Severity: ERROR

Corrective Action: Modify the volume's maxfiles (maximum number of files) to increase the inodes
on the affected volume. If you need assistance, contact NetApp technical support.

Description: This message occurs when a volume is out of inodes, which refer to individual files,
other types of files, and directories. If your system is configured to do so, it generates and
transmits an AutoSupport (or 'call home') message to NetApp technical support and to the
configured destinations. Successful delivery of an AutoSupport message significantly improves
problem determination and resolution.
```

Note: In a FlexGroup volume, if any member volume runs out of inodes, the entire FlexGroup volume reports being out of inodes, even if other members have available inodes.

Starting in ONTAP 9.3, a FlexGroup volume takes per-member inode numbers into account when deciding which member volumes are most optimal for data ingest.

In addition to the callhome message, the following EMS messages are available:

```
Message Name: fg.inodes.member.nearlyFull
Severity: ALERT

Corrective Action: Adding capacity to the FlexGroup by using the "volume modify -files +X"
command is the best way to solve this problem. Alternatively, deleting files from the FlexGroup
might work, although it can be difficult to determine which files have landed on which
constituent.
```



```
Description: This message occurs when a constituent within a FlexGroup is almost out of inodes.
This constituent will receive far fewer new create requests than average, which might impact the
FlexGroup's overall performance, because those requests are routed to constituents with more
inodes.
```

```
Message Name: fg.inodes.member.full
Severity: ALERT

Corrective Action: Adding capacity to the FlexGroup by using the "volume modify -files +X"
command is the best way to solve this problem. Alternatively, deleting files from the FlexGroup
may work, but it is difficult to determine which files have landed on which constituent.

Description: This message occurs when a constituent with a FlexGroup has run out of inodes. New
files cannot be created on this constituent. This might lead to an overall imbalanced
distribution of content across the FlexGroup.
```

```
Message Name: fg.inodes.member.allOK
Severity: NOTICE

Corrective Action: (NONE)

Description: This message occurs when conditions that led to previous
"fg.inodes.member.nearlyFull" and "fg.inodes.member.full" events no longer apply for any
constituent in this FlexGroup. All constituents within this FlexGroup have sufficient inodes for
normal operation.
```

These messages can be used for monitoring, or for triggering scripts that automatically increase inode counts to help avoid space errors.

64-Bit File Identifiers

By default, NFS in ONTAP uses 32-bit file IDs. 32-bit file IDs are limited to 2,147,483,647 maximum unsigned integers. With the 2 billion inode limit in FlexVol, this value fits nicely into the architecture.

However, because FlexGroup volumes can officially support up to 400 billion files in a single container (and theoretically, many more), the implementation of 64-bit file IDs was needed. 64-bit file IDs support up to 9,223,372,036,854,775,807 unsigned integers.

Best Practice 12: 64-Bit File Identifiers
NetApp strongly recommends enabling the NFS server option `-v3-64bit-identifiers` at the a**dvanced privilege** level before you create a FlexGroup volume, especially if your file system exceeds or might exceed the two billion inode threshold.

The 64-bit file identifier option is set to “off/disabled” by default. This was by design, to make certain that legacy applications and operating systems that require 32-bit file identifiers were not unexpectedly affected by ONTAP changes before administrators could properly evaluate their environments. Check with your application and OS vendor for their support for 64-bit file IDs before enabling them. Alternatively, create a test SVM and enable it to see how applications and clients react with 64-bit file IDs. Most modern applications and operating systems can handle 64-bit file IDs without issue.

This option can currently be enabled only with the advanced privilege level on the command line:

```
cluster::> set advanced
cluster::*> nfs server modify -vserver SVM -v3-64bit-identifiers enabled
```

After enabling or disabling this option, you must remount all clients. Otherwise, because the file system IDs change, the clients might receive stale file handle messages when attempting NFS operations. For more information about how enabling or disabling FSID change options can affect SVMs in high-file-count environments, see “How FSIDs Operate with SVMs in High-File-Count Environments,” later in this document.

If a FlexGroup volume does not exceed two billion files, you can leave this value unchanged. However, to prevent any file ID conflicts, the inode maximum on the FlexGroup volume should also be increased to no more than 2,147,483,647.

```
cluster::*> vol show -vserver SVM -volume flexgroup -fields files
```

Note: This option does not affect SMB operations and is unnecessary with volumes that use only SMB.

NFSv3 Versus NFSv4.x – File IDs

NFSv3 and NFSv4.x use different file ID semantics. Now that FlexGroup volumes support NFSv4.x, ONTAP 9.7 provides two different options for enabling/disabling 64-bit file IDs.

When you use both NFSv3 and NFSv4.x in an SVM and you want the 64-bit ID option to apply to both protocols, you must set both options.

If only one option is set and volumes are accessed by both protocols, you might see undesired behavior between protocols. For instance, NFSv3 might be able to create and view more than 2 billion files, whereas NFSv4.x would throw an error.

The options are:

```
-v3-64bit-identifiers [enabled/disabled]
-v4-64bit-identifiers [enabled/disabled]
```

Using Quota Enforcement to Limit File Count

Starting with ONTAP 9.5, it’s possible to set up a quota policy that prevents a FlexGroup volume from exceeding two billion files if 32-bit file handles are still being used by way of quota enforcement.

Because quota policies don’t apply to files created below the parent volume, create a qtree inside the FlexGroup volume. Then create a default quota rule with two billion files as the limit to help reduce the risk of users overrunning the 32-bit file ID limitations.

```
cluster::*> qtree create -vserver DEMO -volume FG4 -qtree twobillionfiles -security-style unix -
oplock-mode enable -unix-permissions 777
cluster::*> quota policy rule create -vserver DEMO -policy-name files -volume FG4 -type tree -
target "" -file-limit 2000000000
cluster::*> quota on -vserver DEMO -volume FG4
[Job 15906] Job is queued: "quota on" performed for quota policy "tree" on volume "FG4" in
Vserver "DEMO".
cluster::*> quota resize -vserver DEMO -volume FG4
[Job 15907] Job is queued: "quota resize" performed for quota policy "tree" on volume "FG4" in
Vserver "DEMO".
cluster::*> quota report -vserver DEMO -volume FG4
Vserver: DEMO

                                    ----Disk----  ----Files-----   Quota
Volume   Tree      Type    ID        Used  Limit    Used   Limit   Specifier
-------  --------  ------  -------  -----  -----  ------  ------   ---------
FG4      twobillionfiles
                   tree    1           0B      -       1
                                                   2000000000   twobillionfiles
FG4                tree    *           0B      -       0
                                                   2000000000   *
2 entries were displayed.
```

After that is done, use file permissions to limit access, preventing users from creating files at the volume level. Apply SMB shares to the qtree rather than the volume, and mounts should occur at the qtree level.

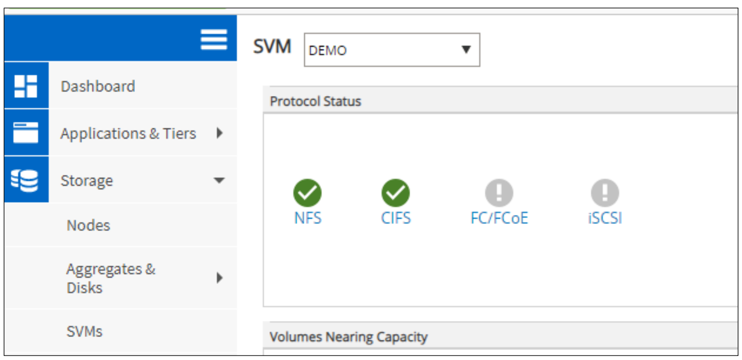
Then, as files are created in the qtree, they count against the limit.


```
[root@centos7 home]# cd /FG4/twobillionfiles/
[root@centos7 twobillionfiles]# ls
[root@centos7 twobillionfiles]# touch new1
[root@centos7 twobillionfiles]# touch new2
[root@centos7 twobillionfiles]# touch new3
[root@centos7 twobillionfiles]# ls
new1  new2  new3
cluster::*> quota report -vserver DEMO -volume FG4
Vserver: DEMO

                                    ----Disk----  ----Files-----   Quota
Volume   Tree       Type    ID        Used  Limit    Used   Limit    Specifier
-------  ---------  ------  -------  -----  -----  ------  ------   ---------
FG4      twobillionfiles
                    tree    1           0B      -       4
                                                      2000000000   twobillionfiles
FG4                 tree    *           0B      -       0
                                                      2000000000   *
```

System Manager Support for the 64-Bit File ID Option (Classic View)

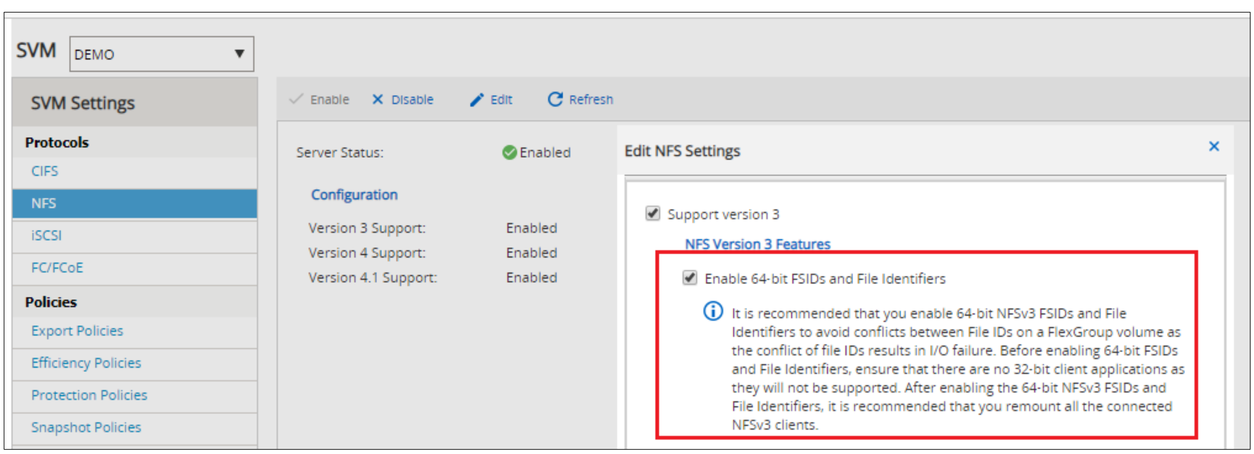
Starting in ONTAP 9.2, it is also possible to enable or disable the NFS server option from ONTAP System Manager. To do this, select Storage > SVMs. Select the desired SVM and then click NFS.

Figure 50) SVM Settings tab.



Click Edit. The dialog box that opens contains a checkbox that you can use to enable or disable 64-bit file identifiers.

Figure 51) Enable or disable 64-bit file identifiers in System Manager.



After you enable or disable this option, all clients must be remounted, because the file system IDs change and the clients might receive stale file handle messages when attempting NFS operations.

ONTAP System Manager: 9.7 and Later

ONTAP 9.7 introduced a new System Manager interface based on REST API capabilities. Because the 64-bit file ID option does not currently exist in the REST API, the only way to modify it in System Manager is to use the classic view. Alternately, use the CLI.

Impact of File ID Collision

If 64-bit file IDs are not enabled, the risk for file ID collisions increases. When a file ID collision occurs, the impact can range from a stale file handle error on the client, to the failure of directory and file listings, to the entire failure of an application. Usually, it is imperative to enable the 64-bit file ID option when you use FlexGroup volumes.

You can check a file's ID from the client using the `stat` command. When an inode or file ID collision occurs, it might look like the following. The inode is 3509598283 for both files.

```
[root@client]# stat libs/
  File: `libs/'
  Size: 12288           Blocks: 24         IO Block: 65536  directory
Device: 4ch/76d Inode: 3509598283  Links: 3
Access: (0755/drwxr-xr-x)  Uid: (60317/   user1)   Gid: (10115/      group1)
Access: 2017-01-06 16:00:28.207087000 -0700
Modify: 2017-01-06 15:46:50.608126000 -0700
Change: 2017-01-06 15:46:50.608126000 -0700

[root@client example]# stat iterable/
  File: `iterable/'
  Size: 4096            Blocks: 8          IO Block: 65536  directory
Device: 4ch/76d Inode: 3509598283  Links: 2
Access: (0755/drwxr-xr-x)  Uid: (60317/   user1)   Gid: (10115/      group1)
Access: 2017-01-06 16:00:44.079145000 -0700
Modify: 2016-05-05 15:12:11.000000000 -0600
Change: 2017-01-06 15:23:58.527329000 -0700
```

A collision can result in issues such as circular directory structure errors on the Linux client and an inability to remove files.

```
rm: WARNING: Circular directory structure.
This almost certainly means that you have a corrupted file system.
NOTIFY YOUR SYSTEM MANAGER.
The following directory is part of the cycle:
  '/directory/iterable'

rm: cannot remove '/directory': Directory not empty
```

Note: This option does not affect SMB operations and is unnecessary with volumes using only SMB.

Effects of File System ID (FSID) Changes in ONTAP

NFS uses a file system ID (FSID) when interacting between client and server. This FSID lets the NFS client know where data lives in the NFS server's file system. Because ONTAP can span multiple file systems across multiple nodes by way of junction paths, this FSID can change depending on where data lives. Some older Linux clients can have problems differentiating these FSID changes, resulting in failures during basic attribute operations, such as `chown` and `chmod`.

An example of this issue can be found in bug 671319. If you disable the FSID change with NFSv3, be sure to enable the `-v3-64bit-identifiers` option in ONTAP 9 (see "64-Bit File Identifiers"). But keep in mind that this option could affect older legacy applications that require 32-bit file IDs.

How FSIDs Operate with SVMs in High-File-Count Environments

The FSID change option for NFSv3 and NFSv4.x provides FlexVol and FlexGroup volumes with their own unique file systems, which means that the number of files allowed in the SVM is dictated by the number of volumes. However, disabling the FSID change options will cause the 32-bit or 64-bit file identifiers to apply to the SVM itself, meaning that the two billion file limit with 32-bit would apply to all volumes. Therefore, the SVM would be limited to two billion files, rather than the FlexVol or FlexGroup volume. Leaving the FSID change option enabled allows volumes to operate as independent file systems with their own dedicated file counts.

NetApp recommends leaving the FSID change option enabled with FlexGroup volumes to help prevent file ID collisions.

How FSIDs Operate with Snapshot Copies

When a Snapshot copy of a volume is created, a copy of a file's inodes is preserved in the file system for access later. The file theoretically exists in two locations.

With NFSv3, even though there are two copies of essentially the same file, the FSIDs of those files are not identical. FSIDs of files are formulated by using a combination of NetApp WAFL inode numbers, volume identifiers, and Snapshot IDs. Because every Snapshot copy has a different ID, every Snapshot copy of a file has a different FSID in NFSv3, regardless of the setting of the `-v3-fsid-change` option. The NFS RFC specification does not require FSIDs for a file to be identical across file versions.

Note: The `-v4-fsid-change` option does not apply to FlexGroup volumes prior to ONTAP 9.7, because NFSv4 is unsupported with FlexGroup volumes in those releases.

Directory Size Considerations

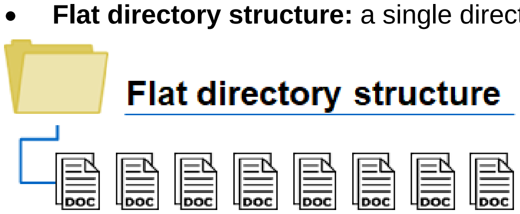
In ONTAP, there are limitations to the maximum directory size on disk. This limit is known as maxdirsize. The `maxdirsize` value for a volume is capped at 320MB, regardless of platform. This means that the memory allocation for the directory size can reach a maximum of only 320MB before a directory can no longer grow larger.

Best Practice 13: Recommended ONTAP Version for High-File-Count Environments
For high-file-count environments, use ONTAP 9.2 or later.

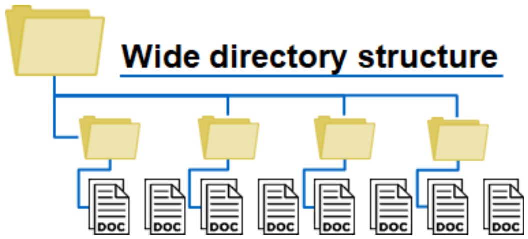
What Directory Structures Can Affect maxdirsize?

The `maxdirsize` value can be a concern when you're using flat directory structures, where a single folder contains millions of files at a single level. Folder structures where files, folders, and subfolders are interspersed have a low impact on `maxdirsize`. There are several directory structure methodologies.

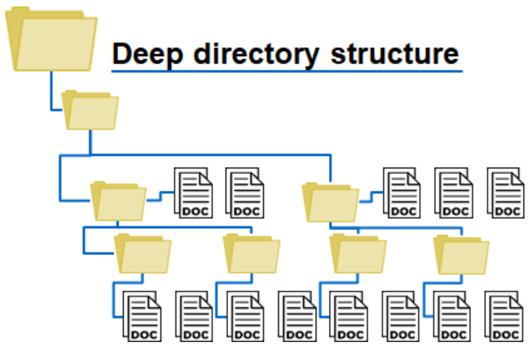
- **Flat directory structure:** a single directory with many files



- **Wide directory structure.** Many top-level directories with files spread across directories



- **Deep directory structures.** Fewer top-level directories, but with many subfolders; files spread across directories



How Flat Directory Structures Can Affect FlexGroup Volumes

Flat directory structures (many files in a single/few directories) have a negative impact on a wide array of file systems, whether they're NetApp systems or not. Areas of impact can include, but are not limited to:

- Memory pressure
- Network performance/latency (particularly during mass queries of files, `GETATTR` operations, `READDIR` operations, and so on)
- CPU utilization

FlexGroup volumes can also have an extra impact on `maxdirsize`. Unlike a FlexVol volume, a FlexGroup volume uses remote hard links inside ONTAP to help redirect traffic. These remote hard links are what allow a FlexGroup volume to deliver scale-out performance and capacity in a cluster.

However, in flat directories, a higher ratio of remote hard links to local files is seen. These remote hard links count against the total `maxdirsize` value, so a FlexGroup volume might approach the `maxdirsize` limit faster than a FlexVol will.

For example, if a directory has millions of files in it and generates roughly 85% remote hard links for the file system, you can expect `maxdirsize` to be exhausted at nearly twice the amount as a FlexVol would.

Querying for Used maxdirsize Values

It is important to monitor and evaluate `maxdirsize` allocation in ONTAP. However, there are no commands for this specific to ONTAP. Instead, `maxdirsize` allocation would need to be queried from the client.

The following command from an NFS client would be able to retrieve the directory size information for a folder inside a FlexGroup volume for the 10 largest directories in a given mount point, while omitting Snapshot copies from the search.

```
# find /mountpoint -name .snapshot -prune -o -type d -ls -links 2 -prune | sort -rn -k 7 | head
```

The following example took less than a second on a dataset in folders with millions of files:


```
[root@centos7 /]# time find /flexgroup/manyfiles/ -name .snapshot -prune -o -type d -ls -links 2 
-prune | sort -rn -k 7 | head
787227871 328976 drwxr-xr-x   2 root     root     335544320 May 29 21:23 
/flexgroup/manyfiles/folder3/topdir_8/subdir_0
384566806 328976 drwxr-xr-x   2 root     root     335544320 May 29 13:14 
/flexgroup/manyfiles/folder3/topdir_9/subdir_0
3605793347 328976 drwxr-xr-x   2 root     root     335544320 May 29 21:23 
/flexgroup/manyfiles/folder3/topdir_0/subdir_0
3471151639 328976 drwxr-xr-x   2 root     root     335544320 May 29 13:45 
/flexgroup/manyfiles/folder3/topdir_4/subdir_0
2532103978 328976 drwxr-xr-x   2 root     root     335544320 May 29 14:16 
/flexgroup/manyfiles/folder3/topdir_2/subdir_0
2397949155 328976 drwxr-xr-x   2 root     root     335544320 May 29 14:15 
/flexgroup/manyfiles/folder3/topdir_1/subdir_0
1994984460 328976 drwxr-xr-x   2 root     root     335544320 May 29 13:43 
/flexgroup/manyfiles/folder3/topdir_6/subdir_0
1860674357 328976 drwxr-xr-x   2 root     root     335544320 May 29 13:18 
/flexgroup/manyfiles/folder3/topdir_5/subdir_0
1458235096 328976 drwxr-xr-x   2 root     root     335544320 May 29 14:25 
/flexgroup/manyfiles/folder3/topdir_3/subdir_0
1325327652 328976 drwxr-xr-x   2 root     root     335544320 May 29 14:25 
/flexgroup/manyfiles/folder3/topdir_7/subdir_0

real    0m0.055s
user    0m0.002s
sys     0m0.035s
```

Using XCP to Check maxdirsize

The XCP Migration Tool is mostly considered a RapidData mover, but it also derives value in its robust file scanning abilities. XCP is able to run `find` commands in parallel as well, so the previous examples can be run even faster on the storage system.

The following XCP command example allows you to run `find` only on directories with more than 2,000 entries:

```
# xcp diag find --branch-match True -fmt "'{size} {name}'.format(size=x.digest, name=x)" 
localhost:/usr 2>/dev/null | awk '{if ($1 > 2000) print $1 " " $2}'
```

And this XCP command helps you find the directory size values:

```
# xcp -match "type == d" -fmt "'{} {}'.format(used, x)" localhost:/usr | awk '{if ($1 > 100000) 
print}' | sort -nr
```

When XCP looks for the directory size values, it scans the file system first. Here’s an example:

```
[root@XCP flexgroup]#  xcp -match "type == d" -fmt "'{} {}'.format(used, x)" 
10.193.67.219:/flexgroup_16/manyfiles | awk '{if ($1 > 100000) print}' | sort -nr

 660,693 scanned, 54 matched, 123 MiB in (24.6 MiB/s), 614 KiB out (122 KiB/s), 5s
 1.25M scanned, 58 matched, 234 MiB in (22.1 MiB/s), 1.13 MiB out (109 KiB/s), 10s
 …
 31.8M scanned, 66 matched, 5.83 GiB in (4.63 MiB/s), 28.8 MiB out (22.8 KiB/s), 7m52s

Filtered: 31816172 did not match
31.8M scanned, 66 matched, 5.83 GiB in (12.6 MiB/s), 28.8 MiB out (62.4 KiB/s), 7m53s.
336871424 10.193.67.219:/flexgroup_16/manyfiles/folder3/topdir_9/subdir_0
336871424 10.193.67.219:/flexgroup_16/manyfiles/folder3/topdir_8/subdir_0
336871424 10.193.67.219:/flexgroup_16/manyfiles/folder3/topdir_7/subdir_0
336871424 10.193.67.219:/flexgroup_16/manyfiles/folder3/topdir_6/subdir_0
336871424 10.193.67.219:/flexgroup_16/manyfiles/folder3/topdir_5/subdir_0
336871424 10.193.67.219:/flexgroup_16/manyfiles/folder3/topdir_4/subdir_0
336871424 10.193.67.219:/flexgroup_16/manyfiles/folder3/topdir_3/subdir_0
```


Best Practice 14: Directory Structure Recommendation
• For the best performance, avoid flat directory structures in ONTAP if at all possible. Wide or deep directory structures work best, as long as the path length of the file or folder does not exceed NAS protocol standards. • If flat directory structures are unavoidable, pay close attention to the `maxdirsize` values for the volume and increase them as necessary. • NFS path lengths are defined by the client OS. • For information about SMB path lengths, see this Microsoft Dev Center link.

Number of Files That Can Fit into a Single Directory with the Default maxdirsize

To determine how many files can fit into a single directory with the default `maxdirsize` setting, use this formula:

- Memory in KB * 53 *25%

Since `maxdirsize` is set to 320MB by default on larger systems, the maximum number of files in a single directory would be 4,341,760 for SMB and NFS. NetApp strongly recommends that you keep the `maxdirsize` value as low as possible, but no higher than 80% of the 320MB limit (256MB, or around 3.4 million files).

Event Management System Messages Sent When maxdirsize Is Exceeded

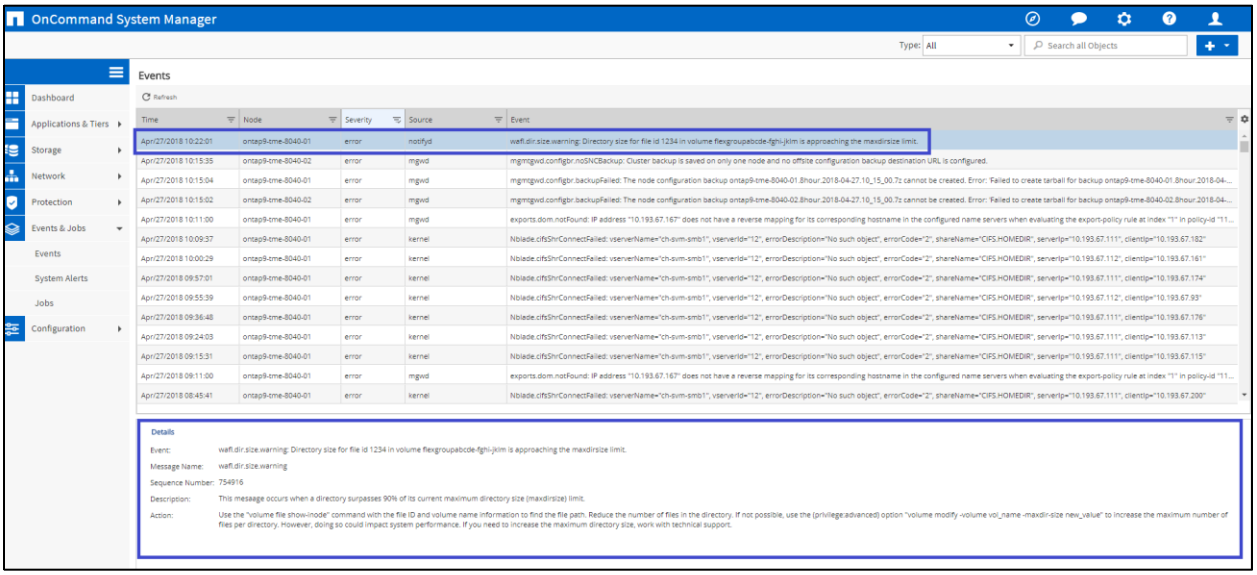
The following event management system (EMS) messages are triggered when `maxdirsize` is either exceeded or close to being exceeded. Warnings are sent at 90% of the `maxdirsize` value and can be viewed with the `event log show` command or with the ONTAP System Manager event section. Active IQ Unified Manager can be used to monitor `maxdirsize`, trigger alarms, and send a notification before the 90% threshold (see Figure 52). These event management system messages also support SNMP traps.

```
Message Name: wafl.dir.size.max
Severity: ERROR

Corrective Action: Use the "volume file show-inode" command with the file ID and volume name
information to find the file path. Reduce the number of files in the directory. If not possible,
use the (privilege:advanced) option "volume modify -volume vol_name -maxdir-size new_value" to
increase the maximum number of files per directory. However, doing so could impact system
performance. If you need to increase the maximum directory size, work with technical support.

Description: This message occurs after a directory has reached its maximum directory size
(maxdirsize) limit.

Supports SNMP trap: true

Message Name: wafl.dir.size.max.warning
Severity: ERROR

Corrective Action: Use the "volume file show-inode" command with the file ID and volume name
information to find the file path. Reduce the number of files in the directory. If not possible,
use the (privilege:advanced) option "volume modify -volume vol_name -maxdir-size new_value" to
increase the maximum number of files per directory. However, doing so could impact system
performance. If you need to increase the maximum directory size, work with technical support.

Description: This message occurs when a directory has reached or surpassed 90% of its current
maximum directory size (maxdirsize) limit, and the current maxdirsize is less than the default
maxdirsize, which is 1% of total system memory.

Supports SNMP trap: true

Message Name: wafl.dir.size.warning
Severity: ERROR
```



```
Corrective Action: Use the "volume file show-inode" command with the file ID and volume name
information to find the file path. Reduce the number of files in the directory. If not possible,
use the (privilege:advanced) option "volume modify -volume vol_name -maxdir-size new_value" to
increase the maximum number of files per directory. However, doing so could impact system
performance. If you need to increase the maximum directory size, work with technical support.

Description: This mesaage occurs when a directory surpasses 90% of its current maximum directory
size (maxdirsize) limit.

Supports SNMP trap: true
```

Figure 52) ONTAP System Manager event screen with `maxdirsize` warning.



Impact of Increasing maxdirsize

When a single directory contains many files, the lookups (such as in a “find” operation) can consume large amounts of CPU and memory. Starting in ONTAP 9.2, directory indexing creates an index file for directory sizes exceeding 2MB to help offset the need to perform so many lookups and avoid cache misses. Usually, this helps large directory performance. However, for wildcard searches and `readdir` operations, indexing is not of much use.

Best Practice 15: Maxdirsize Maximums
Values for `maxdirsize` are hard coded to not exceed 4GB. To avoid performance issues, NetApp recommends setting `maxdirsize` values no higher than 1GB if possible.

Do FlexGroup Volumes Avoid maxdirsize Limitations?

In FlexGroup volumes, each member volume has the same `maxdirsize` setting. Even though a directory could potentially span multiple FlexVol member volumes and nodes, the `maxdirsize` performance effect can still come into play, because directory size is the key component, not individual FlexVol volumes. Directory size is tied to the parent volume and does not divide up across other member volumes. Therefore, the overall size of a directory is still an issue. Thus, FlexGroup volumes do not provide relief for environments facing `maxdirsize` limitations.

Best Practice 16: Avoiding maxdirsize Issues
Newer platforms offer more memory and CPU capacity, and AFF systems provide performance benefits. However, the best way to reduce the performance effect in directories with large numbers of files is to spread files across more directories in a file system.

Effect of Exceeding maxdirsize

When `maxdirsize` is exceeded in ONTAP, an `out of space` error (`ENOSPC`) is issued to the client and an event management system message is triggered. To remediate this problem, a storage administrator must increase the `maxdirsize` setting or move files out of the directory. For more information about remediation, see KB 000002080 on the NetApp Support site. For examples of the `maxdirsize` event management system events, see Event Management System Examples.

Special Character Considerations

Most common text characters in Unicode (when they are encoded with UTF-8 format) use encoding that is equal to or smaller than three bytes. This common text includes all modern written languages, such as Chinese, Japanese, and German. However, with the popularity of special characters such as the emoji, some UTF-8 character sizes have grown beyond 3 bytes. For example, a trophy symbol is a character that requires 4 bytes in UTF-8 encoding.

Special characters include, but are not limited to, the following:

- Emojis
- Music symbols
- Mathematical symbols

When a special character is written to a FlexGroup volume, the following behavior occurs:

```
# mkdir /flexgroup4TB/🏆
mkdir: cannot create directory '/flexgroup4TB/\360\237\217\206': Permission denied
```

In the preceding example, `\360\237\217\206` is hex `0xF0 0x9F 0x8F 0x86` in UTF-8, which is a trophy symbol.

ONTAP software did not natively support UTF-8 sizes that are greater than three bytes in NFS, as indicated in bug 229629. To handle character sizes that exceed three bytes, ONTAP places the extra bytes into an area in the operating system known as `bagofbits`. These bits are stored until the client requests them. Then the client interprets the character from the raw bits. FlexVol supports `bagofbits`, and FlexGroup volumes added support for `bagofbits` in ONTAP 9.2.

Best Practice 17: Special Character Handling in FlexGroup Volumes
For special character handling with FlexGroup volumes, use ONTAP 9.2 and later.

Also, ONTAP has an event management system message for issues with `bagofbits` handling.

```
Message Name: wafl.bagofbits.name
Severity: ERROR

Corrective Action: Use the "volume file show-inode" command with the file ID and volume name
information to find the file path. Access the parent directory from an NFSv3 client and rename
the entry using Unicode characters.

Description: This message occurs when a read directory request from an NFSv4 client is made to a
Unicode-based directory in which directory entries with no NFS alternate name contain non-Unicode
characters.
```

To test `bagofbits` functionality in FlexGroup, use the following command:


```
# touch "$(echo -e "file\xFC")"
```

In ONTAP 9.1, this would fail:

```
# touch "$(echo -e "file\xFC")"
touch: cannot touch `file\374': Permission denied
```

In ONTAP 9.2 and later, this command succeeds:

```
# touch "$(echo -e "file\xFC")"
# ls -la
-rw-r--r--.  1 root root     0 May  9  2017 file?
```

Support for utf8mb4 Volume Language

As mentioned before, special characters might exceed the supported three bytes UTF-8 encoding that is natively supported. ONTAP then uses the `bagofbits` functionality to allow these characters to work.

This method for storing inode information is not ideal, so starting in ONTAP 9.5, utf8mb4 volume language support was added. When a volume uses this language, special characters that are four bytes in size will be stored properly and not in `bagofbits`.

Volume language is used to convert names sent by NFSv3 clients to Unicode, and to convert on-disk Unicode names to the encoding expected by NFSv3 clients. In legacy situations in which NFS hosts are configured to use non-UTF-8 encodings, you will want to use the corresponding volume language. Use of UTF-8 has become almost universal these days, so the volume language is likely to be UTF-8.

NFSv4 requires use of UTF-8, so there is no need to use non-UTF-8 encoding for NFSv4 hosts. Similarly, CIFS uses Unicode natively, so it will work with any volume language. However, use of utf8mb4 is recommended because files with Unicode names above the basic plane are not converted properly on non-utf8mb4 volumes.

Volume language can only be set on a volume at creation by using the `–language` option. You cannot covert a volume's language. To use files with a new volume language, create the volume and migrate the files by using a utility like the XCP Migration Tool.

Best Practice 18: UTF-8 or utf8mb4?
If you're running ONTAP 9.5 or later, it is best to use the utf8mb4 volume language to help prevent issues with filename translation unless clients are unable to support the language.

Use of Change Notifications with SMB

SMB change notifications are how SMB clients are informed of a file's existence in a share without needing to close a session or refresh a window. The SMB clients are in constant contact with the SMB server during SMB sessions, and the SMB server sends periodic updates to the client regarding any file changes in the share. This feature is used, for example, for applications that must write files and then be able to immediately read the files in SMB shares. This functionality is controlled through the `changenotify` share property. ONTAP 9.4 and later versions automatically set this share property on new SMB shares, even if change notifications are not needed.

Support for SMB change notifications was added for FlexGroup volumes in ONTAP 9.2. However, in rare cases, this share property can create performance issues. This is because a FlexGroup volume spans multiple nodes in a cluster and latency can occur because of the number of change notification requests, particularly in high-file-count environments.

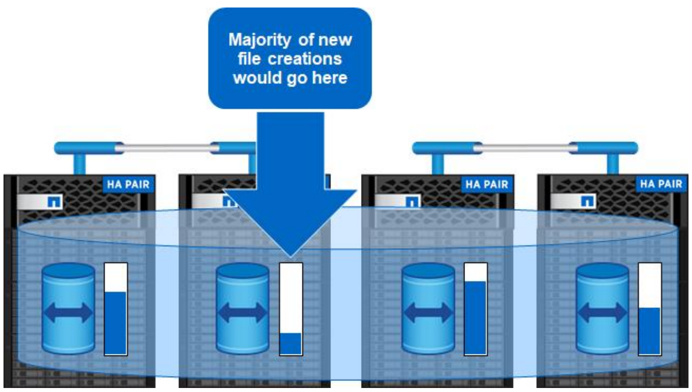
ONTAP 9.5 introduced inherited change notifications. This feature adds an improved algorithm that divides change notifications better and expends less CPU for operations. It should help prevent performance issues when you use SMB change notifications.

Best Practice 19: SMB Change Notification Recommendation
• In ONTAP 9.4 and earlier, disable SMB change notifications unless necessary for applications. You can do this by using System Manager or the `cifs share properties remove` command in the CLI. • If change notifications are not needed, disable them on FlexGroup volumes to avoid potential performance impact. • If using SMB change notifications, use ONTAP 9.5 or later.

File Deletions/FlexGroup Member Volume Balancing

In general, a FlexGroup volume spreads data across multiple member volumes evenly on ingest of data. This makes file deletions operate a bit more efficiently on a FlexGroup volume as compared to a FlexVol volume. This is because the system is able to use more hardware and WAFL affinities to spread out the delete load more efficiently and use less CPU per node for these operations. However, overall performance of file deletions might be slower because of remote access across the FlexGroup volume as compared to FlexVol volumes. In rare cases, the deletion of files (especially sets of large files) can create artificial hot spots in a FlexGroup volume by way of capacity imbalances.

Figure 53) Capacity imbalance after deletion of larger files.



A FlexGroup volume's workload balance can be viewed with the following **diag-privilege**-level command:

```
cluster::*> set diag
cluster::*> node run * flexgroup show [flexgroup name]
```

This displays the following output:

- Member volume dataset ID (DSID)
- Member volume capacities (used and available, in blocks)
- Member volume used %
- Urgency, target, and probability percentages (used in ingest calculations)

For more information, see "flexgroup show" later in this document.

Rebalancing Data Within a FlexGroup

It is not possible to rebalance the workload in a FlexGroup volume to even out capacities, but ONTAP generally does a good job of balancing the ingest load, so a rebalance is not necessary. In the rare case in which a member volume becomes a hot spot, you should analyze the workload. You can use the XCP Migration Tool to scan folders and files to identify file sizes and anomalies. For an example, see the section "Using XCP to Scan Files Before Migration."

After the files are identified, either delete them, move them, add space to the member volumes, or add more member volumes to help balance the ingest load in a FlexGroup volume. These options provide more affinities until the less-full members catch up to other member volumes or until you increase the size of the existing FlexGroup member volumes to provide some relief. In most cases, imbalances in data capacity won't be noticeable to the end user.

Why Doesn't a FlexGroup Volume Rebalance Data?

As a FlexGroup volume ingests data, it has three goals:

- The volume should encourage all its member FlexVol volumes to participate in hosting the workload in parallel. If only a subset of member volumes is active, the FlexGroup volume should distribute more new data toward the underactive members.
- The FlexGroup volume should prevent any member FlexVol volume from running out of free space, unless all other members are also out of free space. When one member has more data than others, the FlexGroup volume should align the underused members by placing new data on them at a higher-than-average rate.
- The FlexGroup volume must minimize the performance losses caused by pursuing the previous two goals. If the FlexGroup volume were to carefully and accurately place each new file where it could be most beneficial, then the previous two goals could be easily achieved. However, the cost of all that careful placement would appear as increased service latency.

Some of these goals are in conflict, so ONTAP employs a sophisticated set of algorithms and heuristics to maintain a balance in the FlexGroup volumes. However, in some scenarios, imbalances such as the following might occur:

- Large files or files that grow over time might be present in a FlexVol member volume.
- Many files might be zipped or tarred into a single file in the same FlexGroup volume as the files themselves.
- A large amount of data might be deleted, and most of that data could be from the same member volume (rare).

In scenarios where FlexGroup member volumes have an imbalance of capacity or files, ONTAP takes extra measures to help the less-allocated member volumes "catch up" to the filled members. As a result, performance can be affected.

However, this performance effect isn't as serious as the performance effect of ONTAP moving data in a FlexGroup volume to rebalance the workloads. Therefore, the current approach to FlexGroup volume imbalances is to enable volume autogrow and set thresholds (~80% full) to help keep the system's free space in check.

Performance Issues When Member Volumes Reach 80%

In versions earlier than ONTAP 9.5P4, the ingest heuristics could encounter an issue that might contribute to performance issues when a member volume capacity reaches 80%. If possible, when using FlexGroup volumes, upgrade to ONTAP 9.5P4 or later. For more information about the issue, check bug 1231125.

Listing Files When a Member Volume Is Out of Space

If a FlexGroup member volume runs out of space, the entire FlexGroup volume reports that it is out of space. Even read operations, such as listing the contents of a folder, can fail when a FlexGroup member is out of space.

Although `ls` is a read-only operation, FlexGroup volumes still require a small amount of writable space to allow it to work properly. ONTAP uses that storage to establish metadata caches. For example, suppose the name `foo` points to an inode with X properties, and the name `bar` points to an inode with Y properties. The amount of space used is negligible—a few kilobytes, or maybe a few megabytes on large systems—and this space is used and released repeatedly. Internally, this space is called the RAL reserve.

Under normal circumstances, even if you manage to fill up a member volume, a bit of space is left for the FlexGroup volume to use as it performs read-only operations like `ls`. However, ONTAP prioritizes other operations over the RAL reserve. If a member volume is 100% full, for example, and you create a Snapshot copy and then try to continue using the volume, the WAFL Snapshot reserve is used as you overwrite blocks and therefore consumes more space. ONTAP prioritizes the Snapshot space and takes space from things like the RAL reserve. This scenario rarely occurs, but it explains why an operation like `ls` might fail because of lack of space.

File Rename Considerations

FlexGroup volumes handle most high-metadata workloads well. However, with workloads that do a large amount of file renames at a time (for example, hundreds of thousands), performance of these operations suffers in comparison to FlexVol volumes. This is because a file rename doesn't move the file in the file system; instead, it just moves the name to a new location. In a FlexGroup volume, moving this name would likely take place as a remote operation and create a remote hard link. Subsequent renames would create more remote hard links to the file's location, which would keep adding latency to operations that occur on that file. If an application's workflow is mostly file renames, you should consider using FlexVol volumes instead of FlexGroup volumes for the rename operations. If the desired final landing spot is a FlexGroup volume after the rename occurs, consider moving the files from the FlexVol volume to the FlexGroup volume after the rename process.

Symlink Considerations

If your workload contains many symlinks (that is, symlink counts in the millions) in a single FlexGroup volume, attempts to resolve that many symlinks might have a negative effect on performance. The negative effect is caused by creating remote hard links artificially in addition to the remote hard links ONTAP creates.

Try to keep the number of symlinks below a few thousand per FlexGroup if possible.

NFS Version Considerations

When a client using NFS attempts to mount a volume in ONTAP without specifying the NFS version (for example, `-o nfsvers=3`), a protocol version negotiation between the client and server takes place. The client asks for the highest versions of NFS supported by the server. If the server (in the case of ONTAP, an SVM serving NFS) has NFSv4.x enabled, the client attempts to mount with that version.

In ONTAP 9.6 and earlier, FlexGroup volumes did not support NFSv4.x, so the mount request fails in those releases. Usually, this error manifests as “access denied,” which can mask what the actual issue is in the environment:

```
# mount demo:/flexgroup /flexgroup
mount.nfs: access denied by server while mounting demo:/flexgroup
```

In ONTAP 9.7 and later, NFSv4.x is supported. This can create a different set of issues, however. Clients will still mount the latest NFS version advertised by the NFS server (in this case, the ONTAP SVM). If NFSv4.x versions are enabled, clients might mount through NFSv4.x when NFSv3 is desired or expected. When NFSv4.x mounts, performance and access permissions behave differently than in NFSv3.

Mapping NFS Connected Clients to Volume Names

To check what version of NFS is being mounted from the cluster, use the `nfs connected-clients show` command available in ONTAP 9.7:


```
cluster::> nfs connected-clients show -node * -vserver DEMO

     Node: node1
  Vserver: DEMO
  Data-Ip: 10.x.x.x
Client-Ip       Volume-Name      Protocol Idle-Time     Local-Reqs Remote-Reqs
--------------- ---------------- -------- ------------- ---------- -----------
10.x.x.x        CIFS             nfs4.1   2d 0h 9m 3s   153        0
10.x.x.x        vsroot           nfs4.1   2d 0h 9m 3s   0          72
10.x.x.x        flexgroup_16__0001
                                 nfs3     0s            0          212087
10.x.x.x        flexgroup_16__0002
                                 nfs3     0s            0          192339
10.x.x.x        flexgroup_16__0003
                                 nfs3     0s            0          212491
10.x.x.x        flexgroup_16__0004
                                 nfs3     0s            0          192345
10.x.x.x        flexgroup_16__0005
                                 nfs3     0s            212289     0
```

To avoid issues with mounting a FlexGroup volume in environments in which NFSv4.x is enabled, either configure clients to use a default mount version of NFSv3 through `fstab` or explicitly specify the NFS version when mounting.

For example:

```
# mount -o nfsvers=3 demo:/flexgroup /flexgroup
# mount | grep flexgroup
demo:/flexgroup on /flexgroup type nfs (rw,nfsvers=3,addr=10.193.67.237)
```

Also, if a FlexGroup volume is junctioned to a parent volume that is mounted to a client with NFSv4.x, traversing to the FlexGroup volume in ONTAP 9.6 and earlier fails, because no NFSv4.x operations are allowed to FlexGroup volumes.

For example, FlexGroup volumes are always mounted to the vsroot (vserver root), which operates as (/) in the NFS export path. If a client mounts vsroot with NFSv4.x, attempts to access a FlexGroup volume in ONTAP 9.6 or earlier from the NFSv4.x mount fail. This includes `ls -la` operations, because they require the ability to do NFSv4.x `GETATTR` operations.

Note in the following example that the information for the FlexGroup volumes is incorrect because of the lack of NFSv4.x support:

```
# mount demo:/ /mnt
# mount | grep mnt
demo:/ on /mnt type nfs (rw,vers=4,addr=10.193.67.237,clientaddr=10.193.67.211)
# cd /mnt/flexgroup
-bash: cd: /mnt/flexgroup: Permission denied
# ls -la
ls: cannot access flexgroup_4: Permission denied
ls: cannot access flexgroup_local: Permission denied
ls: cannot access flexgroup_8: Permission denied
ls: cannot access flexgroup_16: Permission denied
drwx--x--x. 12 root root   4096 Mar 30 21:47 .
dr-xr-xr-x. 36 root root   4096 Apr  7 10:30 ..
d??????????  ? ?    ?         ?            ? flexgroup_16
d??????????  ? ?    ?         ?            ? flexgroup_4
d??????????  ? ?    ?         ?            ? flexgroup_8
```

Compare that to the NFSv3 mount:

```
# ls -la
drwx--x--x. 12 root root 4096 Mar 30 21:47 .
dr-xr-xr-x. 36 root root 4096 Apr  7 10:30 ..
drwxr-xr-x.  6 root root 4096 May  9 15:56 flexgroup_16
drwxr-xr-x.  5 root root 4096 Mar 30 21:42 flexgroup_4
drwxr-xr-x.  6 root root 4096 May  8 12:11 flexgroup_8
```


As a result, be sure to avoid using NFSv4.x in any path where a FlexGroup volume resides in ONTAP 9.6 or earlier. If NFSv4.x is desired, upgrade ONTAP to 9.7 or later.

Enabling and Using NFSv4.x with FlexGroup Volumes

FlexGroup volumes function identically to FlexVol volumes when you configure NFSv4.x in your environment. TR-4067: NFS Best Practice and Implementation Guide covers NFS in detail for use with ONTAP and applies to FlexGroup volumes as well. Rather than focusing on performance, the benefits of using NFSv4.x with workloads include:

- **Security.** NFSv4.x greatly improves security with NFS through integration of ancillary protocols (such as NLM, NSM, mountd, portmapper) into a single port over 2049. Fewer firewall ports being open helps reduce the threat vectors available.

  Additionally, NFSv4.x includes Kerberos encryption (krb5, krb5i, and krb5p) as part of its RFC requirements, meaning that a client/server will not be compliant with the RFC unless it includes Kerberos support.

  NFSv4.x also provides better masking of UID/GID information by requiring the client and server matching domain IDs in their configurations, which helps make spoofing users harder—particularly when using Kerberos encryption.

  Finally, NFSv4.x offers granular ACL support that mimics the functionality of Windows NTFS ACLs. This provides the ability to add more users and groups to an ACL than NFSv3 offered with mode bits, as well as allowing more ACL functionality beyond basic read/write/execute (rwx). NFSv4.x ACLs can even be applied to datasets that will mount only NFSv3, which can offer granular security on files and folders even if NFSv4.x isn't being used.
- **Improved locking.** NFSv3 locking was done outside the NFS protocol, using ancillary protocols like NSM and NLM. This often resulted in stale locks when clients or servers had outages, which prevented access to files until those stale locks were cleared.

  NFSv4.x provides locking enhancements by way of a leasing mechanism that holds a lease for a specified time and keeps that lease if the client/server communication is intact. If there are any issues with that communication (whether network or server outage), the lease will expire and release the lock until it is reestablished.

  Additionally, locking in NFSv4.x is integrated within the NFS packets, providing more reliable and efficient locking concepts than NFSv3.
- **Data locality and parallel access.** NFSv4.x offers data locality functionality for scale-out NAS environments, such as NFSv4.x referrals, which can redirect mount requests to volumes in ONTAP according to the location on a node to ensure local access to the mount.

  NFSv4.1 also offers parallel NFS support, which establishes a metadata server on mount and then redirects data I/O across the namespace. To do this, it uses a client/server communication that keeps track of data according to node and data LIF location. This concept is similar to that of asymmetric logical unit access (ALUA) for SAN. For more information about pNFS in ONTAP, see TR-4063: Parallel Network File System Configuration and Best Practices for Clustered Data ONTAP 8.2 and Later.

  For more information, see the section on pNFS with FlexGroup volumes.

NFSv4.x Performance Enhancements in ONTAP

In general, NFSv4.x is less performant than NFSv3 because NFSv4.x is stateful, so it has more to do for each protocol operation. NFSv4.x overhead comes in the form of locking and leasing, ACLs, compound calls, and communication of state IDs between the client and server, as well as the processing of each packet.

ONTAP 9.2 and later brought NFSv4.x performance somewhat closer to NFSv3 performance for streaming I/O workloads, such as SAP HANA and databases. ONTAP 9.6 and later added some metadata workload performance enhancements for NFSv4.x that improved results on EDA benchmarks to be more in line with NFSv3.

One of the weak points for performance with NFSv4.x includes workloads with high metadata ingest. FlexGroup volumes work best with these types of workloads, so if you're considering NFSv4.x for these workloads, NetApp strongly recommends using FlexGroup volumes.

One of the benefits of using NFSv4.x is that it does not use RPC slot tables in its operations, so it is not susceptible to RPC slot exhaustion.

If you're using Kerberos with NFS, there is also a small performance impact to operations for processing overhead of the encrypted packets. The impact varies depending on several factors, including:

- ONTAP version
- Hardware being used
- Network latency/WAN latency/cloud region
- Performance headroom on the cluster
- Kerberos encryption being used (krb5, krb5i, or krb5p)
- ONTAP 9.2 and later versions offer AES-NI offloading of NFS Kerberos packets, and ONTAP 9.2 is the minimum version if you're considering Kerberized NFS. For more information about configuring and managing Kerberos with NFS, see TR-4073: Secure Unified Authentication and TR-4616: NFS Kerberos in ONTAP.

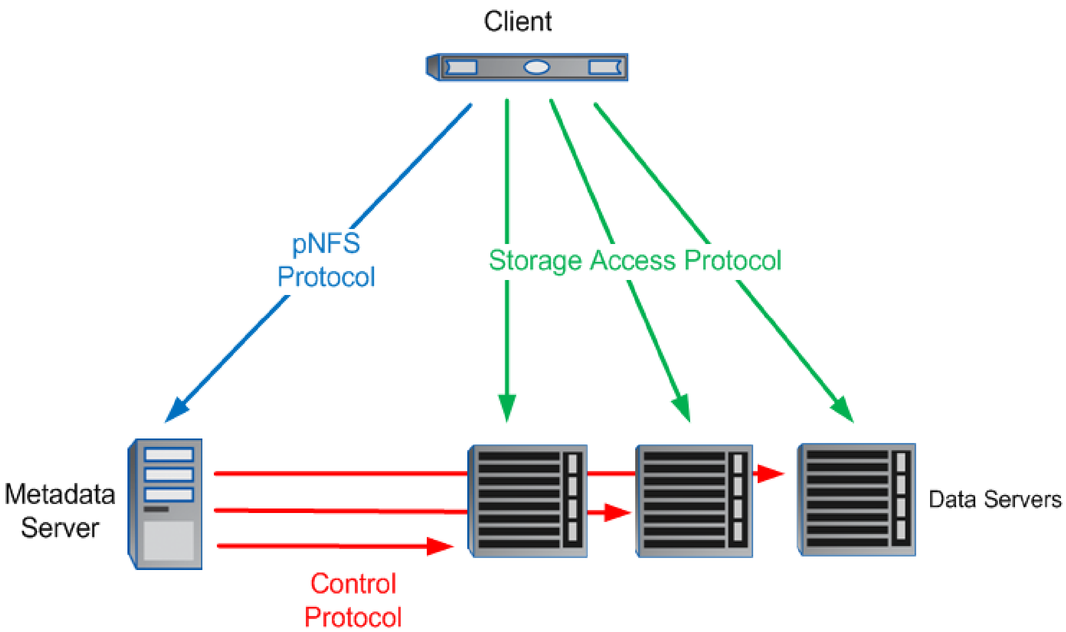
pNFS with FlexGroup Volumes

ONTAP 9.7 offers support for NFSv4.x, which includes NFSv4.1 and its RFC mandatory features. Included in those features is parallel NFS (pNFS), which provides localization of reads and writes across multiple volumes and nodes in a cluster. ONTAP provides the file version of pNFS and does not use the striping or block versions of the feature.

How pNFS Works in ONTAP

If pNFS has been enabled on the NFS server in an SVM, clients that support pNFS and mount by using NFSv4.1 will first connect to a specific node in the cluster with a single TCP connection that acts as a metadata server. This connection will service pNFS operations, such as client/server communications for data layout, LIF location, and pNFS mappings to help redirect I/O traffic to the local volumes and data LIFs in the cluster. The metadata server also services NFS metadata operations such as `getattr` operations and `setattr` operations.

Figure 54) pNFS diagram.



The pNFS architecture includes three main components:

The metadata server that handles all nondata I/O traffic. It is responsible for all metadata operations, such as `GETATTR`, `SETATTR`, `LOOKUP`, `ACCESS`, `REMOVE`, and `RENAME` operations. The metadata server also provides information about the layout of files.

- **Data servers that store file data and respond directly to client read and write requests.** Data servers handle pure `READ` and `WRITE` I/O.
- **One or more clients that are able to access data servers directly.** This access is based on metadata received from the metadata server.

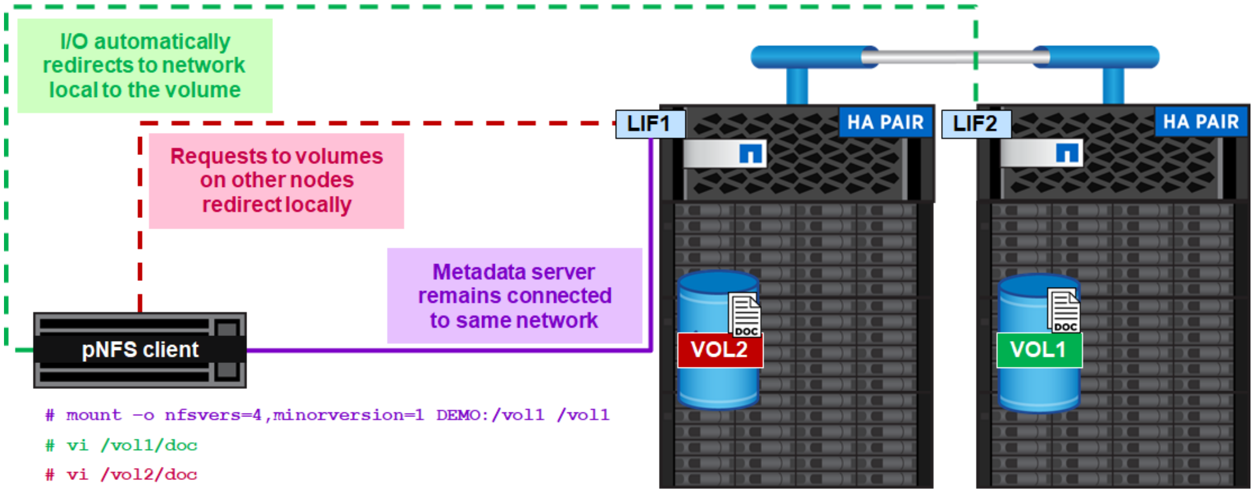
There are three types of protocols used between the clients, metadata server, and data servers:

- **A control protocol used between the metadata server and data servers.** This protocol synchronizes file system data.
- **The pNFS protocol, used between clients and the metadata server.** This is essentially the NFSv4.1 protocol with a few pNFS-specific extensions. It is used to retrieve and manipulate layouts that contain the metadata that describes the location and storage access protocol required to access files stored on numerous data servers.
- **A set of storage access protocols used by clients to access data servers directly.** The pNFS specification currently has three categories of storage protocols: file based, block based, and object based. Data ONTAP 8.1 and later support file-based storage protocol and access the data servers over NFSv4.1.

When a read or write request is performed by a client over pNFS, the client and server negotiate where to send those requests by using the data layout mappings. For example, if a file lives on volume1 (which lives on node1) in a cluster, but the metadata server is connected to node2, then the data layout mapping informs the client to perform the reads/writes over a network connection local to node1.

If a volume is moved (for example, with a nondisruptive volume move operation), the data layout table is updated and ONTAP redirects local traffic to the volume on the next request. This process is similar to how ALUA works in SAN environments, where a path can switch based on locality of the block device.

Figure 55) pNFS operations diagram.



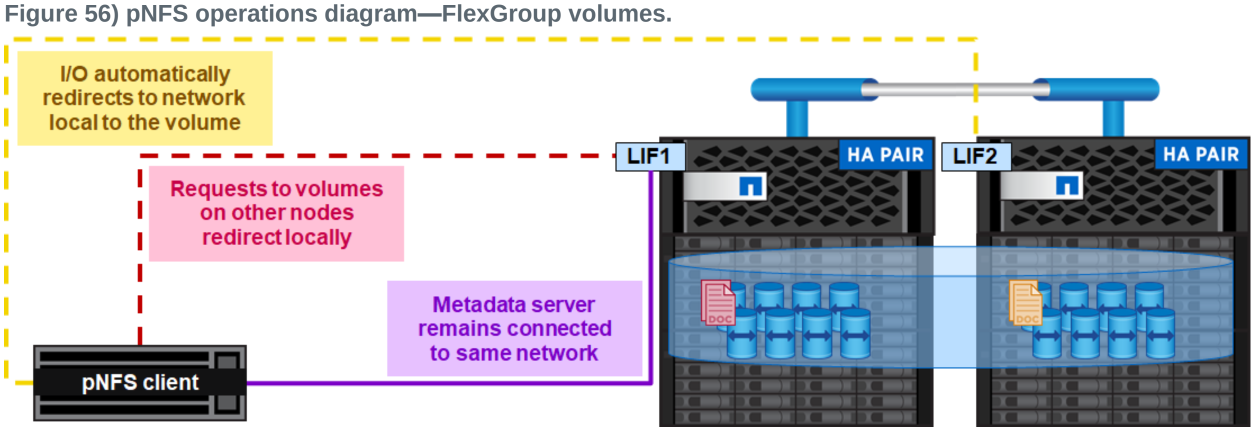
How pNFS Works with FlexGroup Volumes

A FlexGroup volume operates as a single entity, but is constructed of multiple FlexVol member volumes. Each member volume contains unique files that are not striped across volumes. When NFS operations connect to FlexGroup volumes, ONTAP handles the redirection of operations over a cluster network.

With pNFS, these remote operations are reduced, because the data layout mappings track the member volume locations and local network interfaces; they also redirect reads/writes to the local member volume inside a FlexGroup volume, even though the client only sees a single namespace. This approach enables a scale-out NFS solution that is more seamless and easier to manage, and it also reduces cluster network traffic and balances data network traffic more evenly across nodes.

FlexGroup pNFS differs a bit from FlexVol pNFS. Even though FlexGroup load-balances between metadata servers for file opens, pNFS uses a different algorithm. pNFS tries to direct traffic to the node on which the target file is located. If multiple data LIFs per node are given, connections can be made to each of the LIFs, but only one of the LIFs of the set is used to direct traffic to volumes per network interface.

Figure 56) pNFS operations diagram—FlexGroup volumes.



pNFS Best Practices

pNFS best practices in ONTAP don’t differ much from normal NAS best practices, but here are a few to keep in mind. In general:

- Use the latest supported client OS version.

- Use the latest supported ONTAP patch release.
- Create a data LIF per node, per SVM to ensure data locality for all nodes.
- Avoid using LIF migration on the metadata server data LIF, because NFSv4.1 is a stateful protocol and LIF migrations can cause brief outages as the NFS states are reestablished.
- In environments with multiple NFSv4.1 clients mounting, balance the metadata server connections across multiple nodes to avoid piling up metadata operations on a single node/network interface.
- If possible, avoid using multiple data LIFs on the same node in an SVM.
- In general, avoid mounting NFSv3 and NFSv4.x on the same datasets. If you can't avoid this, check with the application vendor to ensure that locking can be managed properly.
- If you're using NFS referrals with pNFS, keep in mind that referrals will establish a local metadata server, but data I/O will still redirect. With FlexGroup volumes, the member volumes might live on multiple nodes, so NFS referrals aren't of much use. Instead, use DNS load balancing to spread out connections.

For more detailed information, see TR-4063: Parallel Network File System Configuration and Best Practices for Clustered Data ONTAP 8.2 and Later.

NFSv4.x General Considerations

When considering NFSv4.x for your SVM, be sure to factor in performance, client/application support, name services infrastructure, and locking mechanisms before deploying. Also consider whether applications can use both NFSv3 and NFSv4.x on the same datasets. For instance, VMware recommends against service datastores over both protocol versions.

If possible, set up a separate SVM to conduct functionality and performance testing before deploying in production.

NFSv4.x configuration generally requires the following to work properly:

- NFS clients that support NFSv4.x.
- NFS mounts that specify NFSv4.x.
- NFS server configuration (NFSv4.x and desired features enabled—such as referrals, pNFS, ACL support, NFSv4 ID domain configured to be identical on client and NFS server).
- Matching user names and groups on client and server (case sensitive; for example, user1@domain.com should exist on both server and client; USER1 and user1 are not considered matches).
- Optional: Name services for UNIX identities, such as NIS or LDAP, can greatly simplify NFSv4.x implementation and functionality.

For more detailed information, see TR-4067: NFS Best Practice and Implementation Guide.

NFS Mount Considerations: READDIRPLUS (READDIR+)

If you are running a version of ONTAP earlier than 9.1P4 and use the `READDIR+` functionality in NFS, you might experience some latency on rename operations. This is caused by bug 1061496, which is fixed in 9.1P4 and later. If you're running a release of ONTAP that is exposed to this bug and are experiencing latencies, consider mounting FlexGroup volumes with the option `-nordirplus` to disable `READDIR+` functionality.

NAS Metadata Effect in a FlexGroup Volume

The overhead for metadata operations affects how a workload performs, which can be anywhere from a 10% to 30% performance hit for remote operations. Most of the metadata effect is related to write metadata. Most read metadata has little to no effect.

- **`getattr`**, **`access`**, **`statfs`**, **`lock`**, **`unlock`**. Little to no FlexGroup overhead.
- **`Readdirplus`**. Before ONTAP 9.3, mostly remote; after 9.3, little remote overhead.
- **`nfs create`**, **`unlink`**, **`lookup`**. Little to no FlexGroup overhead under heavy load.
- **`nfs mkdir`**, **`rmdir`**, **`lookup dir`**. 50% to 100% remote access, so high overhead.
- **CIFS `open`/`close`**. High overhead.

Virtualization Workload Considerations

Although NetApp does not currently recommend using virtualization workloads on a FlexGroup volume, there is no technical reason why it is not possible. ONTAP does not prevent using virtualization workloads on a FlexGroup volume, but it's important to consider the following:

- Virtual machines and Snapshot copies start out as small files and grow over time; this can affect the balance and load distribution in a FlexGroup volume.
- FlexGroup volumes support VAAI starting in ONTAP 9.7, which is used to offload copy operations from vSphere to storage. Note that copy offload is not always faster than host copy, and vSphere only offloads operations on cold VMs for NFS storage.
- With virtualization workloads, VMDK files in the same FlexGroup datastore could live in multiple FlexVol member volumes across the cluster. As a result, creating Snapshot copies of these VMDKs could have unpredictable results because of how FlexGroup volumes coordinate Snapshot copies across member volumes and pause I/O during Snapshot creation.
- FlexGroup volumes are not currently supported by NetApp tools for vSphere such as the Virtual Storage Console (VSC), VASA Provider (for vVols), Storage Replication Adapter (for Site Recovery Manager), SnapCenter for vSphere, and so on.
- It is possible to use FlexGroup volumes with virtualization backup products such as Veeam or Rubrik, because these products use file protocols to perform backups. However, using FlexGroup volumes in disaster recovery scenarios to host datastores in failure events will present the same challenges as using FlexGroup volumes for virtualization workloads in production.
- Because FlexGroup volumes have not been fully tested for virtualization workloads, it's possible that other limitations might apply. Before deploying in production, be sure to fully test the virtualization workload on the FlexGroup volume.

Databases on FlexGroup Volumes

Usually, databases (such as Oracle) create a few small files when they are deployed. In a FlexGroup volume, small numbers of small files tend to favor local placement to their parent folder. This means that an Oracle deployment of eight databases might all land inside the same FlexGroup member volume. Not only does this provide no benefits from load distribution across nodes in a cluster, it can also present a problem as the files grow over time. Eventually, the files start to fill the member volume to capacity, and there is a need for remediation steps to move around data.

Database workloads, in theory, would work well in a single namespace that can span a cluster. However, because the files are likely to grow over time and latency-sensitive databases might run on volumes that traverse the cluster network, NetApp currently recommends placing database files in FlexVol volumes.

At-Rest Encryption Considerations

ONTAP 9.2 introduced support for NetApp Volume Encryption (NVE) for FlexGroup volumes. Implementing this feature with FlexGroup volumes follows the same recommendations and best practices as stated for FlexVol volumes. Re-keying an existing FlexGroup volume is possible in ONTAP 9.5 and later. See "Rekeying a FlexGroup Volume or Encrypting Existing FlexGroup Volumes" for details.

Generally speaking, NVE requires the following:

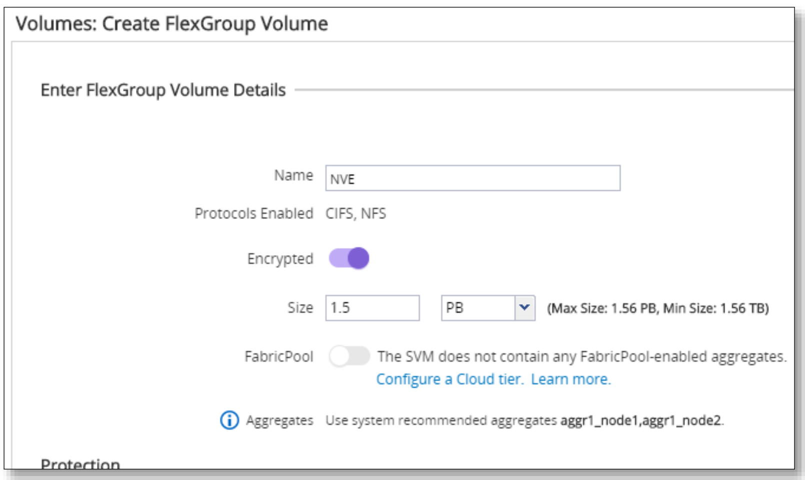
- A valid NVE license
- A key management server (on-box or off-box as of ONTAP 9.3)
- A cluster-wide passphrase (32 to 256 characters)
- AFF or FAS hardware that supports AES-NI offloading

For information about implementing and managing NVE with FlexGroup and FlexVol volumes, see the NetApp Encryption Power Guide and the Scalability and Performance Using FlexGroup Volumes Power Guide on the support site for your release of ONTAP.

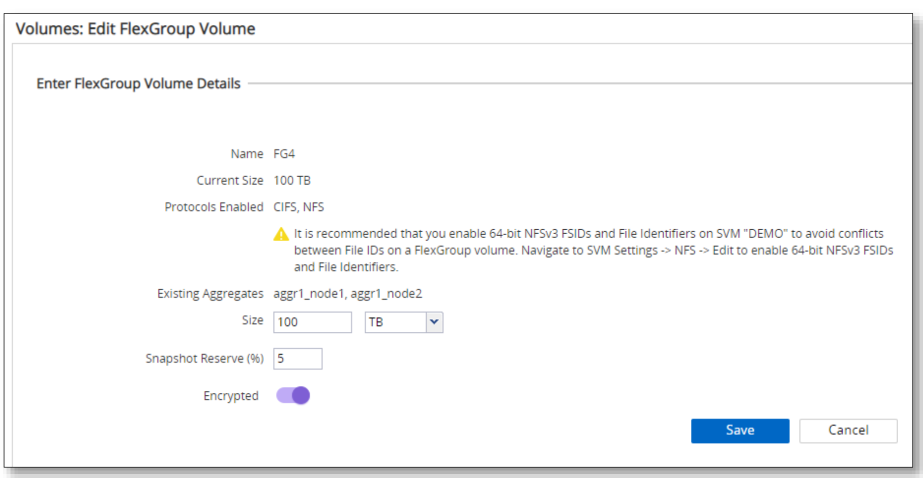
ONTAP 9.6 added NetApp Aggregate Encryption (NAE), which allows you to encrypt at the aggregate level. FlexGroup volumes can use NAE, provided all aggregates that contain member volumes belonging to the same FlexGroup volume are encrypted.

Encrypting Your FlexGroup Volume

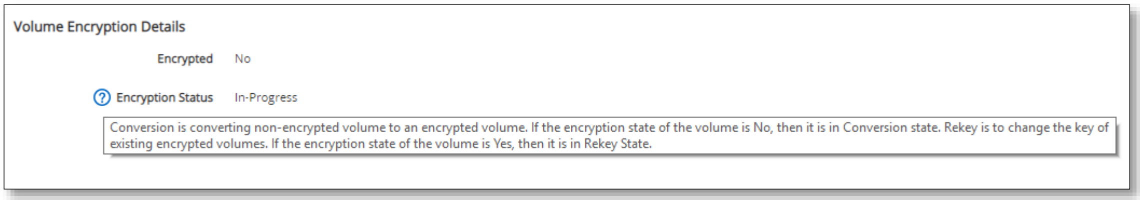
The simplest way to encrypt a FlexGroup volume is by using System Manager. For new volumes, select Encrypted on the creation screen.



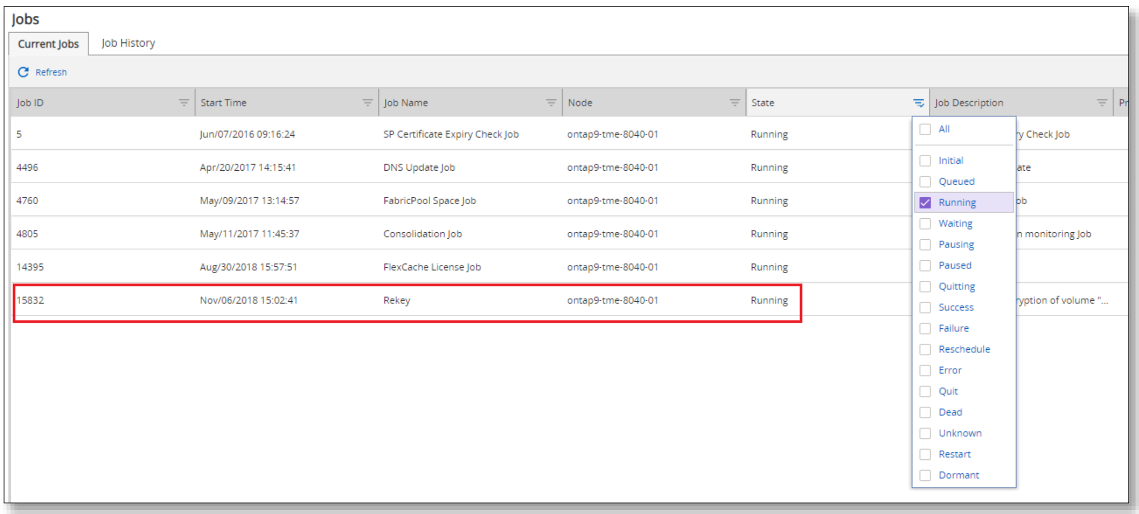
For existing FlexGroup volumes, select the volume and click Edit. Then toggle the Encrypted value to purple.



After this action, the FlexGroup volume takes a while to encrypt. In this example, the 100TB FlexGroup volume took 30 to 45 minutes.



You can view the progress of this process in System Manager in Events and Jobs > Jobs. Use Running as a filter. There, you can see a Rekey job.



For more detailed information, use the CLI and the `job show -instance` command.

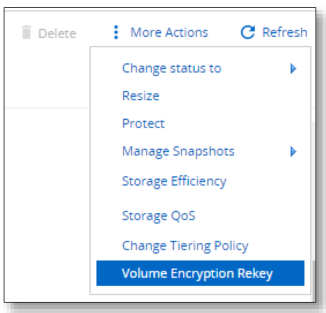
```
cluster::*> job show -id 15832 -instance

                      Job ID: 15832
              Owning Vserver: cluster
                        Name: Rekey
                 Description: Conversion to encryption of volume "FG4".
                    Priority: High
                        Node: node1
                    Affinity: Cluster
                    Schedule: @now
                  Queue Time: 11/06 15:02:41
                  Start Time: 11/06 15:02:41
                    End Time: -
              Drop-dead Time: -
                  Restarted?: false
                       State: Running
                 Status Code: 0
           Completion String:
                    Job Type: VOL_REKEY
                Job Category: VOPL
                        UUID: eb19d1dc-e1fe-11e8-88fc-00a0986b1223
          Execution Progress: -
                   User Name: admin
                     Process: mgwd
  Restart Is or Was Delayed?: false
Restart Is Delayed by Module: -
```

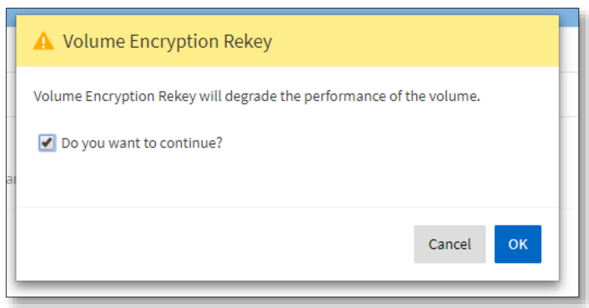

When this command finishes, the volume appears as encrypted.

Rekeying a FlexGroup Volume or Encrypting Existing FlexGroup Volumes

ONTAP 9.5 added support for both rekeying FlexGroup volumes and encrypting FlexGroup volumes that have not yet been encrypted. The process is essentially the same as for a FlexVol volume.

In ONTAP System Manager, click the More Actions menu and select Volume Encryption Rekey.





FlexCache Volume Considerations

ONTAP 9.5 also included NetApp FlexCache. This feature provides a sparse volume that can accelerate performance for NAS workloads and prevent volume hot spots in a cluster or across a WAN. The FlexCache cache volume is powered by FlexGroup volumes, and the underlying protocol that redirects the pointers and blocks is the remote access layer (RAL). The RAL is also what makes a FlexGroup volume a FlexGroup volume. ONTAP 9.6 increased the limit of maximum cache volumes per node to 100, so FlexCache has more scalability in current releases.

Starting in ONTAP 9.7, FlexGroup volumes can be origin volumes for FlexCache. For more information about FlexCache, see TR-4743: FlexCache in NetApp ONTAP.

FlexClone

Starting in ONTAP 9.7, NetApp FlexClone is supported for use with FlexGroup volumes. This feature provides storage administrators with a way to create instant, space-efficient copies (backed by Snapshot technology) of volumes to use for testing, development, backup verification, and a variety of other use cases. There are no specific considerations for use with FlexGroup volumes, except that a FlexClone copy of a FlexGroup volume will use the same number of member volumes as the FlexGroup parent volume. As a result, the volume count on a node can start to add up as FlexClone copies are created.

For example, if you have a FlexGroup volume that contains 16 member volumes and then create a FlexClone copy of that FlexGroup volume, you now have used 32 volumes in the system.

```
cluster::*> volume clone create -vserver DEMO -flexclone FGclone -type RW -parent-vserver DEMO -
parent-volume flexgroup_16

cluster::*> vol show -vserver DEMO -volume flexgroup_16*,FGclone* -fields name -sort-by name
vserver volume               name-ordinal
```



```
------- ------------------ ------------
DEMO    FGclone            -
DEMO    flexgroup_16       -
DEMO    FGclone__0001      base
DEMO    FGclone__0002      base
DEMO    FGclone__0003      base
DEMO    FGclone__0004      base
DEMO    FGclone__0005      base
DEMO    FGclone__0006      base
DEMO    FGclone__0007      base
DEMO    FGclone__0008      base
DEMO    FGclone__0009      base
DEMO    FGclone__0010      base
DEMO    FGclone__0011      base
DEMO    FGclone__0012      base
DEMO    FGclone__0013      base
DEMO    FGclone__0014      base
DEMO    FGclone__0015      base
DEMO    FGclone__0016      base
DEMO    flexgroup_16__0001 base
DEMO    flexgroup_16__0002 base
DEMO    flexgroup_16__0003 base
DEMO    flexgroup_16__0004 base
DEMO    flexgroup_16__0005 base
DEMO    flexgroup_16__0006 base
DEMO    flexgroup_16__0007 base
DEMO    flexgroup_16__0008 base
DEMO    flexgroup_16__0009 base
DEMO    flexgroup_16__0010 base
DEMO    flexgroup_16__0011 base
DEMO    flexgroup_16__0012 base
DEMO    flexgroup_16__0013 base
DEMO    flexgroup_16__0014 base
DEMO    flexgroup_16__0015 base
DEMO    flexgroup_16__0016 base
```

8.12 Other Hardware Considerations

For the most consistent level of performance, use NetApp Flash Cache™ cards or NetApp Flash Pool™ aggregates in a cluster on any node that participates in a FlexGroup volume. Flash Cache cards are expected to provide the same performance benefits for FlexGroup volumes that they provide for FlexVol volumes.

Advanced Disk Partitioning

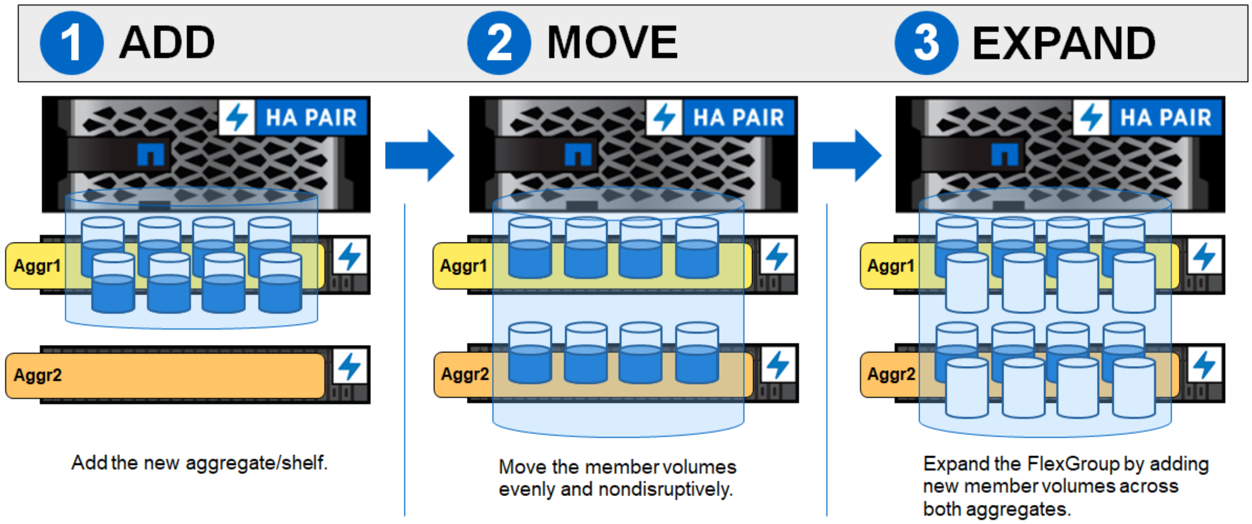
FlexGroup volumes have no bearing on the use of ADP. No special considerations need to be made.

Adding Disks/Aggregates/Nodes

When adding disks to an existing aggregate that contains FlexGroup member volumes, no action is required.

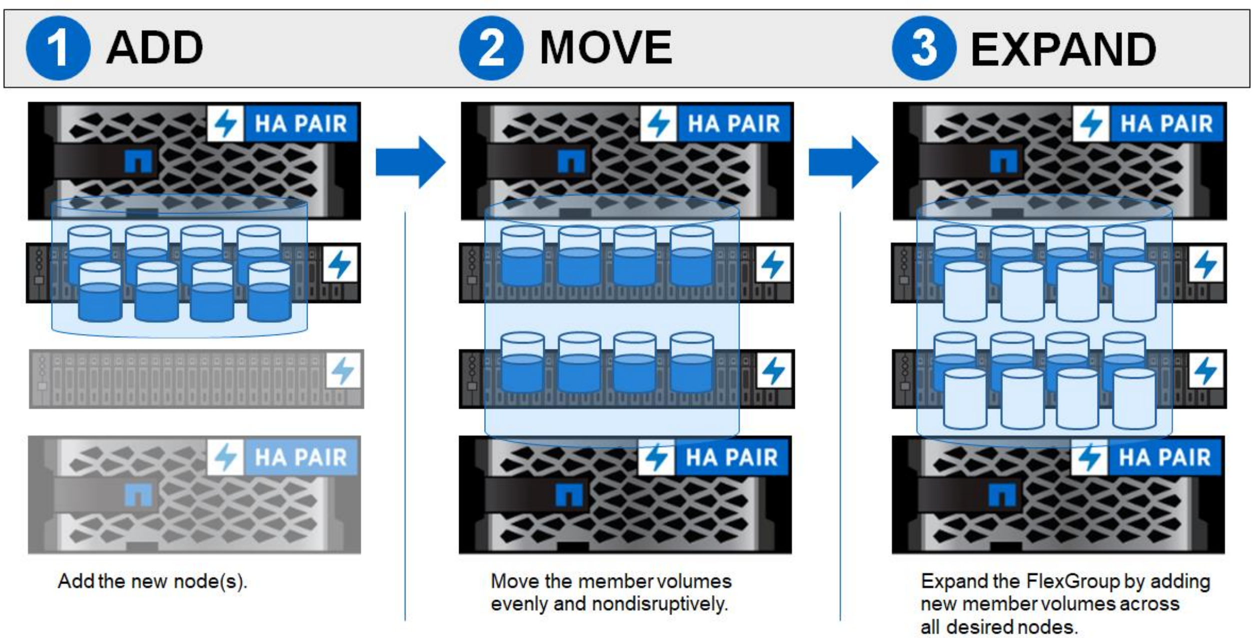
When adding aggregates to nodes, if the FlexGroup volume must span the new aggregates, you can use nondisruptive volume moves to move member volumes to the new aggregates without needing a maintenance window. Then you would create member volumes in the FlexGroup volume spanning new and old aggregates.

Figure 57) Adding aggregates with FlexGroup volumes.



When adding new nodes to a cluster, follow the same steps for adding aggregates to a cluster. Use `volume move` and `volume expand` commands to adjust the member volumes.

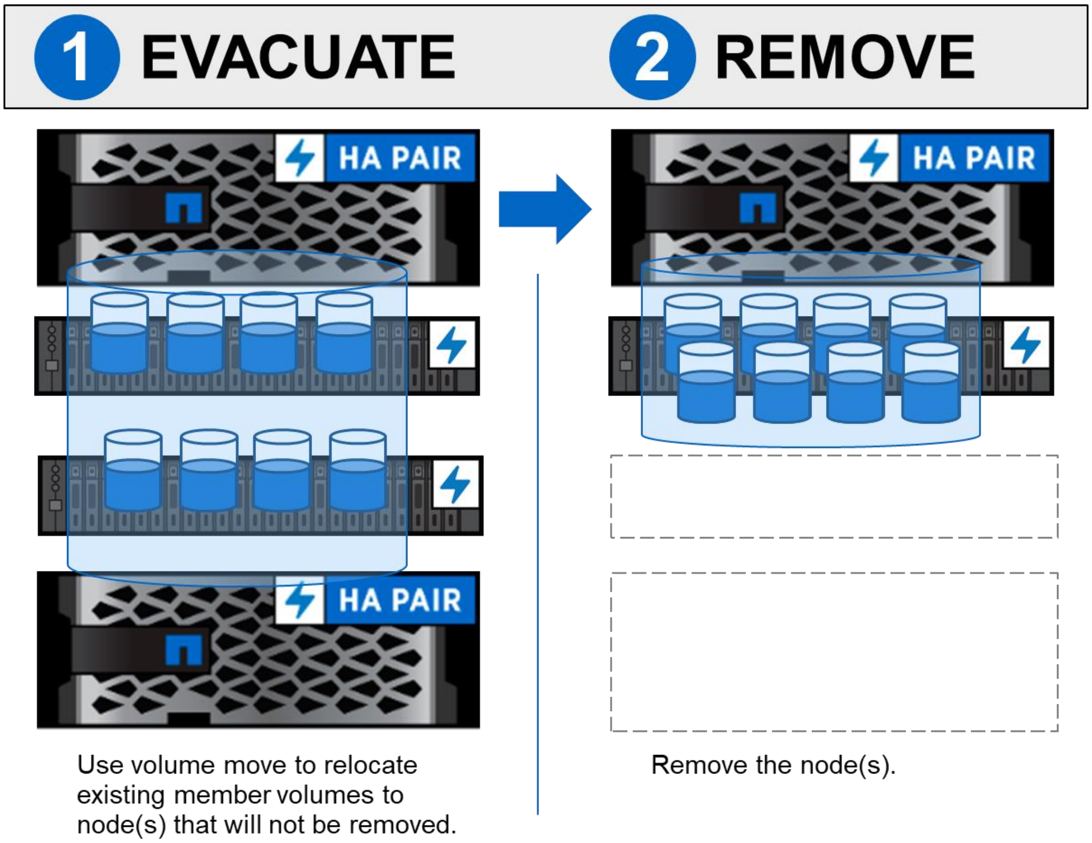
Figure 58) Adding nodes and expanding the FlexGroup volume.



Removing Nodes

You can also remove nodes from a cluster by using nondisruptive volume moves of member volumes. For example, if you want to remove two nodes from an eight-node cluster and each node has 16 member volumes, then you would use volume move to distribute 32 member volumes across the remaining six nodes. Because 32 volumes do not evenly divide across six nodes, use the next divisible node count to evenly distribute so that four nodes get eight member volumes per node. If there isn’t enough space for four nodes to take on eight member volumes each, place six member volumes on two nodes (12) and five member volumes on four nodes (20).

Figure 59) Removing nodes that contain FlexGroup member volumes.



Note: Member volumes cannot be removed from FlexGroup volumes.

NetApp SyncMirror (Mirrored Aggregates)

FlexGroup volumes can reside on aggregates that participate in a NetApp SyncMirror® configuration, which is a way to replicate aggregates internally for extra data protection functionality. For more information about SyncMirror, see this NetApp Support site page.

Note: SyncMirror does not provide the same functionality as StrictSync (NetApp SnapMirror Synchronous), which was new to ONTAP 9.5. FlexGroup volumes currently do not support StrictSync. For the latest data protection information regarding FlexGroup volumes, see TR-4678.

9 FlexGroup Sample Designs

NetApp ONTAP FlexGroup offers benefits that Infinite Volume does not offer, in that a FlexGroup volume can be managed like a normal NetApp FlexVol volume. The following design variations are allowed with a FlexGroup volume.

FlexGroup volumes can:

- Share SVM as a FlexVol volume
- Share the same physical disks and aggregates as a FlexVol volume
- Be mounted to other FlexGroup or FlexVol volumes
- Be mounted below the FlexGroup level, similar to FlexVol volumes
- Share export policies and rules with FlexVol volumes

- Enforce quotas (starting in ONTAP 9.5)

FlexGroup volumes ideally should not:

- Be configured to span mixed disk or aggregate types (for example, member volumes of the same FlexGroup volume on SATA and SSD)
- Span nodes of different hardware types
- Span aggregates with uneven free capacity
- Span aggregates with uneven performance allocation

FlexGroup Volume Maximums

The stated supported limits for a FlexGroup volume are 200 constituent volumes, 20PB, and 400 billion files. However, these are simply the tested limits in a 10-node cluster. When you factor in the maximum volumes that are allowed per node (1,000) with 24 nodes (24,000 total), the limits can potentially expand dramatically.

Ultimately, there is no architectural limitation for a FlexGroup volume other than the underlying hardware capacities. If you want to exceed these limits, contact your NetApp sales representative to begin a qualification process.

To suggest the possibilities for future capacity expansion, Table 12 and Table 13 show the untested limits of what is possible for FlexGroup volumes. The example in Table 12 is a maximized 24-node cluster (NetApp FAS9000). The example in Table 13 is a 24-node NetApp AFF A700 cluster. Theoretical limits might vary depending on platforms being used.

When the per-node volume limit is 1,000 per node (2,000 per HA pair), the following considerations apply:

- Each node needs a root volume (24 volumes).
- Each SVM needs a minimum of one vsroot volume (one volume per SVM).
- The hardware has a maximum capacity (400TB aggregates before ONTAP 9.2, 800TB per aggregate [AFF only] in ONTAP 9.2 and later, 14400TB total raw capacity per node, and so on).
- Volumes have capacity limits of 2 billion files and 100TB.

The node root and vsroot volumes remove at least 25 available volumes from the configuration. This example shows a cluster with only a single data SVM to maximize the potential capacity.

Note: The following configurations are theoretical. Do not try to use these configurations without review and approval from NetApp engineering. For details, contact your NetApp representative or email flexgroups-info@netapp.com.

Table 12) Architectural maximum limits for FlexGroup volumes on FAS9000.

Cluster Size	Architectural Maximum Member Volumes per FlexGroup Volume	Architectural Maximum Capacity per FlexGroup Volume	Theoretical Maximum Inodes per FlexGroup Volume
24 nodes	23,975 (1,000 volumes per node * 24 nodes – 25 reserved volumes)	176PB (based on architectural hardware limits of the FAS9000)	~47 trillion inodes (based on 2 billion inodes * 23,975 FlexGroup members)

Table 13) Architectural maximum limits for FlexGroup volumes on AFF A700.

Cluster Size	Architectural Maximum Member Volumes per FlexGroup Volume	Architectural Raw Maximum Capacity per FlexGroup Volume	Possible Effective Capacity	Theoretical Maximum Inodes per FlexGroup Volume
24 nodes	23,975 (1,000 volumes per node * 24 nodes – 25 reserved volumes)	~180PB (based on 32TB SSD and architectural hardware limits of the AFF A700)	Up to ~700PB (workload dependent, with storage efficiencies at ratio of 5:1)	~47 trillion inodes (based on 2 billion inodes * 23,975 FlexGroup members)

FlexGroup Sample Design 1: FlexGroup Volume, Entire Cluster (24 Nodes)

A FlexGroup volume can span an entire 24-node cluster, thus gaining the benefits of using all of the available hardware in the cluster with a single distributed namespace. In addition to using all your available hardware, you get the added benefit of gaining more potential capacity and more volume affinities in workloads. Table 14 breaks down the potential volume affinities, capacity, and maximum inodes in various cluster sizes. For more information about volume affinities, see the section "Volume Affinity and CPU Saturation."

Table 14) Cluster maximums, various cluster sizes with eight FlexVol members per node.

Cluster Size	Available Affinities per FlexGroup Volume	Maximum Capacity per FlexGroup Volume*	Maximum Inodes per FlexGroup Volume*
2	16 (before ONTAP 9.4) 32 (ONTAP 9.4 and later)	1600TB (~1.6PB) 3200TB (~3.2PB)	32,641,751,216 65,283,502,432
4	32 (before ONTAP 9.4) 64 (ONTAP 9.4 and later)	3200TB (~3.2PB) 6400TB (~6.4PB)	65,283,502,432 130,567,004,864
8	64 (before ONTAP 9.4) 128 (ONTAP 9.4 and later)	6400TB (~6.4PB) 12800TB (~12.8PB)	130,567,004,864 261,134,009,728
12	96 (before ONTAP 9.4) 192 (ONTAP 9.4 and later)	9600TB (~9.6PB) 19200TB (~19.2PB)	195,850,507,296 391,701,014,592
16	128 (before ONTAP 9.4) 256 (ONTAP 9.4 and later)	12800TB (~12.8PB) 25600TB (~25.6PB)	261,134,009,728 522,268,019,456
20	160 (before ONTAP 9.4) 320 (ONTAP 9.4 and later)	16000TB (~16PB) 32000TB (~32PB)	326,417,512,160 652,835,024,320
24	192 (before ONTAP 9.4) 384 (ONTAP 9.4 and later)	19200TB (~19.2PB) 38400TB (~38.4PB)	391,701,014,592 783,402,029,184

*Keep in mind that eight FlexVol members per node is a best practice, not a requirement. FlexGroup volumes can have as many—or as few—FlexVol members as desired. However, current testing shows that eight members produce the best results. ONTAP 9.4 increased the maximum affinities per node to 16, so the best practice recommendations might change.

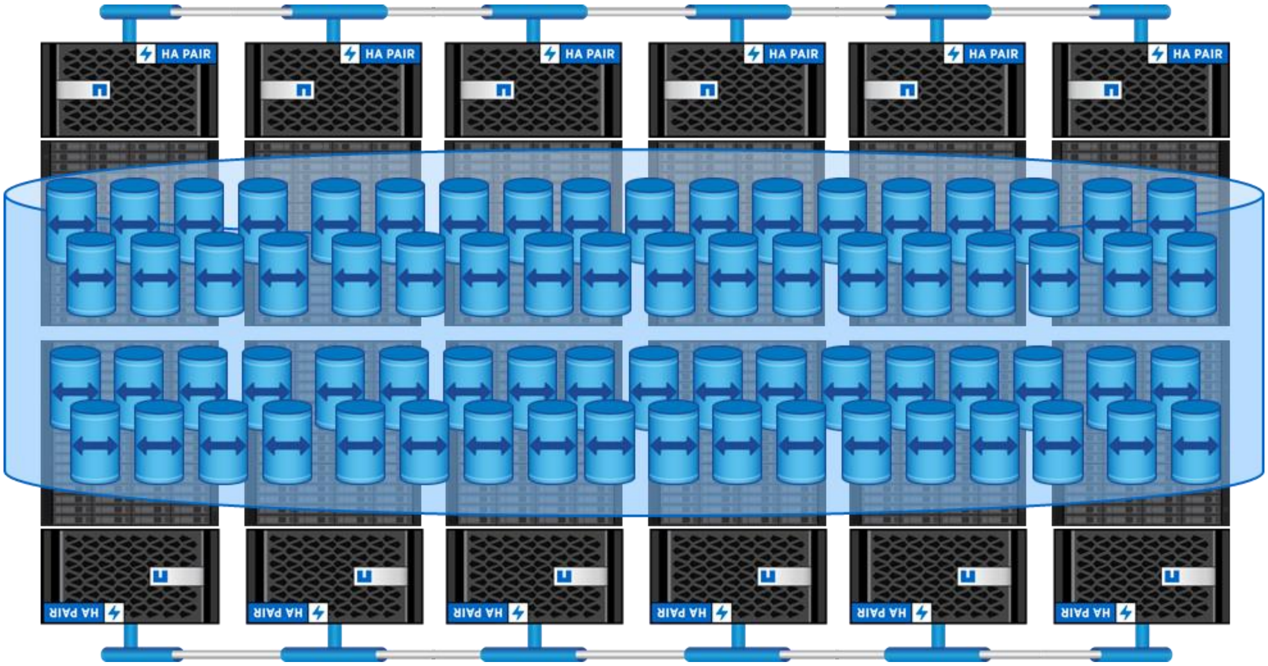
Considerations

- If you use an entire cluster to host a FlexGroup volume, keep in mind the information in the section "Cluster Considerations."

Use Cases

- Immense capacity (archives, scratch space, and media repositories)
- Workloads that require immense compute power in addition to storage (EDA)

Figure 60) FlexGroup volume, entire cluster (24 nodes).



FlexGroup Sample Design 2: Multiple Nodes, Aggregates, Partial Cluster

Sometimes, storage administrators might not want to span a FlexGroup volume across the nodes of an entire cluster. The reasons include, but are not limited to, the following:

- Mix of hardware or FAS (that is, some nodes are AFF)
- Mix of aggregate or disk types (that is, hybrid aggregates on the same node)
- Desire to dedicate nodes to specific tasks, storage tiers, or tenants
- In these scenarios, the FlexGroup volume can be created to use only specific aggregates, whether on the same node or on multiple nodes. If a FlexGroup volume has already been created, the member FlexVol volumes can be moved nondisruptively to the desired nodes and aggregates. For details, see "When to Use Nondisruptive Volume Move."

Considerations

When you try to create a FlexGroup volume on a mix of nodes and aggregates, the automated commands are not of much use. Instead, use `volume create` or the GUI, where it is possible to specify aggregates on FlexGroup creation. For already-created FlexGroup volumes, the command line is the only option.

Use Cases

- Mixed workloads (high performance + archive)
- Mixed cluster hardware
- Nodes with hybrid aggregates

Figure 61) Multiple nodes, partial cluster.

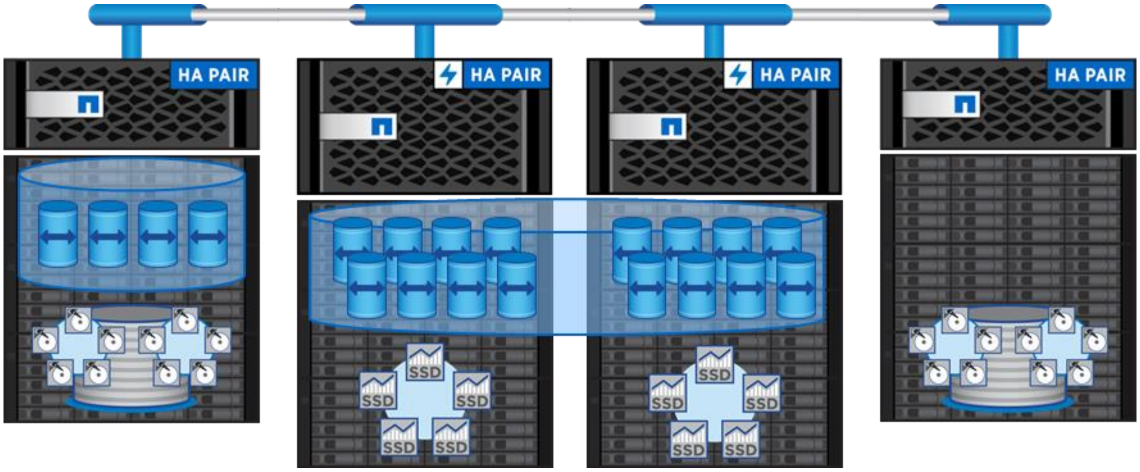
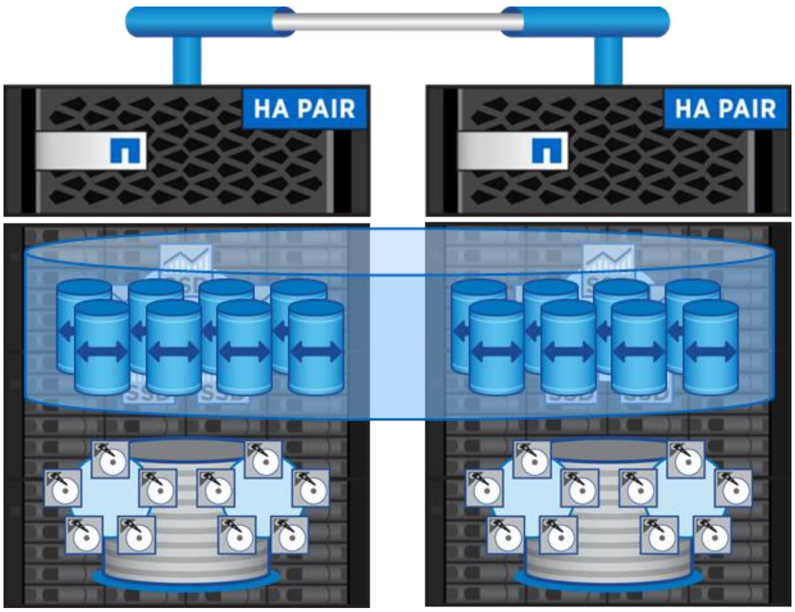


Figure 62) Multiple aggregates, partial cluster.



FlexGroup Sample Design 3: FlexGroup, Single Node

An ONTAP cluster uses a back-end 10GB/40GB cluster network to pass reads and writes from a node that receives an I/O request on a data LIF to the node that owns the physical data. When traffic is remote, a small penalty is incurred (about 5% to 10%) for remote I/O as these packets are processed. When traffic is all local to the node that owns the data, no cluster back end is used. Also, NAS operations get special bypass consideration to direct requests to disk even faster, so there is a benefit to going locally to a node.

With FlexGroup, there is no manual intervention of control over where a data requests lands; ONTAP controls that portion for simplicity’s sake. Because of this aspect, if a FlexGroup volume spans multiple nodes in a cluster, there is going to be indirect traffic over the cluster interconnects.

Although FlexGroup concurrency more than outweighs any performance penalty for remote traffic, you can achieve some performance gains by isolating a FlexGroup volume to a single node.

Figure 63 shows a single FlexVol volume that is accessed 100% locally on an AFF A700 node versus a single FlexGroup volume with eight FlexVol members that is also accessed 100% locally. The test used was a Git clone during a compilation of the GCC library. The same testing equipment and data described in AFF A700 Testing in section 7.3 were used.

This test shows that a clusterwide FlexGroup volume gives marginally better completion times because more hardware can be used. As extra threads are added to a local FlexGroup volume, the completion times start to get longer because the hardware can't keep up as well. In these tests, 64 threads appear to be the tipping point for local FlexGroup performance. However, both FlexGroup volumes are two to three times faster than a local FlexVol volume and have a more gradual performance curve.

Figure 63) Git clone completion times comparison.

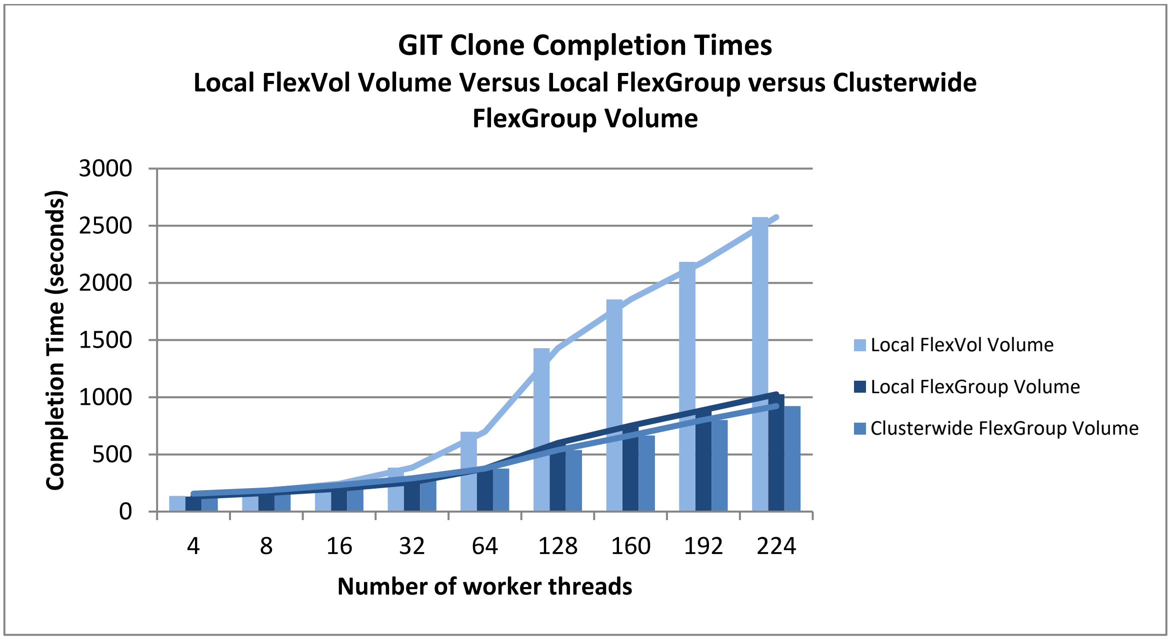
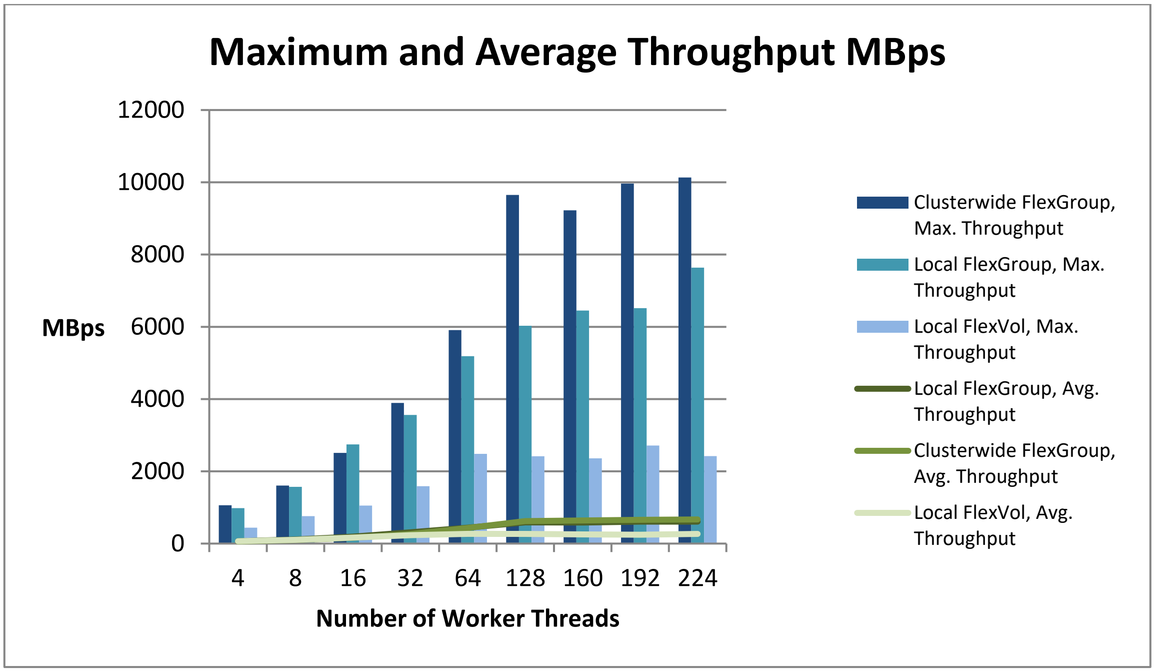


Figure 64 shows average and maximum throughput for the local FlexVol volume versus the local FlexGroup volume. For good measure, the clusterwide FlexGroup volume was also added for comparison. The local FlexGroup volume shows better overall throughput than the clusterwide FlexGroup volume until it reaches 16 threads. Then the local FlexGroup volume starts to lag behind slightly until 64 threads. At that point, the clusterwide FlexGroup volume zooms past the local FlexGroup volume. Then the clusterwide FlexGroup volume takes advantage of having extra hardware to work with. The FlexVol volume doesn't really compare favorably in this workload.

Figure 64) Average and maximum throughput comparison.



In Figure 65 and Figure 66, we compare read and write throughput, respectively, with the local and clusterwide FlexGroup volumes. At the 64-thread tipping point, the local FlexGroup volume starts to show a shift. Read throughput increases, while write throughput decreases. The cluster-wide FlexGroup volume shows the opposite trend.

Figure 65) Maximum read throughput comparison.

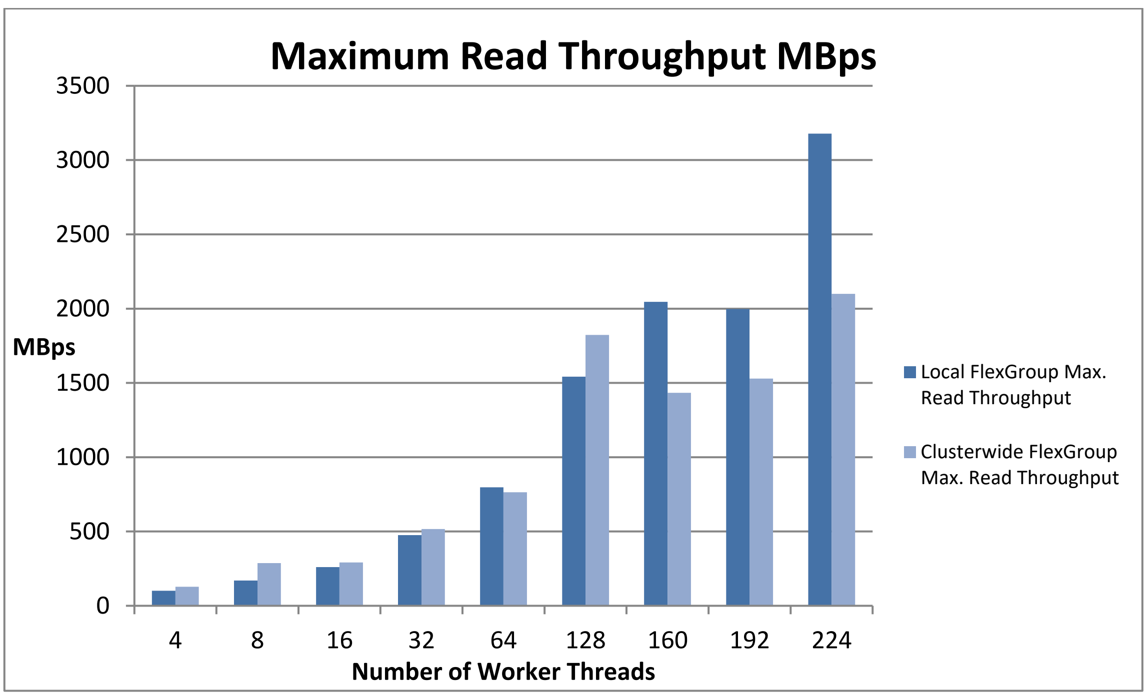


Figure 66) Maximum write throughput comparison.

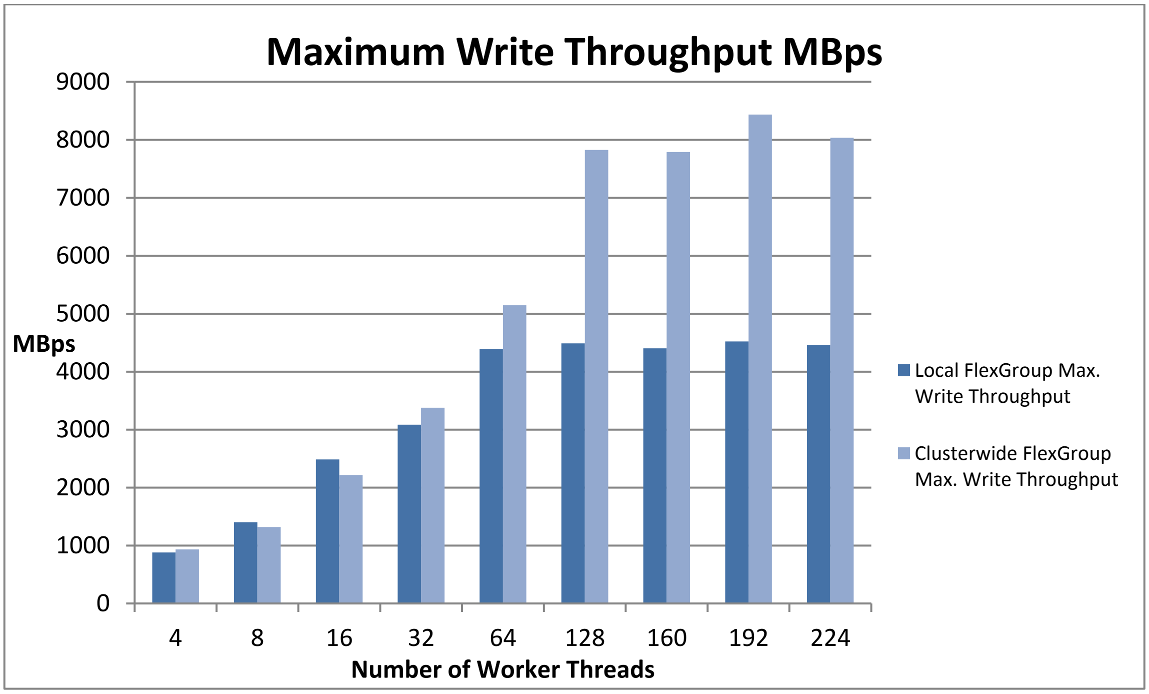
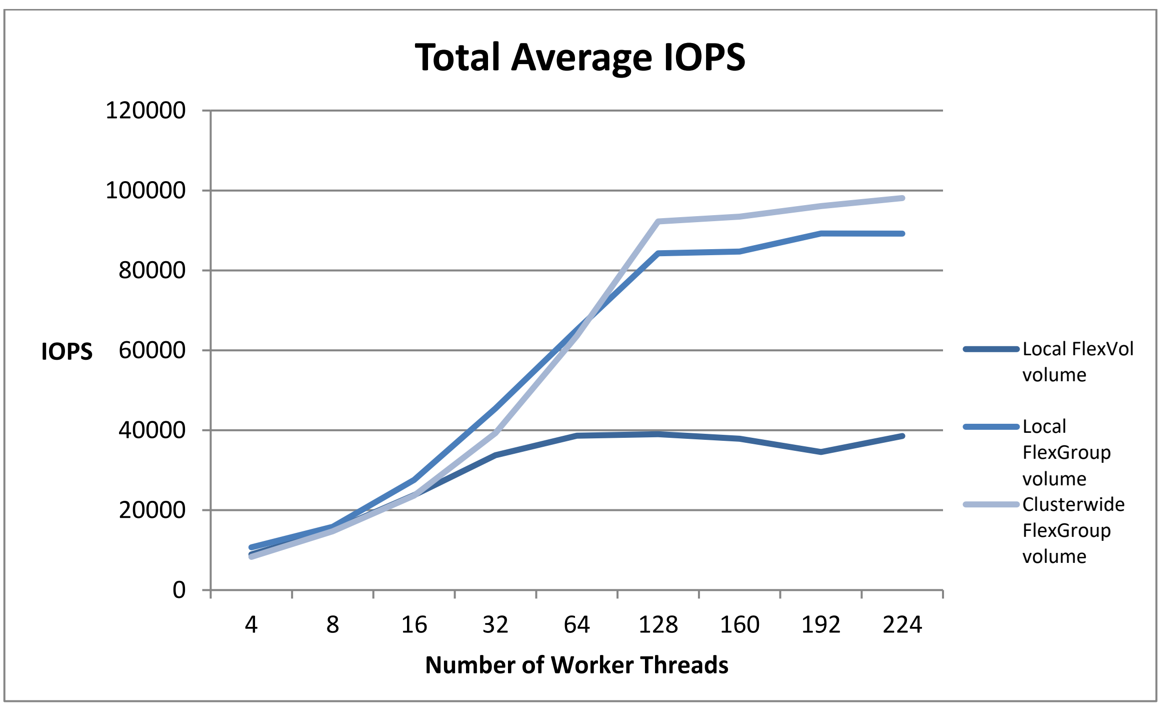


Figure 67 displays the total average IOPS for a local FlexVol volume versus the local and cluster-wide FlexGroup configurations. The FlexGroup configurations produce twice the IOPS that the FlexVol volume does, with the local FlexGroup volume outperforming the cluster-wide FlexGroup volume until the 64-thread tipping point.

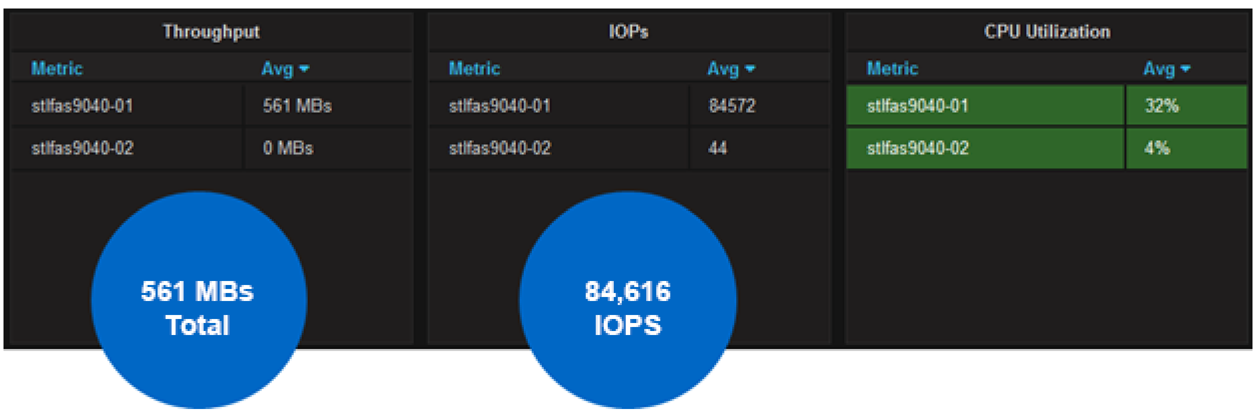
Figure 67) Total average IOPS comparison.



In this test, 64 worker threads appear to be a sweet spot. Let’s look at the average CPU utilization for a single-node FlexGroup volume versus a FlexGroup volume that spans the HA pair at just above 64 threads. Keep in mind that using more CPUs is a good thing; it means that work is being performed. That work is evidenced by the greater number of IOPS and the higher throughput for a FlexGroup volume that spans multiple nodes under the same workload.

Figure 68) Average CPU utilization, throughput, and IOPS for a FlexGroup volume—AFF A700 HA pair, 128 threads.

Figure 69) Average CPU utilization, throughput, and IOPS for a FlexGroup volume—single-node AFF A700, 128 threads.



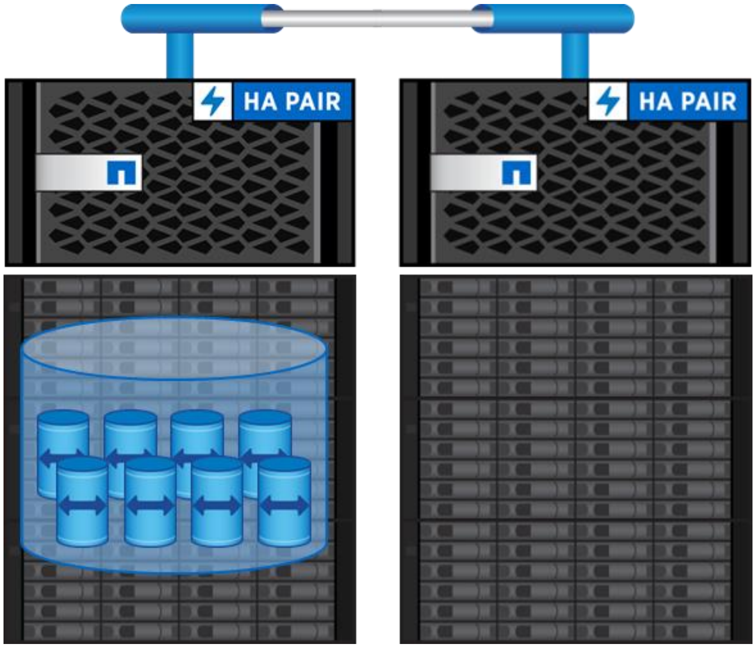
Considerations

When you use a single node for a FlexGroup volume, the gains that are realized by removing cluster interconnect traversal from the equation disappear relatively quickly. They disappear as load is added to the node and CPU, RAM, network bandwidth, and disk utilization becomes an issue. Usually, it makes more sense to spread the FlexGroup volume across multiple nodes rather than trying to save minimal cluster interconnect bandwidth. This architecture consideration becomes even more apparent with the introduction of 40GB Ethernet networks on the cluster back end with newer NetApp systems.

Use Cases

- High read workloads
- Need to isolate workloads to nodes
- Need to keep traffic off the cluster network

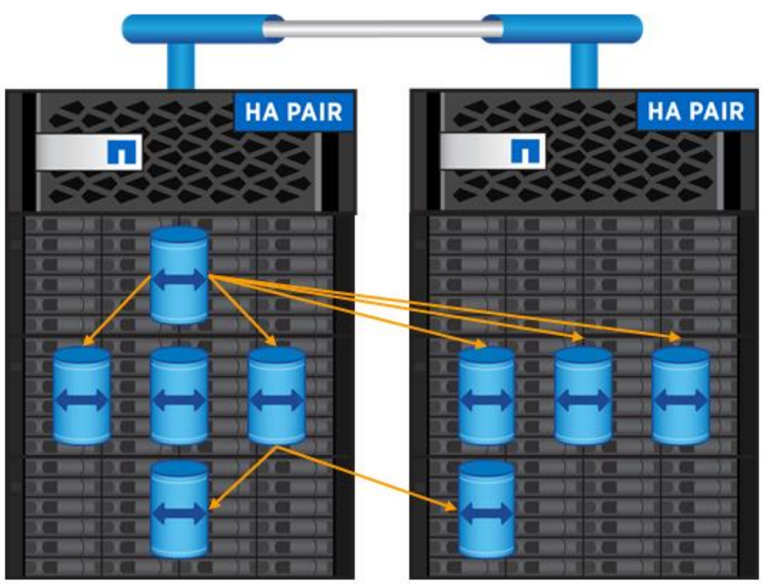
Figure 70) FlexGroup volume, single node.



FlexGroup Sample Design 4: FlexGroup Volumes Mounted to FlexGroup Volumes

With FlexVol volumes in ONTAP, you can mount volumes to other volumes to span the cluster and get >100TB in capacity, which was not possible with a single FlexVol volume. This method of designing a file system compares favorably with FlexGroup in terms of performance; often, performance with this design is a bit better than with FlexGroup volumes. However, the management overhead of creating multiple FlexVol volumes across multiple nodes and mounting them to each other in the namespace takes valuable personnel hours. In addition, adding capacity and volume affinities carry similar management headaches.

Figure 71) FlexVol volumes that are mounted to each other within a cluster namespace.



FlexVol volumes that are mounted to each other also do not provide true large buckets of storage, because each FlexVol volume is considered to be a folder in the file system to clients with 100TB limits. These folders also prevent the applications from controlling the narrative in file system layout. Instead, storage administrators must create the folder structure, forcing the application teams to reorganize their applications to fit into the folder constraints.

Although the concept of mounting FlexVol volumes to each other creates some complications, it also offers the following distinct benefits (aside from potential performance gains):

- Granular control over subvolume export policies and rules
- Granular control over data location
- The ability to mirror NetApp SnapMirror volumes at a more succinct level
- The ability to create a multitenant data organization without needing to create SVMs

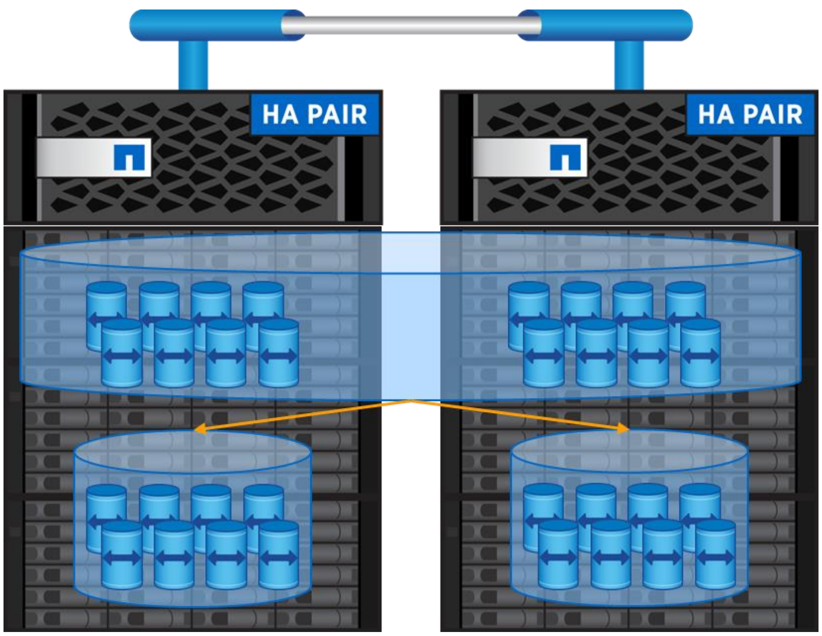
With FlexGroup volumes, the goal is to eliminate the management headaches of the multimounted FlexVol volume by providing a large bucket of storage that is easy to set up and manage.

However, the innovative performance capabilities of FlexGroup and FlexVol volumes, and the unified management capabilities of the two technologies, give storage administrators even greater freedom with the file system configuration. With a FlexGroup volume, you can reap the performance benefits of volume concurrency along with the granular control of mounting FlexGroup volumes to one another. When combined with the notion of a single-node FlexGroup volume, this methodology can also create more visibility in the physical location of data.

Use Cases

- More granular control over export policies and rules
- Greater control over the physical location of data
- An increased amount of data in the same volume that can be mirrored with SnapMirror (because of the 32 members per FlexGroup SnapMirror constraint present in ONTAP 9.4 and earlier)
- Mixing file-size workloads to get a more even distribution across the FlexGroup volume (that is, large files versus many small files)

Figure 72) FlexGroup volume mounted to FlexGroup volume.



FlexVol Volumes Mounted to FlexGroup Volumes

NetApp FlexVol volumes can also mount to FlexGroup volumes, and conversely. This configuration is another possibility with a FlexGroup solution.

Use Cases

- More granular control over export policies and rules
- Greater control over the physical location of data
- Features and functionality that are supported in FlexVol that aren't supported in FlexGroup use cases where a workload might occasionally create a large file or a small file that grows over time (for example, if a set of files gets zipped up to a larger zip file)

10 FlexGroup Management Best Practices

This section covers tips and best practices for managing NetApp ONTAP FlexGroup volumes.

10.1 Viewing FlexGroup Volumes

FlexGroup volumes are created through the NetApp ONTAP GUI or through the command line and are designed to behave, from a storage administrator's perspective, as a regular NetApp FlexVol volume. However, the FlexGroup volume is not simply a FlexVol volume; instead, it is made up of a series of FlexVol volume members that act in concordance across the FlexGroup volume. NetApp ONTAP uses these member volumes on ingest of data to provide multiple affinities across the file system, which provides capacity and performance gains.

Usually, a FlexGroup volume can be managed at the FlexGroup level. For instance, when growing a FlexGroup volume, you should run the `volume size` command at the FlexGroup level so that all members are given equivalent capacities. This step helps avoid performance imbalances due to the ingest heuristics being thrown off unnaturally, and it helps to keep all member FlexVol volumes at roughly the same free capacity. Maintaining equal free capacity helps avoid scenarios in which a FlexVol member filling prematurely would cause FlexGroup-level Snapshot failures or `ENOSPC` (out of space) messages for the entire FlexGroup volume.

However, there are instances in which you want to view the individual FlexVol members, such as the following:

- To view the member capacity usage (are we getting close to full?)
- To view individual member performance (do I need to use `volume move`?)

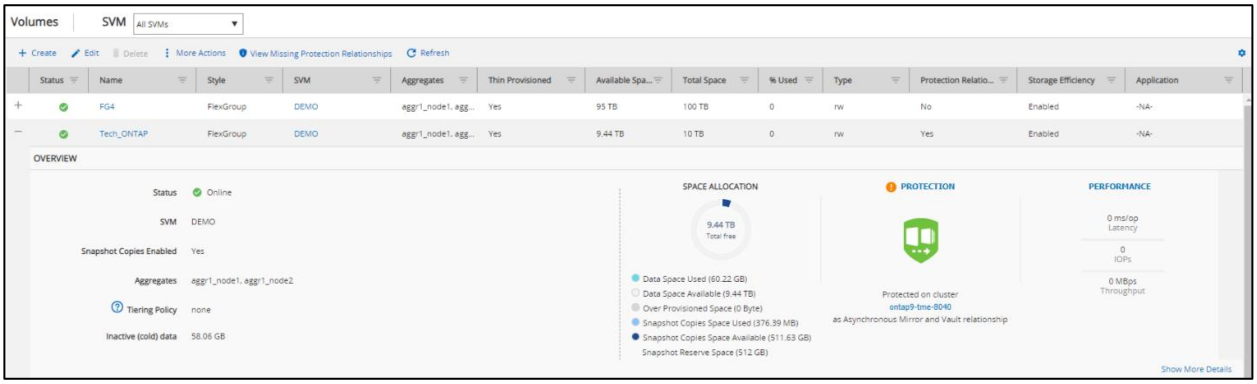
The following sections offer guidance on viewing FlexGroup volumes.

ONTAP System Manager

With ONTAP System Manager, you can view and manage a FlexGroup volume at the FlexGroup level through the FlexGroup tab; however, there are no views for member volumes. This is by design—a FlexGroup volume should be simple to manage. ONTAP System Manager provides useful information about the FlexGroup volume in these views, such as data protection information, real-time performance, and capacity information. For more information about FlexGroup volume use with ONTAP System Manager, see TR-4557.

Note: Keep in mind that ONTAP System Manager cannot provide space allocation information for FlexGroup volumes that are thin-provisioned.

Figure 73) ONTAP System Manager FlexGroup volume view.



Active IQ Performance Manager

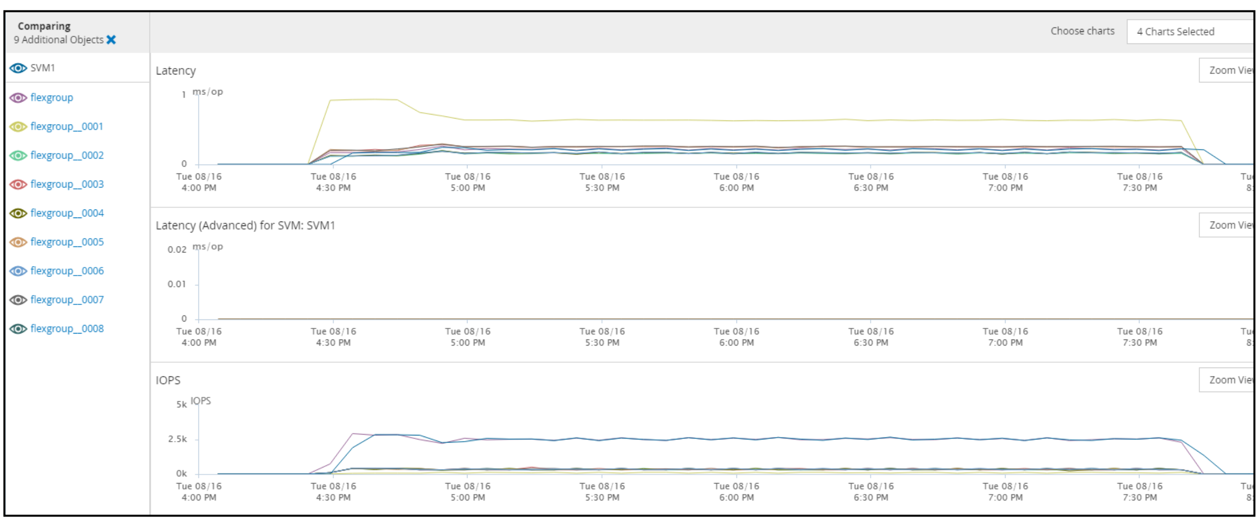
NetApp [Active IQ Performance Manager]() collects an archive of performance statistics for ONTAP, including FlexGroup member volumes. This granular view of the FlexGroup volume allows storage administrators to evaluate individual member FlexVol volumes for performance anomalies and to take corrective actions as needed, such as the following:

- Adding more space
- Adding more members (`volume expand`)
- Nondisruptive volume move

Note: These tasks cannot be carried out in Active IQ Performance Manager. Currently, only the command line and the System Manager GUI can carry out these tasks.

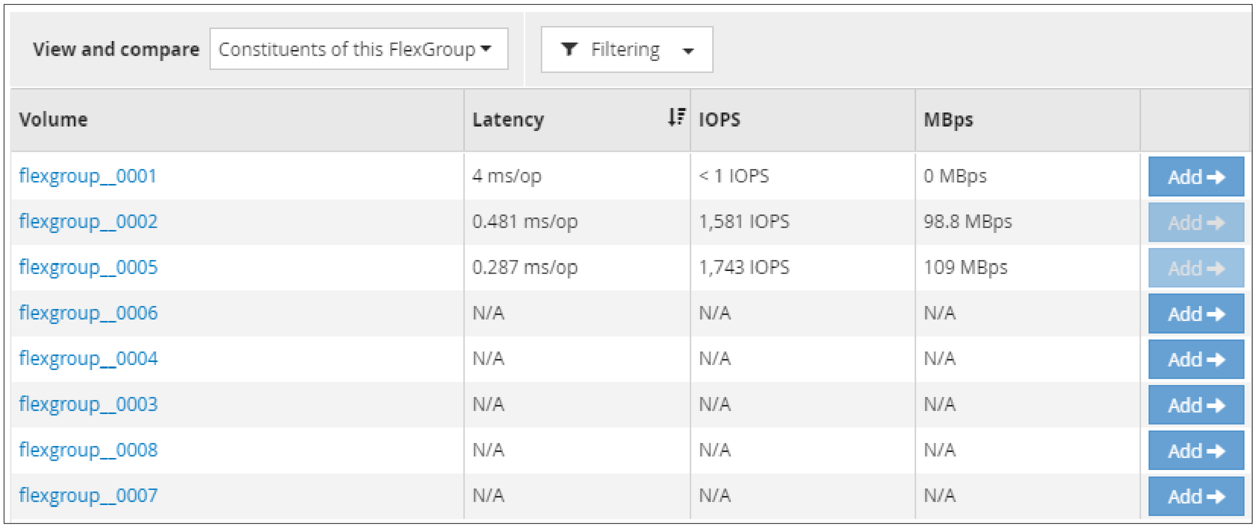
Figure 74 shows several FlexVol members and their corresponding performance. Each line represents a FlexVol member.

Figure 74) Active IQ Performance Manager FlexGroup volume view.



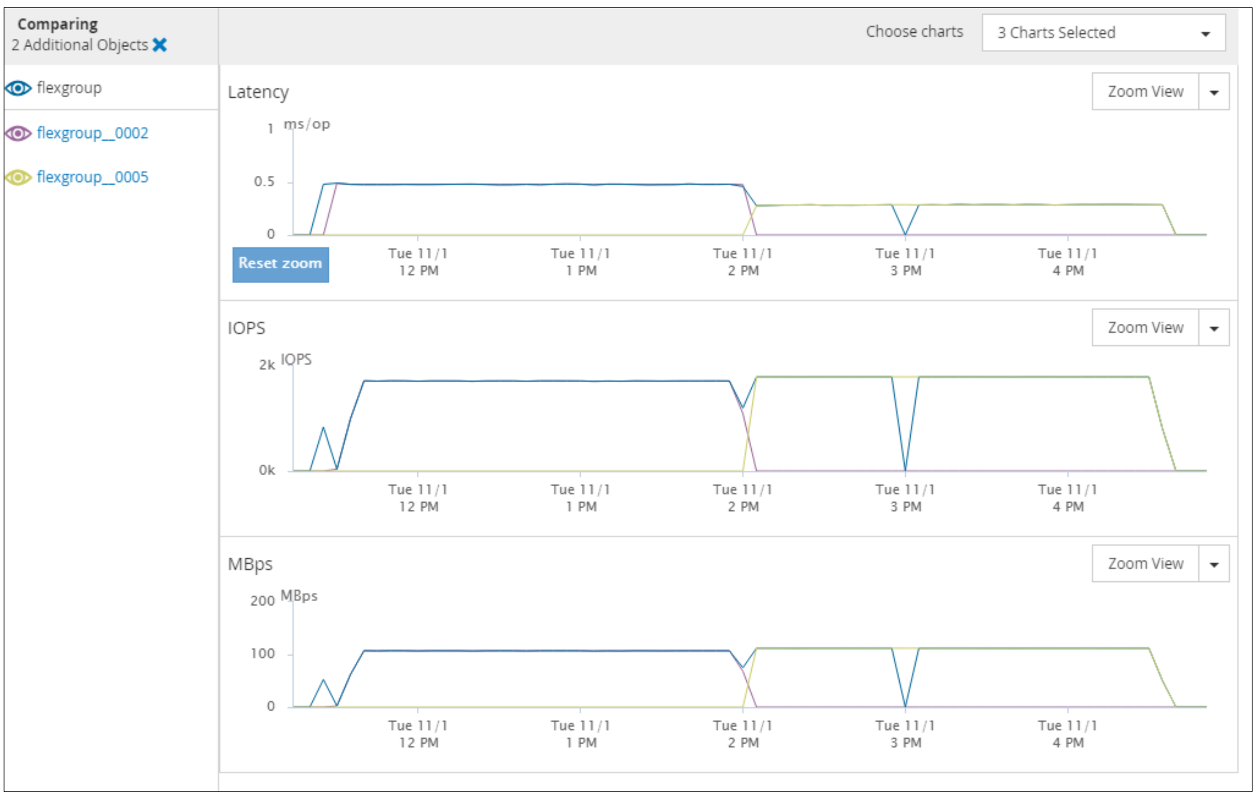
In Figure 75, two 1TB files were written to a FlexGroup volume. In the chart, we can see which member volumes took on that workload (members 2 and 5), and we see a summary of the workload performance. In Figure 76, we can see the IOPS and MBps graphs.

Figure 75) Member volume performance chart.


View and compare: Constituents of this FlexGroup | Filtering

Volume	Latency	IOPS	MBps	
flexgroup__0001	4 ms/op	< 1 IOPS	0 MBps	Add
flexgroup__0002	0.481 ms/op	1,581 IOPS	98.8 MBps	Add
flexgroup__0005	0.287 ms/op	1,743 IOPS	109 MBps	Add
flexgroup__0006	N/A	N/A	N/A	Add
flexgroup__0004	N/A	N/A	N/A	Add
flexgroup__0003	N/A	N/A	N/A	Add
flexgroup__0008	N/A	N/A	N/A	Add
flexgroup__0007	N/A	N/A	N/A	Add

Figure 76) Member volume graphs.



Active IQ Unified Manager

With Active IQ Unified Manager, storage administrators can use a single dashboard to review the health of a NetApp ONTAP cluster. It can also integrate with Active IQ Performance Manager to give a single management view.

With Active IQ Unified Manager, you can review FlexGroup volume capacity, configurations, and storage efficiencies in a graphical format. For member volume information about performance, use Active IQ Performance Manager. For capacity information about member volumes, use the command line.

Figure 77) Active IQ Unified Manager; FlexGroup capacity view.

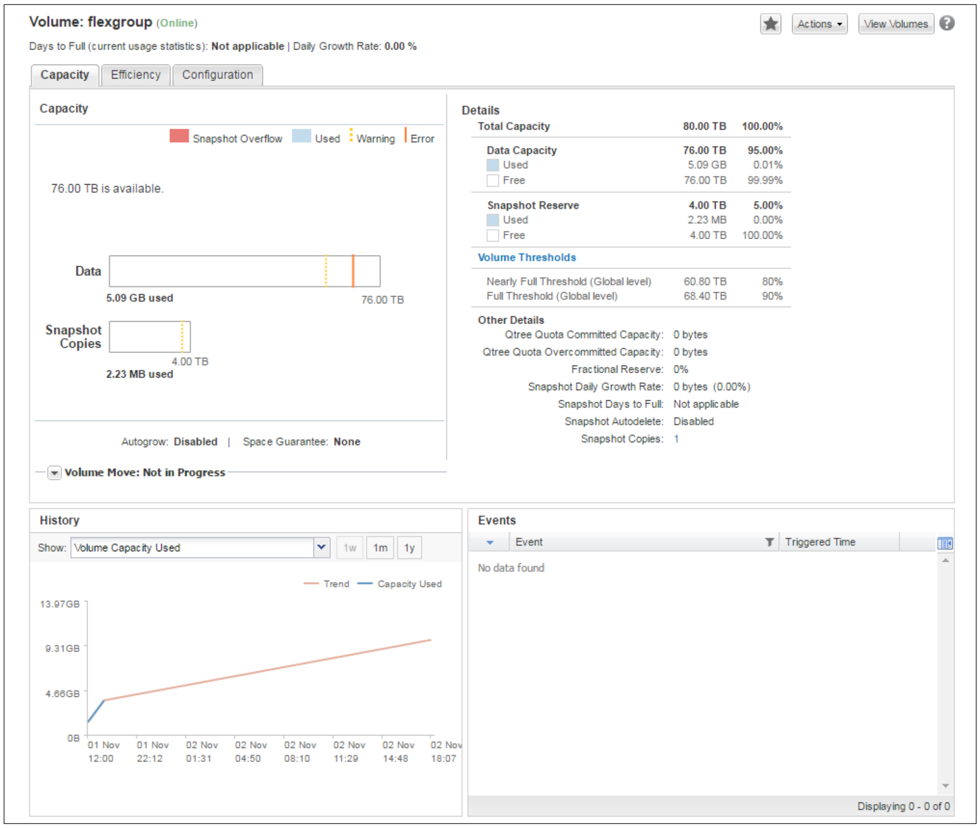
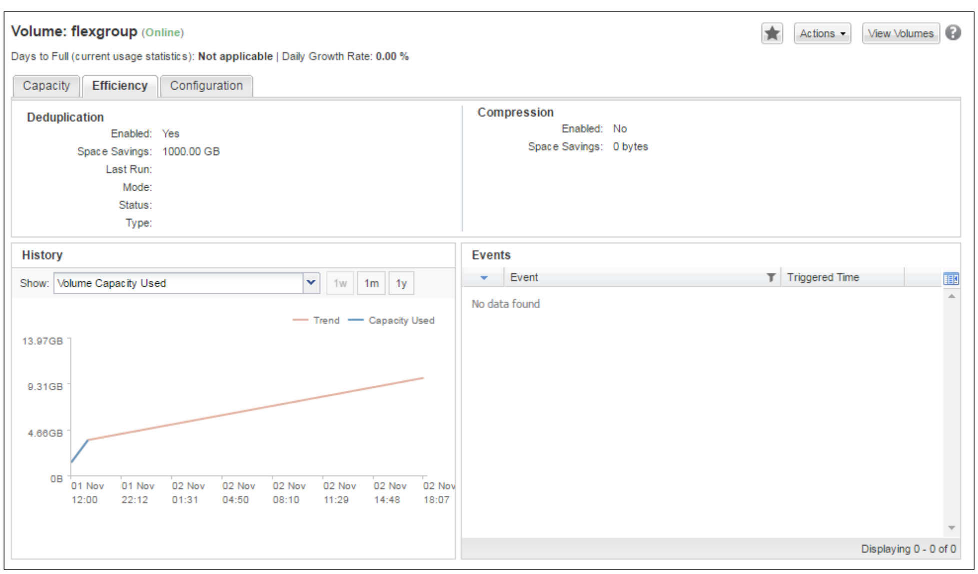


Figure 78) Active IQ Unified Manager FlexGroup efficiency view.



Command Line

The CLI is another way to view FlexGroup volume information. Each privilege level gives a different set of options for viewing the FlexGroup volume properties.

Admin Privilege Level

- Total capacity (total, available, and used: calculated from all the member volumes), storage efficiencies
- NetApp Snapshot reserve or Snapshot policy
- List of aggregates and nodes that the FlexGroup volume spans
- Volume style and extended volume style (tells us whether the volume is a FlexGroup volume)
- Security style, owner, or group
- Junction path
- Maximum files and inodes
- Member volume information (through `-is-constituent true` or `volume show-space`)

Advanced Privilege Level

- Maximum directory size
- FlexGroup master set ID (MSID)
- Whether the volume was transitioned from 7-Mode (important for FlexVol to FlexGroup volume conversion)
- FlexGroup maximum member volume sizes

Diag Privilege Level

- Detailed member volume information (capacity, used, and so on)
- FlexGroup ingest statistics (`flexgroup show`)

Note: Member volume space information can be seen in the **admin privilege** level by using the command `volume show-space`. For details, see the section "Capacity Monitoring and Alerting with the Command Line."

10.2 Monitoring FlexGroup Capacity

This section covers various methods of monitoring a FlexGroup volume's capacity, including viewing total storage efficiency savings. Monitoring FlexGroup capacity is also possible with the NetApp FPolicy support introduced in ONTAP 9.4.

Total FlexGroup Capacity

The total FlexGroup capacity is a number that is derived from the following:

- **Total space.** Total combined allocated space for a FlexGroup volume (member volume capacity * number of members).
- **Available space.** The amount of space that is available in the most allocated member volume.

You can view the total FlexGroup capacity in ONTAP System Manager, in Active IQ Unified Manager, or through the CLI at the admin privilege level.

```
cluster::> vol show -fields size,used,percent-used,available -vserver SVM  -volume flexgroup
vserver volume    size available used   percent-used
------- --------- ---- --------- ------ ------------
SVM     flexgroup 80TB 15.66TB   5.08GB 80%
```

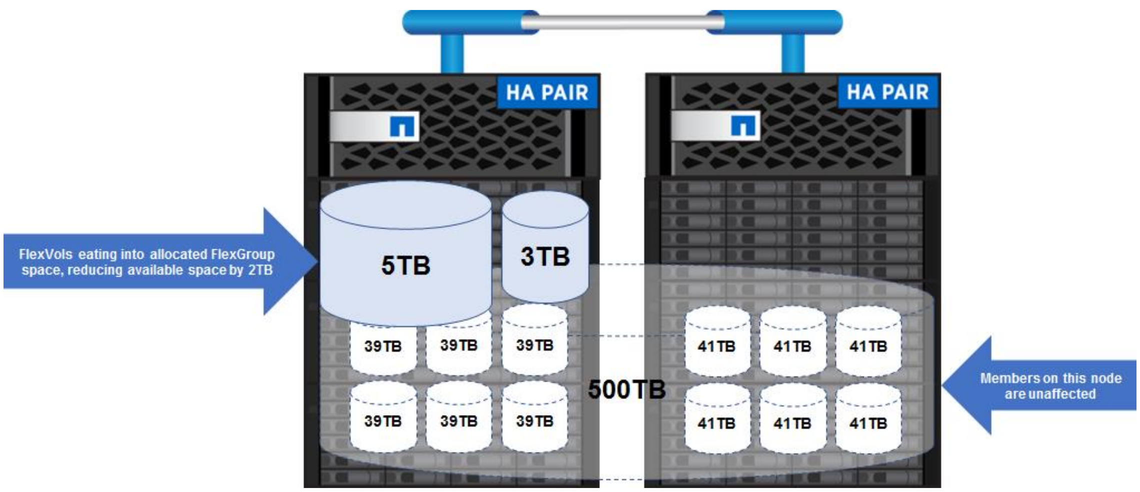

Effect of Overprovisioning or Thin Provisioning in a FlexGroup Volume

Overprovisioning or thin provisioning with a FlexGroup volume can be useful. These functions can help prevent the unwanted frequent remote placement that occurs when a FlexGroup member becomes closer to full (see TR-4557 for details). It also removes the virtual cap on a FlexVol volume and instead relies on the physical available space in an aggregate.

Thin provisioning should be used with the following caveats in mind:

- When a volume is out of space, it is truly out of space because the aggregate is out of space; more disk space must be added to remediate space issues.
- The space allocated does not necessarily reflect the actual space available; it is possible to allocate volumes that are much larger than their physical space.
- If your system is sharing aggregates with other volumes (FlexGroup or FlexVol volumes), you should use thin provisioning with caution. Existing FlexVol or FlexGroup volumes on the same aggregates as FlexGroup volumes can potentially affect how data is ingested. Existing volumes reduce the amount of space available for member volumes because used space eats into other volume allocations. Figure 79 illustrates how other volumes can eat into volumes with thin provisioning enabled.

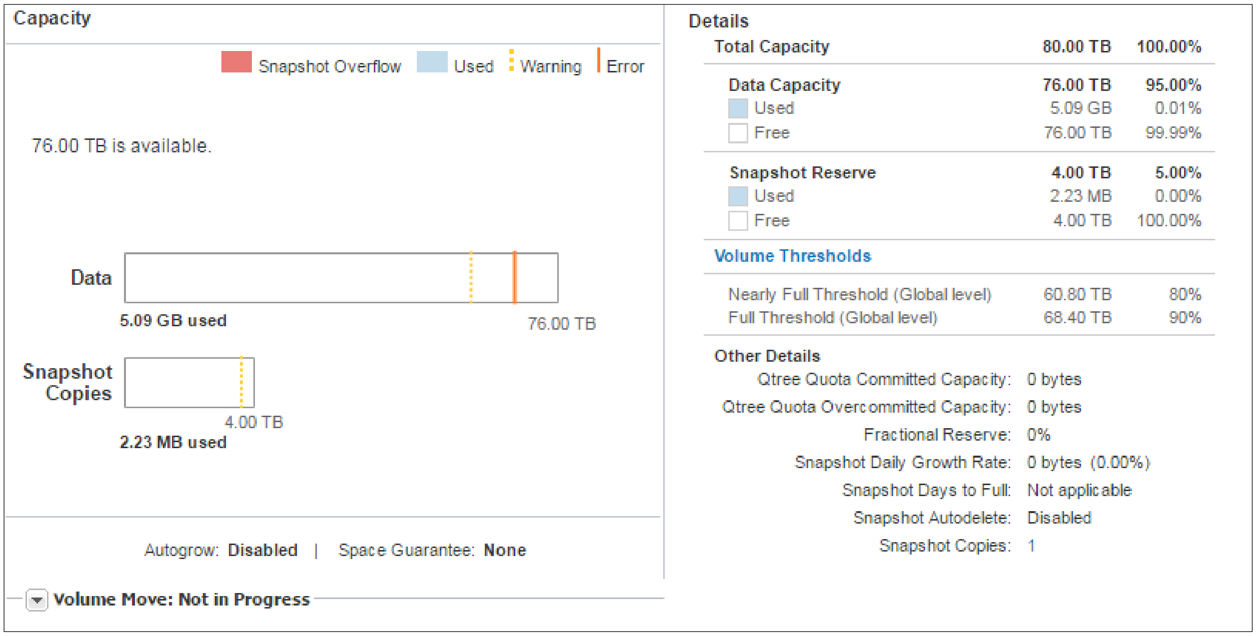
Figure 79) Capacity impact when thin-provisioned FlexGroup volumes exist with space-guaranteed FlexVol volumes.



- Using volume space guarantees can protect against other datasets affecting volume capacities, but they don't offer the most efficient use of your storage capacity.
- If you use thin provisioning, NetApp strongly recommends using Active IQ Unified Manager to monitor and alert for capacity.

Figure 80 shows that the FlexGroup volume has a total capacity of 80TB and 5GB used. Also, 4TB have been reserved for Snapshot copies (5%). The available space is 76TB.

Figure 80) FlexGroup capacity breakdown—Active IQ Unified Manager.



However, in the following CLI output, a few anomalies stand out:

- In Active IQ Unified Manager, the FlexGroup volume shows as having 76TB available, but in the CLI, only 11.64TB are available.
- The FlexGroup volume shows as having 11.64TB available, but the member FlexVol volumes all show roughly 5.8TB available.
- The percentage used for the FlexGroup volume shows as 85%, even though we have used only 5GB, which is a negligible amount of space compared with 80TB (5GB of 81920GB is less than 1%).
- The FlexGroup volume shows as 85% used, but the member FlexVol volumes all show as 41% used, despite each having a different amount of space per FlexVol member.

Example:

```
cluster::> volume show -is-constituent true -fields size,used,percent-used,available -vserver SVM
-volume flexgroup* -sort-by volume
vserver volume             size  available used    percent-used
------- ------------------ ----- --------- ------- ------------
SVM     flexgroup          80TB  11.64TB   5.08GB  85%
SVM     flexgroup__0001    10TB  5.81TB    147.5MB 41%
SVM     flexgroup__0002    10TB  5.83TB    145.2MB 41%
SVM     flexgroup__0003    10TB  5.81TB    144.9MB 41%
SVM     flexgroup__0004    10TB  5.83TB    148.0MB 41%
SVM     flexgroup__0005    10TB  5.81TB    4.08GB  41%
SVM     flexgroup__0006    10TB  5.83TB    147.6MB 41%
SVM     flexgroup__0007    10TB  5.81TB    145.3MB 41%
SVM     flexgroup__0008    10TB  5.83TB    146.5MB 41%
9 entries were displayed.
```

The anomalies are due to ONTAP calculating against the aggregate's available space. The FlexVol member volumes show equivalent available values depending on the aggregates where they are located.

```
cluster::> volume show -is-constituent true -fields available,aggregate -vserver SVM -volume
flexgroup* -sort-by aggregate
vserver volume             aggregate   available
------- ------------------ ----------- ---------
SVM     flexgroup__0001    aggr1_node1 5.81TB
SVM     flexgroup__0003    aggr1_node1 5.81TB
```



```
SVM     flexgroup__0005    aggr1_node1 5.81TB
SVM     flexgroup__0007    aggr1_node1 5.81TB
SVM     flexgroup__0002    aggr1_node2 5.83TB
SVM     flexgroup__0004    aggr1_node2 5.83TB
SVM     flexgroup__0006    aggr1_node2 5.83TB
SVM     flexgroup__0008    aggr1_node2 5.83TB

cluster::> storage aggregate show -aggregate aggr1* -fields availsize
aggregate   availsize
----------- ---------
aggr1_node1 5.81TB
aggr1_node2 5.83TB
2 entries were displayed.
```

Using thin provisioning means that you must consider the aggregate capacity and the volume footprint when monitoring space.

Best Practice 20: Using Thin Provisioning
If you use thin provisioning with ONTAP, it is important to use tools such as Active IQ Unified Manager or to set up monitoring through the CLI to track the available space in your storage system.

Capacity Monitoring and Alerting with the Command Line

When you use thin provisioning, you should use the command `storage aggregate show-space` with `volume show -is-constituent true`, `volume show-space`, and `storage aggregate show`. This approach provides better total visibility into space usage overall. In the command line, you can also use the `-sort-by` option to organize the list.

Note: To get an accurate portrayal of the space that is being used, pay attention to the `Physical Used` portion of the `volume show-space` command. You can find an example in "Command Examples."

Event Management System Messages

Event management system messages alert storage administrators about the capacity of volumes in ONTAP. The messages are listed in this section. You can view them in the command line with the command `event route show -messagename [message] -instance`. For an example of these messages, see Event Management System Examples.

Unmodifiable values:

- Severity level
- Corrective actions
- Description
- SNMP support

Modifiable values:

- Destinations
- Allowed drops or intervals between transmissions

When an event management system message that has SNMP support is triggered, an SNMP trap fires to the configured SNMP server. This action is specified through the `destinations` value. For more information about configuring event management system destinations, see the Express Guide for your specific version of ONTAP.

The default values for Nearly Full (Warning) and Full (Error) are as follows:

```
cluster::*> vol show -vserver SVM -volume flexgroup -fields space-nearly-full-threshold-
percent,space-full-threshold-percent
vserver volume       space-nearly-full-threshold-percent space-full-threshold-percent
```



```
------- ------       ---------------------------------- ---------------------------
SVM     flexgroup    95%                                98%
```

Event management system messages for `volume.full` look like the following:

```
11/28/2016 18:26:34 cluster-01
DEBUG         monitor.volume.full: Volume flexgroup@vserver:05e7ab78-2d84-11e6-a796-00a098696ec7
is full (using or reserving 99% of space and 0% of inodes).
```

In the preceding example, the following values are provided:

- The type of object
- The name of the volume
- The SVM (called `vserver` in the CLI) universal unique identifier (UUID)
- Percentage of space used
- Percentage of inodes used

You can use these values when testing event management system messages. When you look for which SVM is affected by the errors, use the UUID string at the advanced privilege level:

```
cluster::*> vserver show -uuid 05e7ab78-2d84-11e6-a796-00a098696ec7
                               Admin      Operational Root
Vserver     Type    Subtype    State      State       Volume     Aggregate
----------- ------- ---------- ---------- ----------- ---------- ----------
SVM         data    default    running    running     vsroot     aggr1_
                                                                 node1
```

Testing Event Management System Messages

To test an event management system message, use the `event generate` command (available at the diag privilege level). Each message has a unique string of values. The values for `volume.full` and `volume.nearlyFull` are listed in the preceding section. The following example shows how to construct a test message for a `volume.nearlyFull` event and the resulting event management system message:

```
cluster::*> event generate -message-name monitor.volume.nearlyFull -values Volume flexgroup
@vserver:05e7ab78-2d84-11e6-a796-00a098696ec7 95 0

cluster::*> event log show -message-name monitor.volume.nearlyFull
Time                Node             Severity      Event
------------------- ---------------- ------------- ---------------------------
11/28/2016 18:36:35 cluster-01
                                     ALERT         monitor.volume.nearlyFull: Volume
flexgroup@vserver:05e7ab78-2d84-11e6-a796-00a098696ec7 is nearly full (using or reserving 95% of
space and 0% of inodes).
```

Modifying the Volume Full and Nearly Full Thresholds

With a FlexVol volume, the default values for `full` and `nearlyFull` are fine because the volume is isolated to a single container. With a FlexGroup volume, by the time a member FlexVol volume reaches the full or nearly full threshold, the application or end user might already be seeing a performance degradation. This decreased performance is due to increased remote file allocation or a FlexGroup volume that is already reporting to be out of space because of a full or nearly full member volume. This approach is necessary for versions earlier than ONTAP 9.3 because the volume autogrow functionality was unavailable for a FlexGroup volume until ONTAP 9.3.

To help monitor for these scenarios, the volume full and nearly full thresholds might require adjustment to deliver warnings and errors before a volume fills up. Volumes have options to adjust these thresholds at the **admin privilege** level.

```
-space-nearly-full-threshold-percent
```



```
-space-full-threshold-percent
```

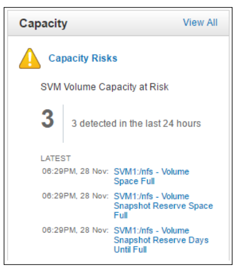
Use the `volume modify` command to adjust these thresholds.

Best Practice 21: Volume Space Threshold Recommendations for FlexGroup
Generally speaking, the nearly full threshold of a FlexGroup volume should be set to 80%, and full should be set to 90%. With these settings, you have enough time to remediate space issues by increasing the member volume sizes or adding more capacity in the form of additional member volumes through [`volume expand`](). These values can vary based on the average file size and the FlexGroup member volume size. For instance, a 1TB FlexGroup member can reach 80% immediately with an average file size of 800GB, but a 100TB FlexGroup member would take longer to hit that threshold.

Capacity Monitoring and Alerting in Active IQ Unified Manager

Active IQ Unified Manager provides methods to monitor and alert on various storage system functionalities, including used and free capacities. On the main Health page, Active IQ OnCommand displays active warnings and errors about capacity.

Figure 81) Active IQ Unified Manager capacity risks.

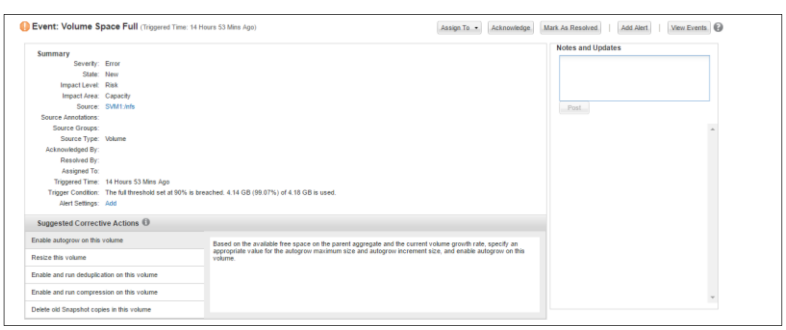


Also, Active IQ Unified Manager has a more detailed view of capacity-related events.

Figure 82) Capacity-related events—detailed view.

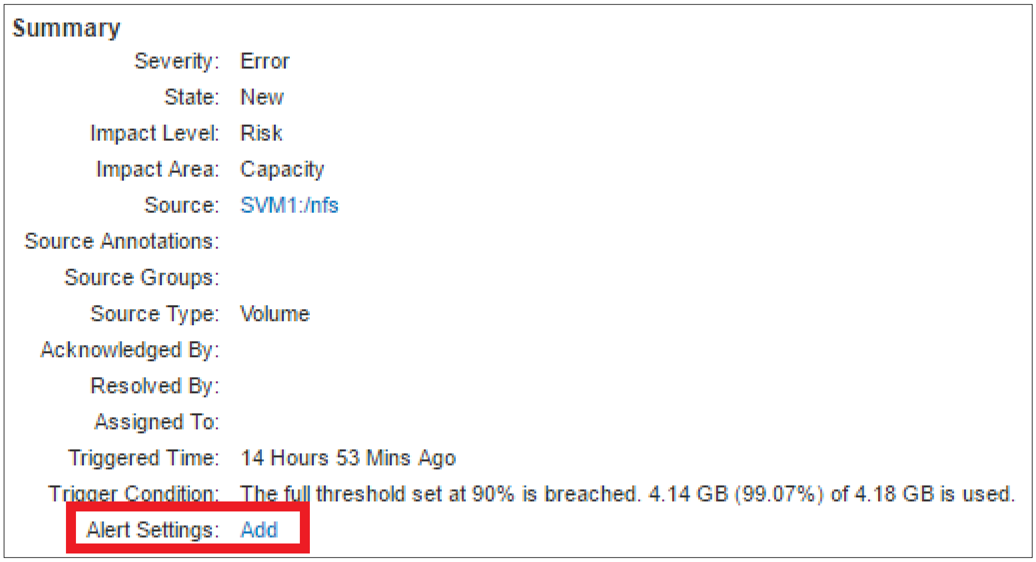
When you click one of the events, a full report of the issue is shown.

Figure 83) Volume full event—detailed view.



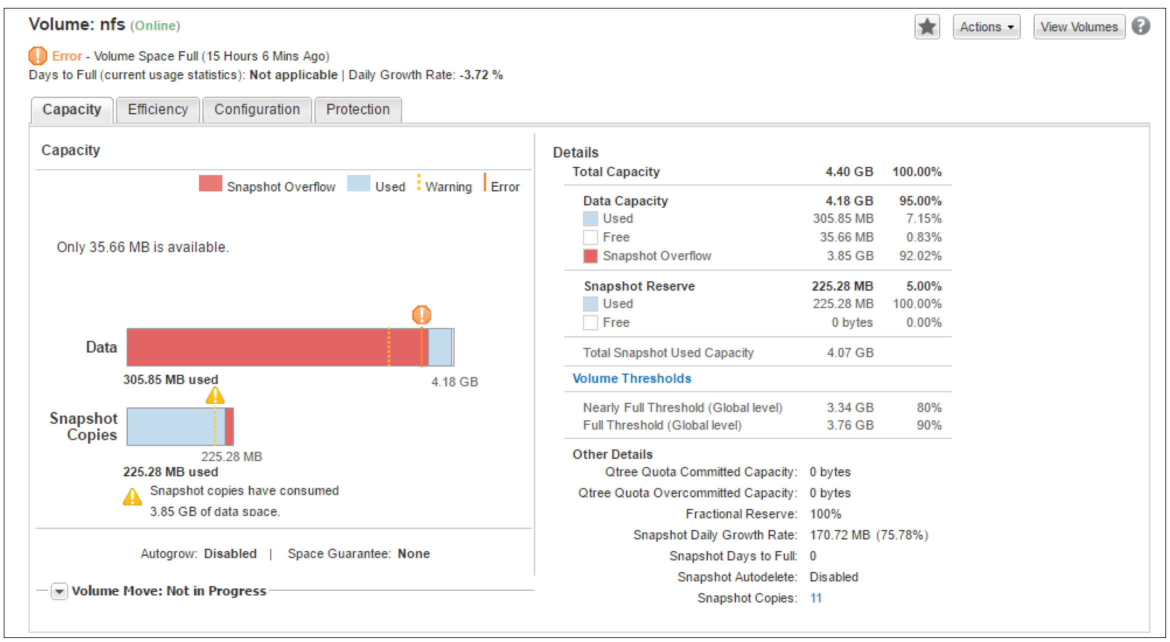
In this detailed view, you can also configure alerts. To do so, click the Add link next to Alert Settings.

Figure 84) Add an alert from an event.



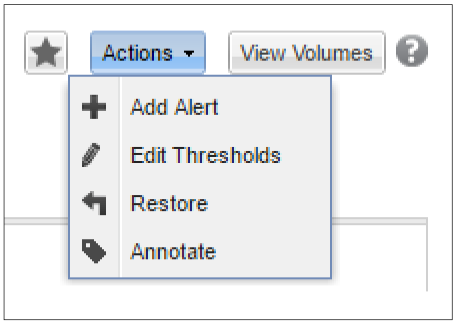
You can also view volume capacities from the Volume screen. When you click Storage > Volumes and select a volume, a screen like the following appears:

Figure 85) Volume capacity detail.



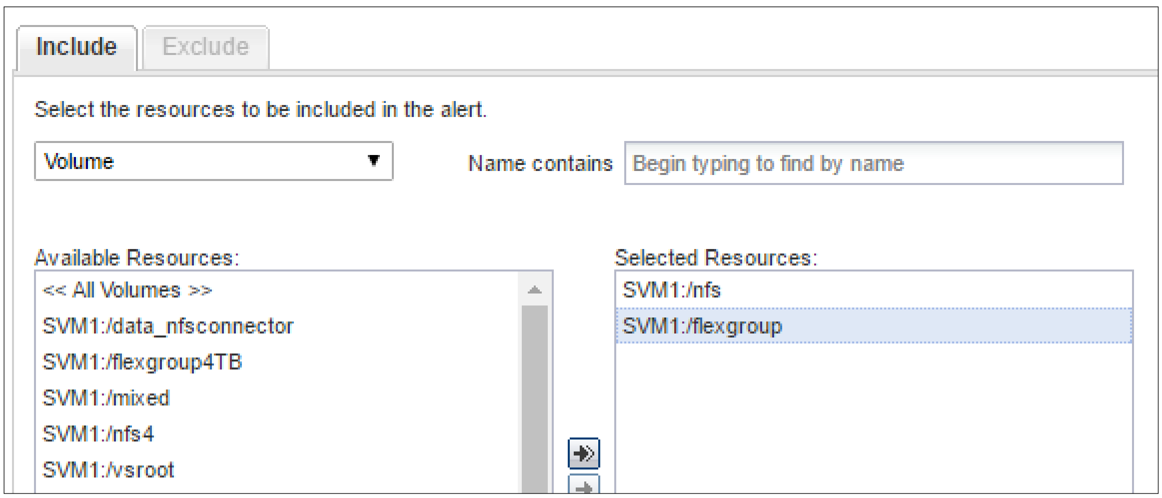
Click the Actions button to create alerts that are specific to the volume.

Figure 86) Adding an alert.



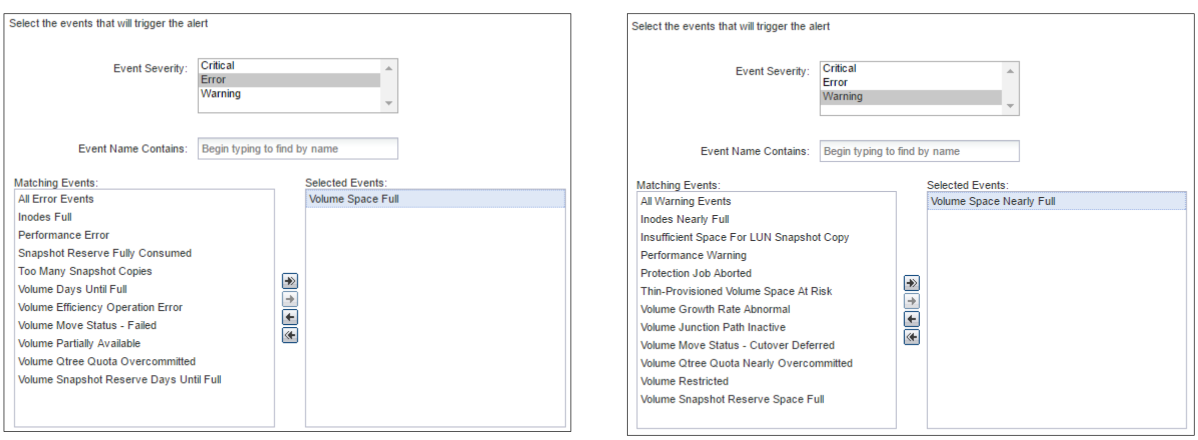
With the alert, you can add one or many volumes to various events (or exclude them).

Figure 87) Including volumes for alerts.



Events are organized by severity and include the Critical, Error, and Warning levels. Volume Space Full is included under the Error level, and Volume Space Nearly Full is under the Warning level.

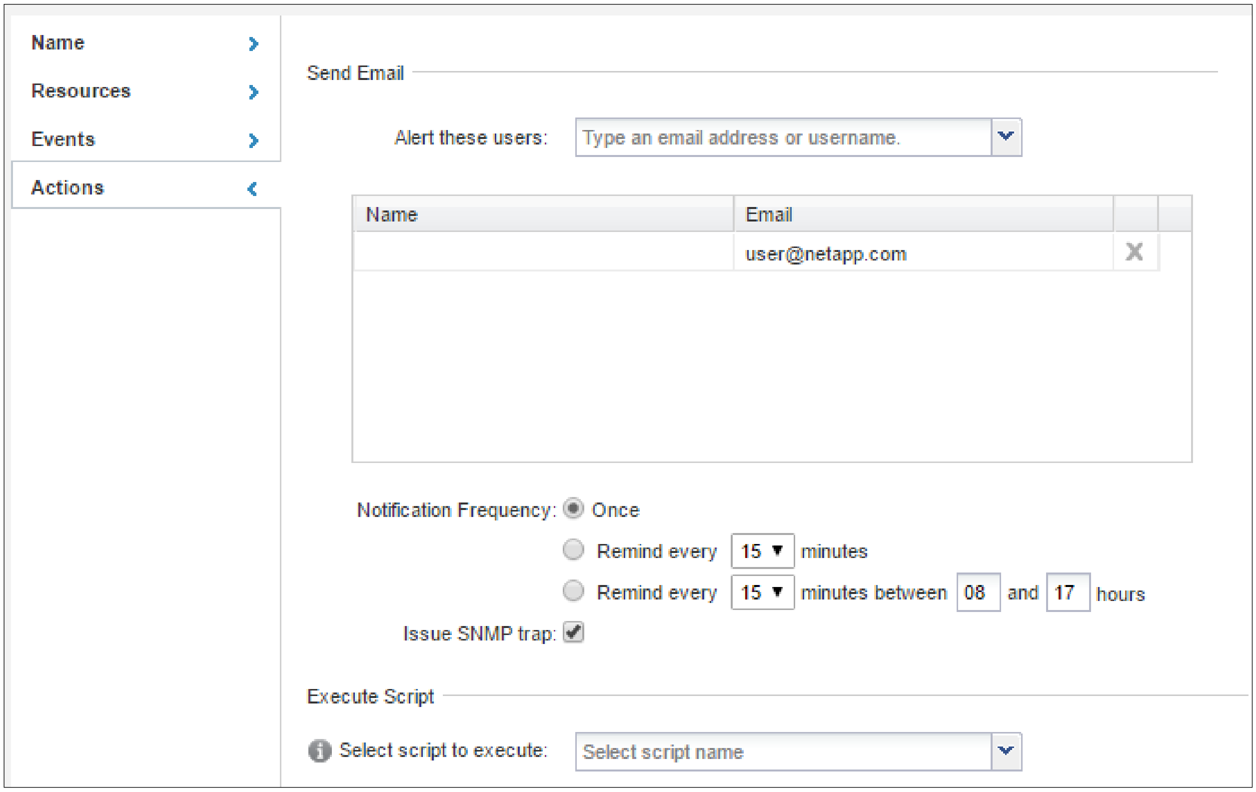
Figure 88) Volume capacity events.



When an event is triggered, the alert mechanism in Active IQ Unified Manager can do the following:

- Send an email to a user, a list of users, and a distribution list
- Trigger an SNMP trap
- Send reminders
- Execute scripts (such as an automated `volume grow` script)

Figure 89) Alert actions.

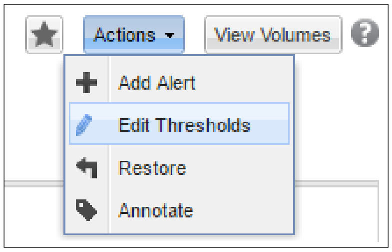


Editing Volume Thresholds in Active IQ Unified Manager

Thresholds for Volume Nearly Full and Volume Full control when an event management system event is triggered by the cluster. This control helps storage administrators stay on top of the volume capacities to prevent volumes from running out of space. In FlexGroup, this approach also involves remote allocation of files and folders, because ingest remoteness increases as a volume gets closer to full. As mentioned earlier, the Volume Nearly Full and Volume Full thresholds should be modified for a FlexGroup volume so that storage administrators are notified about potential capacity issues earlier than the defaults provide. For more information, see "Best Practice 21: Volume Space Threshold Recommendations for FlexGroup," earlier in this document.

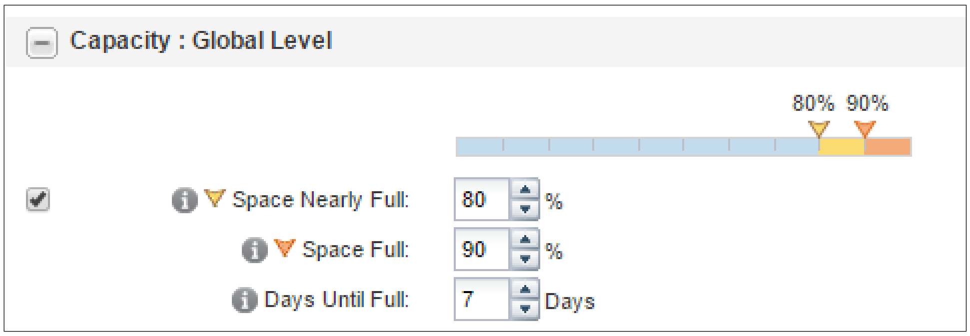
The command line provides a method to modify the thresholds, as does ONTAP System Manager. Under the Actions button of the volume detail, select Edit Thresholds to modify the volume threshold on a per-volume basis. With a FlexGroup volume, the setting is applied to the whole FlexGroup volume, and thresholds are set on each member volume individually.

Figure 90) Editing volume thresholds.



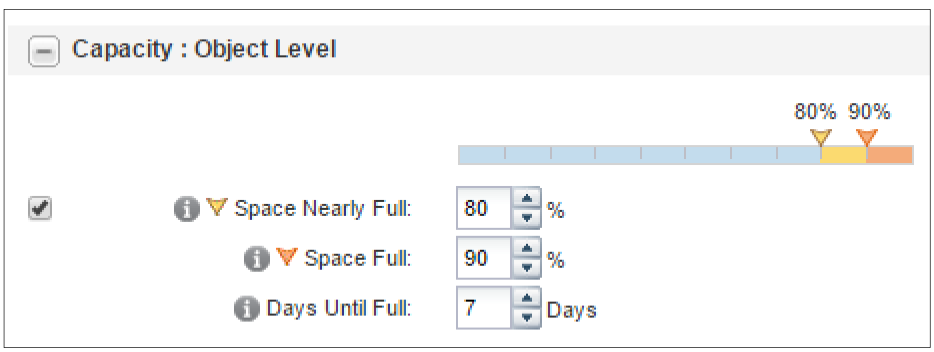
When you initially select the checkbox under Capacity: Global Level, the defaults are as shown, as in Figure 91. These defaults are unaffiliated with the ONTAP event management system volume thresholds. Rather, they are specific to Active IQ Unified Manager.

Figure 91) Capacity: Global Level.



Changing the values modifies the threshold to be an Object Level.

Figure 92) Capacity: Object Level.



On the cluster, the volume-level threshold options are unchanged.

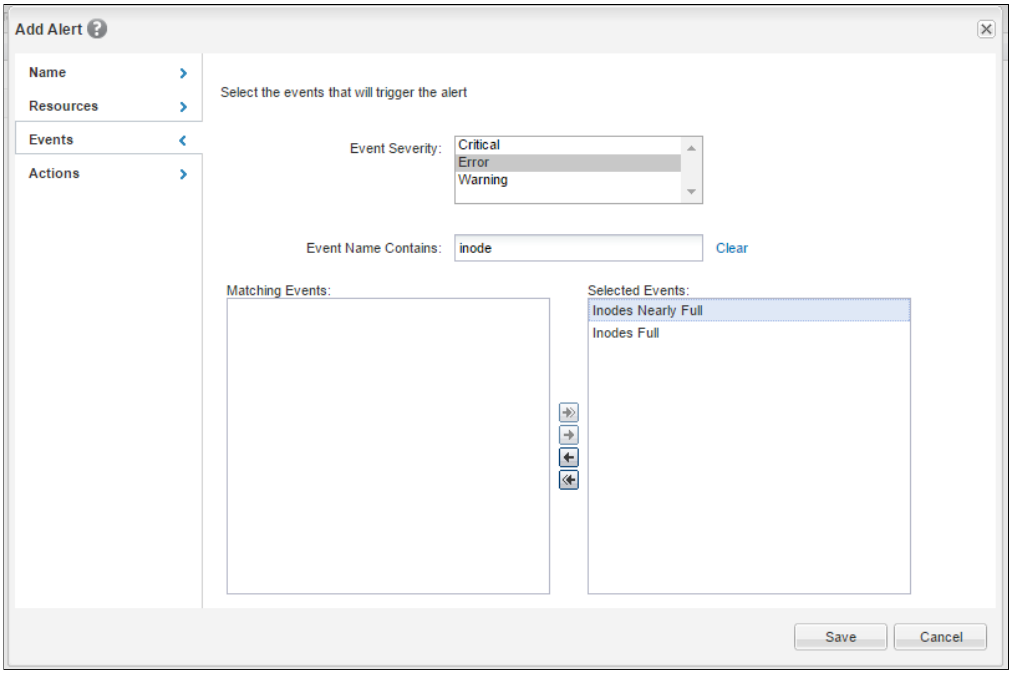
```
cluster::*> vol show -fields space-nearly-full-threshold-percent,space-full-threshold-percent -
sort-by space-nearly-full-threshold-percent -volume flexgroup
vserver volume    space-nearly-full-threshold-percent space-full-threshold-percent
------- --------- ----------------------------------- ----------------------------
SVM     flexgroup 95%                                 98%
```

You can use Active IQ Unified Manager alerting along with the cluster's event management system alerting and event destination logic, or independently of this logic.

Inode Monitoring

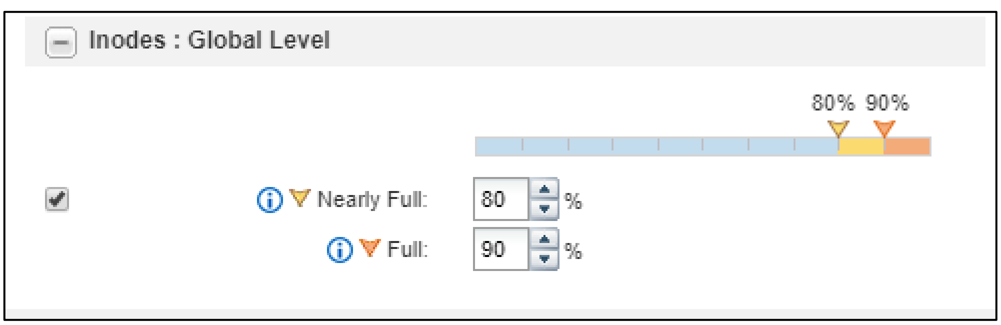
Active IQ Unified Manager also enables you to alert on inode count in FlexGroup volumes with the Inodes Nearly Full (Warning) and Inodes Full (Error) events. Alerts for inodes are configured similarly to the alerts for capacity.

Figure 93) Inode alerts.



You can also edit inode thresholds from the Edit Thresholds window for more granular control over alerting.

Figure 94) Inode thresholds.



Host-Side Capacity Considerations with Thin Provisioning

When using a FlexGroup volume, the client usually reports the available space, the used space, and so on in a way that reflects what the storage administrator has provisioned. This reporting is especially true when the volume space guarantee is set to `volume`, because ONTAP returns the expected capacities to the client.

However, when you use thin provisioning and overprovisioning for your physical storage, the client values do not reflect the expected used capacity of the FlexGroup volume. Instead, they reflect the used capacity in the physical aggregate. This approach is no different from the behavior of FlexVol volumes.

In the following example, there are three FlexGroup volumes:

- `flexgroup` has 80TB allocated and is thin provisioned across two aggregates with about 10TB available.
- `flexgroup4TB` has 4TB allocated with a space guarantee of `volume`.
- `flexgroup4TB_thin` has 4TB allocated and is thin provisioned across two aggregates with about 4TB available.

The following output shows that the cluster sees the proper used space in the volumes.

```
cluster::> vol show -fields size,used,percent-used,space-guarantee,available -vserver SVM
-volume flexgroup*,!*__0* -sort-by size
vserver volume            size available used    percent-used space-guarantee
------- ----------------- ---- --------- ------- ------------ ---------------
SVM     flexgroup4TB      4TB  3.77TB    30.65GB 5%           volume
SVM     flexgroup4TB_thin 4TB  3.80TB    457.8MB 5%           none
SVM     flexgroup         80TB 10.13TB   5.08GB  87%          none
3 entries were displayed.
```

However, the client sees the used capacity of the overprovisioned FlexGroup volume named `flexgroup` as 66TB, rather than the 5GB that is seen on the cluster. This total includes the total available size of the physical aggregate (5.05TB + 5.08TB = ~10TB) and subtracts that from the total size.

The volumes that are not overprovisioned report space normally.

```
# df -h
Filesystem                         Size  Used Avail Use% Mounted on
10.193.67.220:/flexgroup           76T   66T   11T  87% /flexgroup
10.193.67.220:/flexgroup4TB        3.9T   31G  3.8T   1% /flexgroup4TB
10.193.67.220:/flexgroup4TB_thin  3.9T  230M  3.8T   1% /flexgroup4TB_thin

cluster::*> aggr show -aggregate aggr1* -fields usedsize,availsize,percent-used,size
aggregate   availsize percent-used size   usedsize
----------- --------- ------------ ------ --------
aggr1_node1 5.05TB    36%          7.86TB 2.80TB
aggr1_node2 5.08TB    35%          7.86TB 2.78TB
2 entries were displayed.
```

The ~11TB of available space comes from the way that the Linux client calculates the space. This client does 1K blocks, so the number 10881745216 is divided into factors of 1,000. ONTAP uses factors of 1,024 to calculate space.

```
# df | grep flexg
10.193.67.220:/flexgroup          85899345920 75017600704 10881745216  88% /flexgroup
10.193.67.220:/flexgroup4TB        4080218944    32143296  4048075648   1% /flexgroup4TB
10.193.67.220:/flexgroup4TB_thin  4080218944      468736  4079750208   1% /flexgroup4TB_thin
```

Also, the `size` portion of the output considers the default 5% that is allocated for Snapshot space. That's why 80TB becomes 76TB in the preceding `df` output.

```
cluster::> vol show -fields size,percent-snapshot-space -vserver SVM -volume flexgroup*,!*__0*
-sort-by size
vserver volume            size percent-snapshot-space
------- ----------------- ---- ----------------------
SVM     flexgroup4TB      4TB  5%
SVM     flexgroup4TB_thin 4TB  5%
SVM     flexgroup         80TB 5%
3 entries were displayed.
```

When the Snapshot space allocation is reduced to 0, `df` reports a more normalized version of the actual size (but still has the strangeness of the `used` space).

```
cluster::*> vol modify -vserver SVM -volume flexgroup -percent-snapshot-space 0
[Job 2502] Job succeeded: volume modify succeeded
```

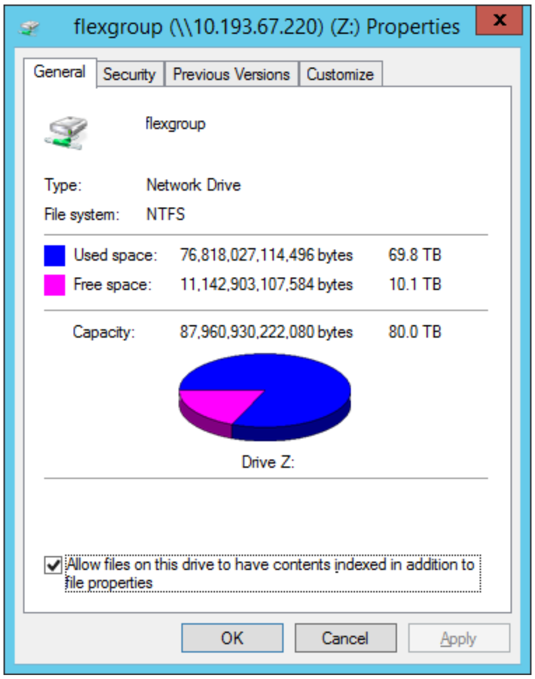


```
# df -h | grep flexgroup
Filesystem                        Size  Used Avail Use% Mounted on
10.193.67.220:/flexgroup           80T   70T   11T  88% /flexgroup
```

Windows Capacity Reporting

Windows reports in very much the same way as the Linux clients. The difference is that Windows uses a factor of 1,024, so the numbers are closer to the ONTAP values. Figure 95 shows just the overprovisioned FlexGroup volume to confirm the behavior.

Figure 95) Windows capacity view.



Viewing FlexVol Member Capacity from the ONTAP Command Line

When FlexGroup volumes are created, each member is evenly divided according to the total capacity and the number of FlexVol members. For example, in the case of an 80TB FlexGroup volume, the FlexVol members would be 10TB apiece. To view member volume capacity, use the `volume show` command at the **diag privilege** level; use `volume show -is-constituent true` or use the `volume show-space` command at the **admin privilege** level. You can find examples in previous sections of this document, such as the section "Effect of Overprovisioning or Thin Provisioning in a FlexGroup Volume" or in the "Command Examples" section at the end of this document.

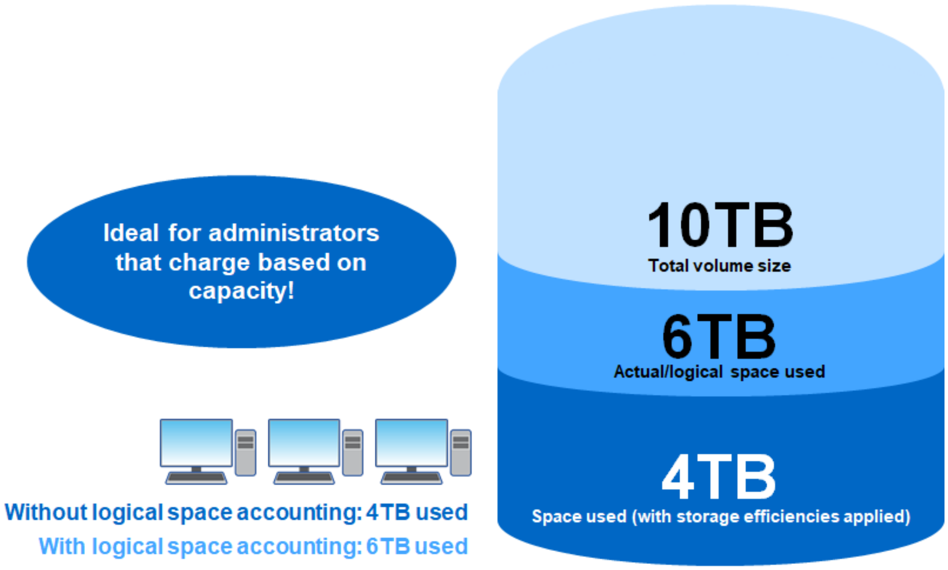
Viewing FlexVol member capacity is useful when you're trying to determine the true available space in a FlexGroup volume. When a FlexGroup volume reports total available space, it considers the total available space on all member volumes. However, when an individual member volume fills to capacity, the entire FlexGroup volume reports as out of space, even if other member volumes show available space. To mitigate this scenario, the FlexGroup ingest algorithms attempt to direct traffic away from a volume that becomes more heavily used than other volumes. For information about how the ingest heuristics operate, see TR-4557.

Logical Space Accounting

Logical space accounting was introduced in ONTAP 9.4. It enables storage administrators to mask storage efficiency savings so that end users avoid overallocating their designated storage quotas.

For example, if a user writes 6TB to a 10TB volume and storage efficiencies save 2TB, logical space accounting can control whether the user sees 6TB or 4TB.

Figure 96) How logical space accounting works.



ONTAP 9.5 enhanced this feature and added quota enforcement support to give more control to storage administrators by preventing new writes according to the logical space.

Currently, FlexGroup volumes do not support this functionality; it is only available for FlexVol volumes.

FabricPool

In ONTAP 9.2, the ability to automatically tier cold data blocks on SSD aggregates to the cloud or on-premises Amazon Simple Storage Service (Amazon S3) object storage was added for FlexVol volumes. This functionality allows storage administrators to preserve more costly SSDs for active workloads, whereas cold or unused data is moved to more cost-effective capacity tiers. This feature is known as FabricPool. You can learn more about the feature in TR-4598: FabricPool Best Practices.

ONTAP 9.5 introduced support for FabricPool for FlexGroup volumes. There are no special considerations to make for FlexGroup volumes; the same FlexVol considerations apply.

Managing Quotas with FlexGroup

FlexGroup originally supported user and group quotas, and tree quotas starting in ONTAP 9.3, with reporting functionality only. Starting in ONTAP 9.5, enforcement of quotas (that is, setting limits) and qtree quotas is supported. ONTAP 9.4 also added support for FPolicy, which can provide quota enforcement from third-party vendors, such as DefendX (formerly NTP).

User and Group Quota Considerations

To implement user or group quotas, the cluster must be able to resolve the specified user or group. This requirement means that the user or group must exist locally on the SVM or within a resolvable name service server, such as Active Directory, LDAP, or NIS. If a user or group cannot be found by the SVM, then the quota rule is not created. If a user or group quota fails to create because of an invalid user, the command line issues this error:

```
Error: command failed: User name user not found. Reason: SecD Error: object not found.
```


ONTAP System Manager delivers a similar message. Use the `event log show` command to investigate the issue further. For more information about configuring name services, see [TR-4073: Secure Unified Authentication] and [TR-4379: Name Services Best Practices Guide].

Creating a User or Group Reporting Quota with the Command Line

To create a user or group reporting quota with the command line for a specific user or group, use the following command at the admin privilege level:

```
cluster::> quota policy rule create -vserver SVM1 -policy-name default -volume flexgroup -type
[user|group] -target [username or groupname] -qtree ""
```

To create a user or group reporting quota with the command line for all users or groups, use the following command at the admin privilege level. The target is provided as an asterisk to indicate `all`:

```
cluster::> quota policy rule create -vserver SVM1 -policy-name default -volume flexgroup -type
[user|group] -target * -qtree ""
```

Before ONTAP 9.5, quota enforcement was unsupported for use with FlexGroup volumes. As a result, you could not set limits for files or disk space usage. ONTAP 9.5 lets you set hard limits for files (`-file limit`) and capacity (`-disk-limit`) with quotas.

This example shows the `quota report` command with FlexGroup volumes and quota enforcement:

```
cluster
::*> quota report -vserver DEMO
Vserver: DEMO

                                    ----Disk----  ----Files-----   Quota
Volume   Tree      Type    ID        Used  Limit    Used   Limit    Specifier
-------  --------- ------  -------  -----  -----  ------  ------    ---------
flexgroup_local
         qtree     tree    1
                                   1.01GB    1GB       5      10    qtree
flexgroup               user    student1, NTAP\student1
                                      4KB    1GB      10      10    student1
```

Creating a Tree Reporting Quota from the Command Line

To create a tree reporting quota with the command line for a specific user or group, use the following command at the admin privilege level:

```
cluster::> quota policy rule create -vserver DEMO -policy-name tree -volume flexgroup_local -type
tree -target qtree
```

To enable quotas, use `quota on` or `quota resize`.

```
cluster::> quota on -vserver DEMO -volume flexgroup_local
[Job 9152] Job is queued: "quota on" performed for quota policy "tree" on volume
"flexgroup_local" in Vserver "DEMO".

cluster::> quota resize -vserver DEMO -volume flexgroup_local
[Job 9153] Job is queued: "quota resize" performed for quota policy "tree" on volume
"flexgroup_local" in Vserver "DEMO".

cluster::> quota show -vserver DEMO -volume flexgroup_local

              Vserver Name: DEMO
               Volume Name: flexgroup_local
               Quota State: on
               Scan Status: -
          Logging Messages: -
          Logging Interval: -
          Sub Quota Status: none
  Last Quota Error Message: -
```



```
Collection of Quota Errors: -
       User Quota enforced: false
      Group Quota enforced: false
       Tree Quota enforced: true
```

The following example shows a `quota report` command on a FlexGroup volume with a tree quota specified:

```
cluster::> quota report -vserver DEMO -volume flexgroup_local
Vserver: DEMO

                                    ----Disk----  ----Files-----   Quota
Volume   Tree      Type    ID        Used  Limit    Used   Limit   Specifier
-------  --------  ------  -------  -----  -----  ------  ------   ---------
flexgroup_local
         qtree     tree    1           0B      -       1       -   qtree
```

Files used, and disk space used, are monitored and increment as new files are created:

```
cluster::> quota report -vserver DEMO -volume flexgroup_local
Vserver: DEMO

                                    ----Disk----  ----Files-----   Quota
Volume   Tree      Type    ID        Used  Limit    Used   Limit   Specifier
-------  --------  ------  -------  -----  -----  ------  ------   ---------
flexgroup_local
         qtree     tree    1
                                  13.77MB      -       4       -   qtree
```

Quota Enforcement Example

When quota enforcement is enabled on a qtree or for a user, ONTAP disallows new file creations or writes after a quota is exceeded. In addition, an event management system message is logged at the DEBUG severity level to notify storage administrators of the quota violation. You can configure these messages so that the system forwards them as SNMP traps or as syslog messages.

In this example, a quota has been set with a hard limit of 1GB and 10 files.

```
cluster::*> quota policy rule show -vserver DEMO

Vserver: DEMO              Policy: tree              Volume: flexgroup_local

                                                Soft             Soft
                          User         Disk     Disk   Files    Files
Type   Target    Qtree   Mapping     Limit    Limit   Limit    Limit  Threshold
-----  --------  ------- -------  --------  -------  ------  -------  ---------
tree   qtree     ""      -             1GB        -      10        -          -
```

When a user tries to copy a 1.2GB file to the qtree, ONTAP reports an “out of space” error.

```
[root@centos7 qtree]# cp /SANscreenServer-x64-7.3.1-444.msi /FGlocal/qtree/
cp: failed to close ‘/FGlocal/qtree/SANscreenServer-x64-7.3.1-444.msi’: No space left on device
```

The file is partially written, but it is unusable because it’s missing data.

```
# ls -alh
total 1.1G
drwxr-xr-x  2 root root  4.0K Jul 19 15:44 .
drwxr-xr-x 11 root root  4.0K Jun 28 15:10 ..
-rw-r--r--  1 root root   14M Dec 12  2017 First Draft TTDD Slide Deck on ONTAP 9.3 - Parisi.pptx
-rw-r--r--  1 root root     0 Dec 12  2017 newfile1
-rw-r--r--  1 root root     0 Dec 12  2017 newfile2
-rw-r--r--  1 root root 1021M Jul 19  2018 SANscreenServer-x64-7.3.1-444.msi
```

ONTAP then reports the quota as exceeded.


```
cluster::*> quota report -vserver DEMO
Vserver: DEMO

                                    ----Disk----  ----Files-----   Quota
Volume   Tree      Type    ID        Used  Limit    Used   Limit   Specifier
-------  --------  ------  -------  -----  -----  ------  ------   ---------
flexgroup_local
         qtree     tree    1
                                   1.01GB    1GB       5      10   qtree
```

The same behavior occurs for file count limits. In this example, the file count limit is 10 and the qtree already has 5 files. An extra 5 files meet our limit.

```
[root@centos7 /]# su student1
sh-4.2$ cd ~
sh-4.2$ pwd
/home/student1
sh-4.2$ touch file1
sh-4.2$ touch file2
sh-4.2$ touch file3
sh-4.2$ touch file4
sh-4.2$ touch file5
touch: cannot touch 'file5': Disk quota exceeded

cluster::*> quota report -vserver DEMO
Vserver: DEMO

                                    ----Disk----  ----Files-----   Quota
Volume   Tree      Type    ID        Used  Limit    Used   Limit   Specifier
-------  --------  ------  -------  -----  -----  ------  ------   ---------
flexgroup_local
         qtree     tree    1
                                   1.01GB    1GB       5      10   qtree
home               user    student1, NTAP\student1
                                      4KB    1GB      10      10   student1
2 entries were displayed.
```

From the event logs, we can see the quota violations.

```
cluster::*> event log show -message-name quota.exceeded
Time                Node             Severity      Event
------------------- ---------------- ------------- ---------------------------
7/19/2018 16:27:54  node02
                                     DEBUG         quota.exceeded: ltype="hard", volname="home",
app="", volident="@vserver:7e3cc08e-d9b3-11e6-85e2-00a0986b1210", limit_item="file",
limit_value="10", user="uid=1301", qtree="treeid=1", vfiler=""
7/19/2018 15:45:02  node01
                                     DEBUG         quota.exceeded: ltype="hard",
volname="flexgroup_local", app="", volident="@vserver:7e3cc08e-d9b3-11e6-85e2-00a0986b1210",
limit_item="disk", limit_value="1048576", user="", qtree="treeid=1", vfiler=""
```

Performance Effect of Using Quotas

Performance effect is always a concern with a new feature. Therefore, we ran a standard NAS benchmark against FlexGroup volumes in ONTAP 9.5 with and without quotas enabled. We concluded that the performance impact for enabling quotas on a FlexGroup volume is negligible.

Figure 97) ONTAP 9.5 performance (operations/sec)—quotas on and off.

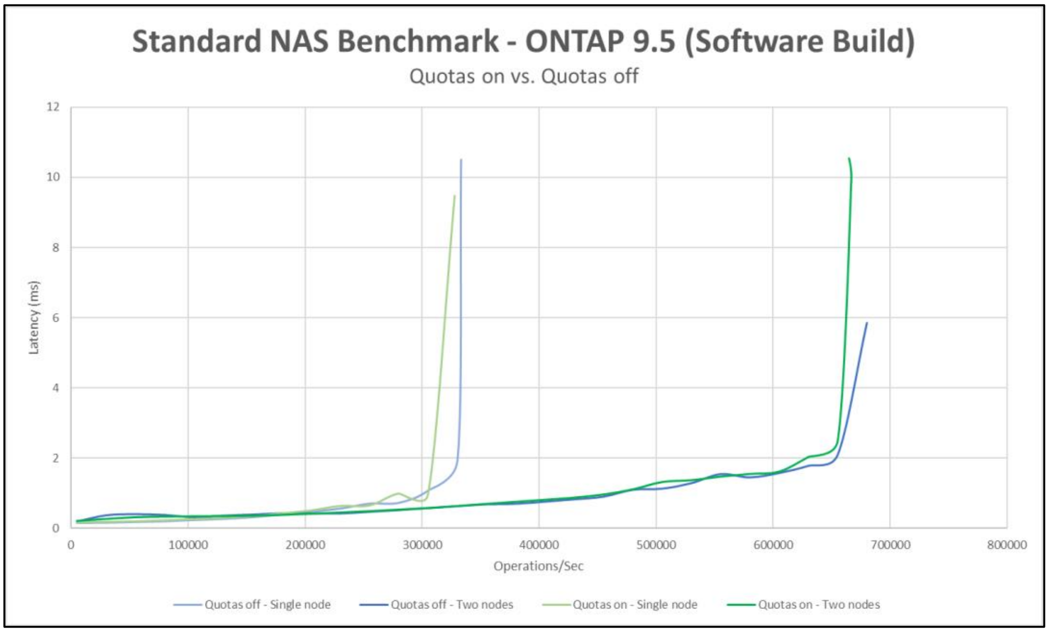
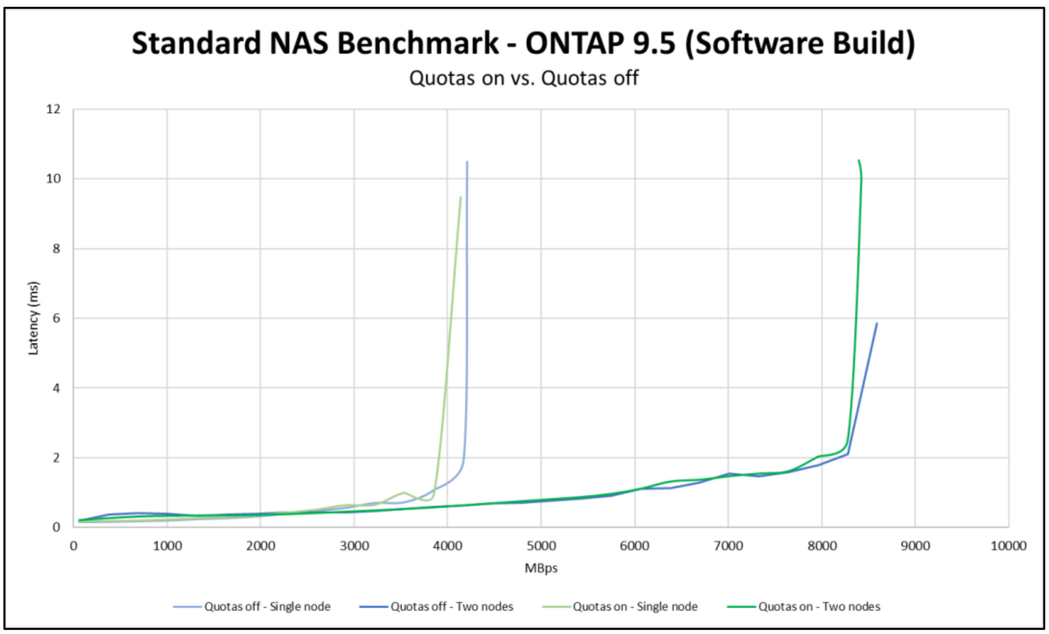


Figure 98) ONTAP 9.5 performance (MBps)—quotas on and off.



User-Mapping Considerations

User mapping in multiprotocol environments (SMB and NFS) for quotas occurs at the member volume level. Eventually, all member volumes agree on the user mapping. However, sometimes there might be a discrepancy, such as when user mapping fails or times out when doing a name mapping that succeeded on another member. This means that at least one member considers the user to be part of a user-mapped pair, and at least one other member considers it to be a discrete record.

At worst, enforcement of the quota rules can be inconsistent until the issue is resolved. For instance, a user might be able to overrun a quota limit.

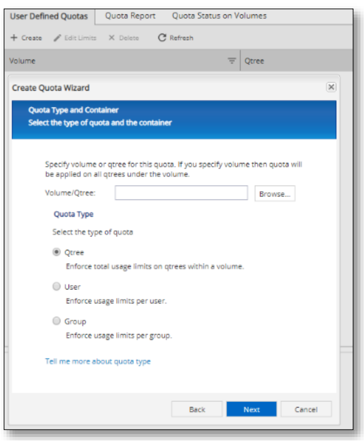
An event management system message is sent for these issues.

```
cluster::*> event route show -message-name fg.quota.usermapping.result -instance

                                      Message Name: fg.quota.usermapping.result
                                          Severity: NOTICE
                                 Corrective Action: (NONE)
                                       Description: This message occurs when the quota mapper
decides whether to map the Windows quota record and the UNIX quota record of a user into a single
multiuser record.
```

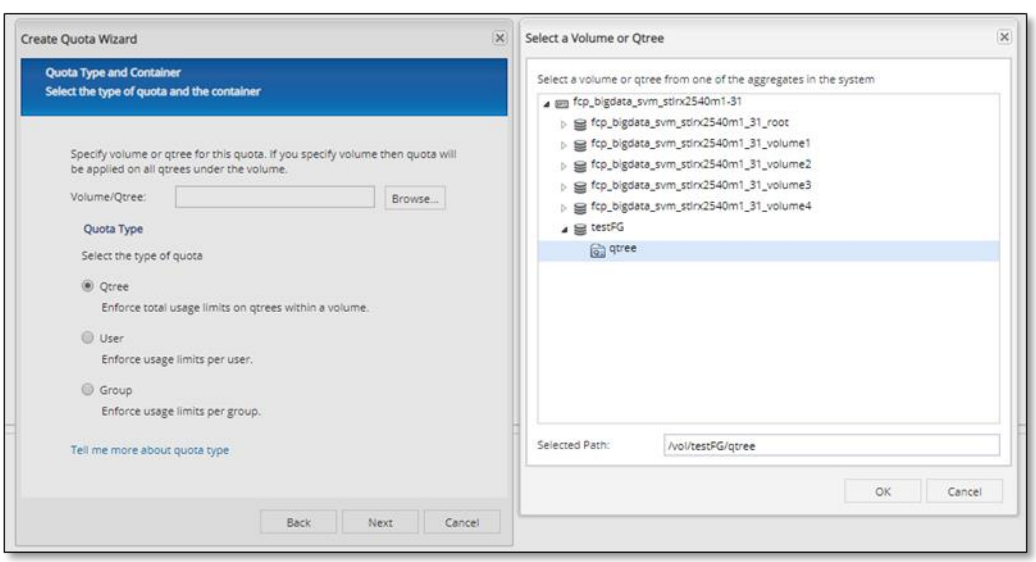
If this issue occurs, contact NetApp technical support for remediation steps.

Creating and Managing Quotas in System Manager: Classic View

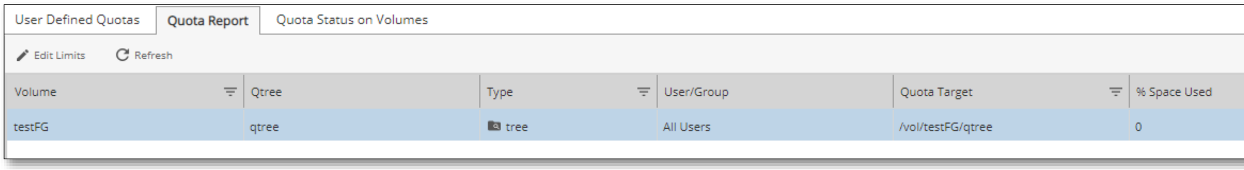
Starting in ONTAP 9.4, you can use System Manager to create, modify, and view quotas in FlexGroup volumes. Navigate to the Storage > Quotas menu option and click +Create on the User-Defined Quotas tab. In the Create Quota wizard, select the type of quota you want to create:



If you are creating a qtree quota, navigate to the desired object.

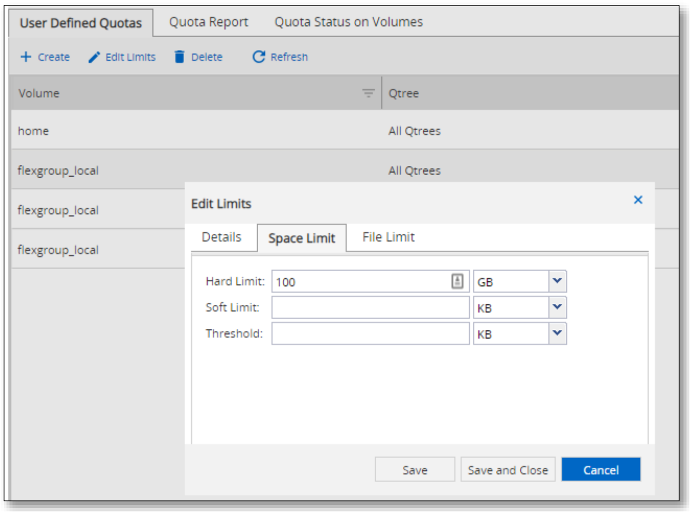


Click Next and review your settings until the wizard completes. You now have a quota policy. To view the policy, select the Quota Report tab:



To create quota rules that allow you to set file and capacity values, click Edit Limits. This option is available only in ONTAP 9.5 and later for FlexGroup volumes. In ONTAP 9.4, to set monitoring limits, use the command line.

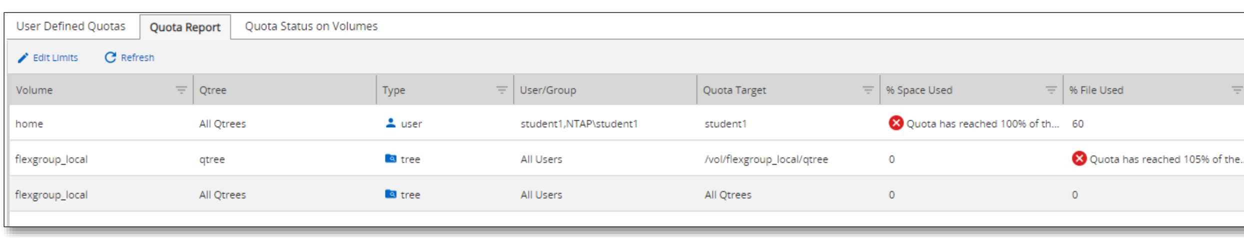
In ONTAP 9.5, you can edit the limits in the System Manager UI.



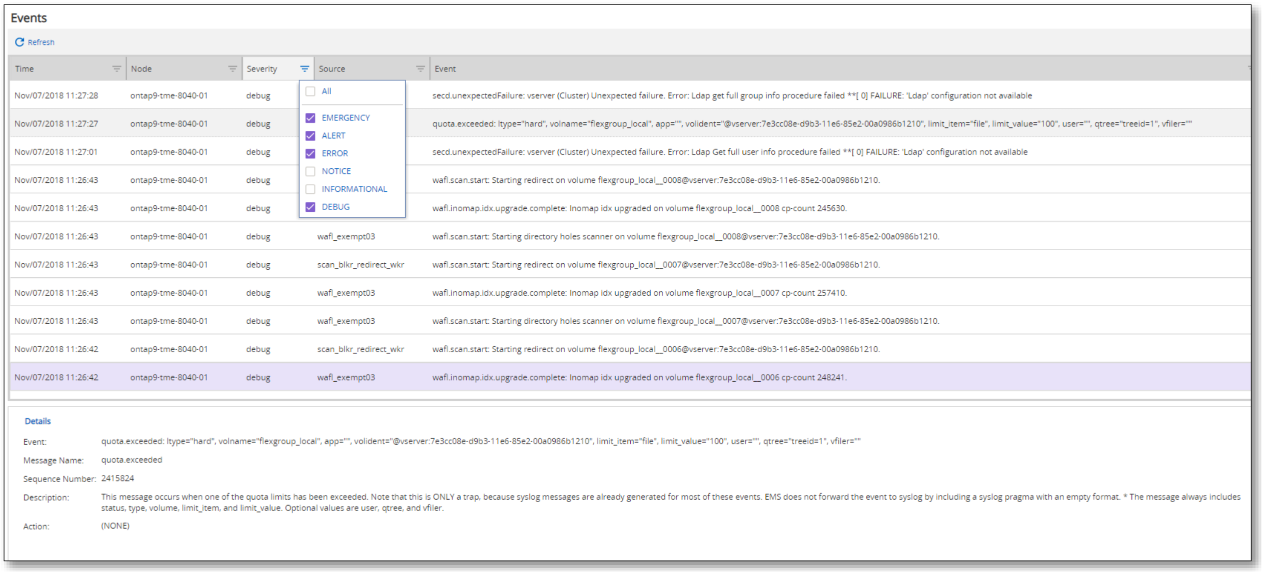
When a quota rule is exceeded, a client attempting to exceed the limit is informed that there is no more space on the file system.

```
OSError: [Errno 28] No space left on device
```

The System Manager UI displays the following quota report:



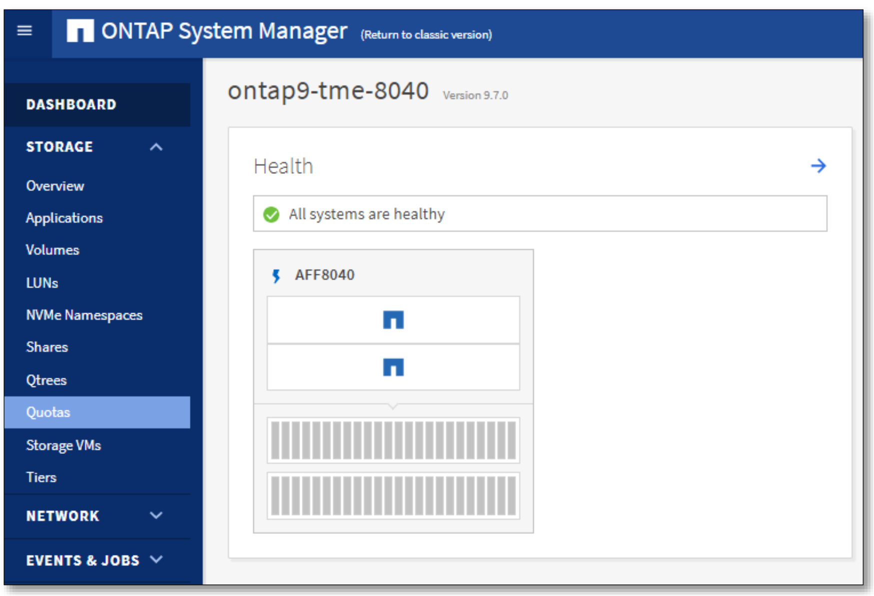
An event message is also generated. The Quota Exceeded error is a DEBUG-level event, so filter by DEBUG in System Manager events:



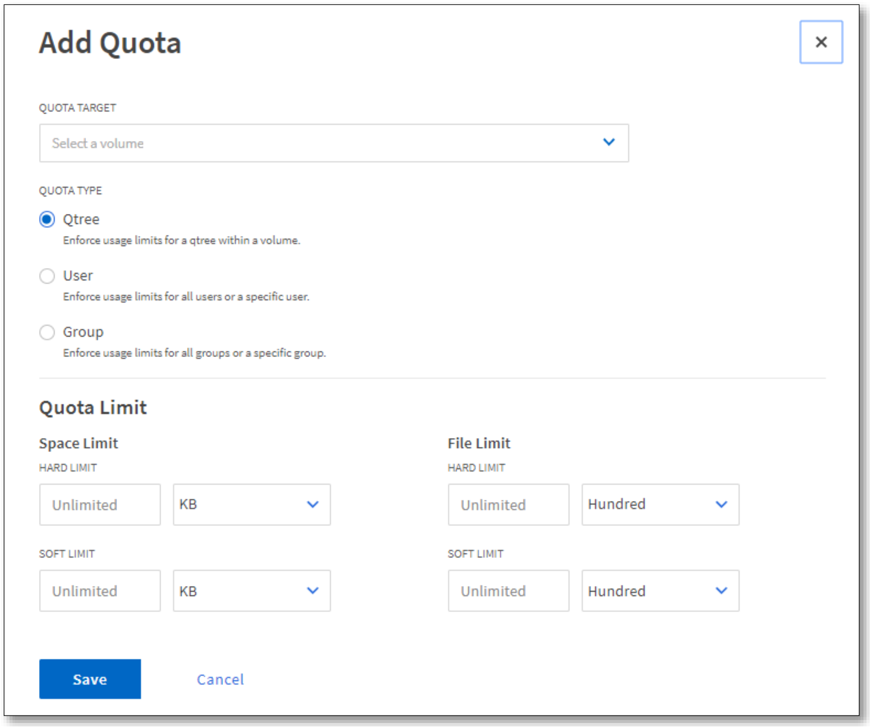
Creating and Managing Quotas in System Manager: ONTAP 9.7 and later

ONTAP 9.7 and later versions offer a new System Manager view to help simplify storage operations. Although the general concept of quota creation is the same, the GUI location changes in the new view.

Here is where to find quotas:



After you select Quotas, click Add. You'll see a single screen, rather than a multistep wizard:



Tree, user, and group quotas all have different options you can apply.

Tree Quota Considerations

SVMs in ONTAP can have a maximum of five quota policies, but only one policy can be active. To see the active policy in an SVM, use the following command:

```
cluster::> vserver show -vserver DEMO -fields quota-policy
vserver quota-policy
------- ------------
DEMO    default
```

The default policy is adequate in most cases and does not need to be changed. When `quota on` is issued, the active policy is used—not the policy that was assigned to a volume. Therefore, it's possible to get into a situation where you think you have applied a quota and rules to a volume, but `quota on` fails.

The following example applies a quota policy to a volume:

```
cluster::*> quota policy show -vserver DEMO -policy-name tree

      Vserver: DEMO
  Policy Name: tree
Last Modified: 10/19/2017 11:25:20
    Policy ID: 42949672962

cluster::*> quota policy rule show -vserver DEMO -policy-name tree -instance

                 Vserver: DEMO
             Policy Name: tree
             Volume Name: flexgroup_local
                    Type: tree
                  Target: tree1
              Qtree Name: ""
            User Mapping: -
              Disk Limit: -
             Files Limit: -
Threshold for Disk Limit: -
```



```
        Soft Disk Limit: -
       Soft Files Limit: -
```

Turning on quotas produces an error because the SVM has `default` assigned for quotas and does not contain any rules.

```
cluster::*> quota on -vserver DEMO -volume flexgroup_local -foreground true

Error: command failed: No valid quota rules found in quota policy default for volume
flexgroup_local in Vserver DEMO.
```

When you add a rule to `default`, the `quota on` command works, but the SVM does not use the new tree policy.

```
cluster::*> quota policy rule create -vserver DEMO -policy-name default -volume flexgroup_local -
type tree -target ""

cluster::*> quota on -vserver DEMO -volume flexgroup_local -foreground true
[Job 8063] Job succeeded: Successful

cluster::*> vserver show -vserver DEMO -fields quota-policy
vserver quota-policy
------- ------------
DEMO    default
```

To use the necessary policy, you must modify the SVM and then turn quotas off and back on.

```
cluster::*> vserver modify -vserver DEMO -quota-policy tree

cluster::*> quota off -vserver DEMO *

cluster::*> quota policy rule delete -vserver DEMO -policy-name default *
1 entry was deleted.

cluster::*> quota on -vserver DEMO -volume flexgroup_local -foreground true
[Job 8084] Job succeeded: Successful
```

This behavior is not unique to FlexGroup volumes; this would happen with FlexVol volumes as well.

10.3 Monitoring FlexGroup Performance

FlexGroup performance can be monitored in many of the same ways that a normal FlexVol volume's performance can be monitored. The same concepts of CPU utilization, disk saturation, NVRAM bottlenecks, and other NetApp WAFL-related performance characteristics apply. Also, NAS performance monitoring doesn't change. You still use the basic CIFS/SMB and NFS statistics that you always have.

The main difference with monitoring FlexGroup performance is that you must consider multiple member FlexVol constituent volumes, and the notion of remote placement of files and folders. These elements add another layer to consider when you want to monitor and isolate performance issues.

Monitoring Performance from the Command Line

From the command line, you have several ways to view performance statistics.

Real-Time Performance Monitoring

To monitor system performance in real time, use the `statistics show-periodic` command.

```
cluster::*> statistics show-periodic ?
  [[-object] <text>]           *Object
  [ -instance <text> ]         *Instance
  [ -counter <text> ]          *Counter
  [ -preset <text> ]           *Preset
  [ -node <nodename> ]         *Node
  [ -vserver <vserver name> ]  *Vserver
```



```
  [ -interval <integer> ]       *Interval in Seconds (default: 2)
  [ -iterations <integer> ]     *Number of Iterations (default: 0)
  [ -summary {true|false} ]     *Print Summary (default: true)
  [ -filter <text> ]            *Filter Data
```

This command provides an up-to-date glimpse into system performance. Leaving the default values alone gives you a cluster-wide view. Specifying an SVM gives you a more granular look, but mainly at the counters that would be specific to an SVM, such as NAS counters, rather than to CPU or disk. When you use SVM-specific statistics, defining the counters that are provided for the object helps reduce the noise on the CLI. You can also get real-time FlexGroup statistics for the ratios of local to remote top-level directories (`tld`), high-level directories (`hld`), regular directories, and files.

For examples of these commands, see "Command Examples."

The FlexGroup statistics also can show various other information and can be gathered over a period of time if you initiate a `statistics start -object flexgroup` command. This command collects statistics over time that can be captured in iterations through an automated tool such as Perfstat or perfarchives.

```
cluster::*> statistics start -object flexgroup
Statistics collection is being started for sample-id: sample_69197
```

Use the following to view the statistics:

```
cluster::*> statistics show -object flexgroup -instance 0

Object: flexgroup
Instance: 0
Start-time: 11/30/2016 16:44:42
End-time: 11/30/2016 17:42:57
Elapsed-time: 3495s
Scope: cluster-01

    Counter                                                     Value
    -------------------------------- --------------------------------
    cat1_tld_remote                                                 2
    cat2_hld_local                                                180
    cat2_hld_remote                                              1292
    cat3_dir_local                                             146804
    cat3_dir_remote                                               283
    cat4_fil_local                                             734252
    cat4_fil_remote                                              1124
    groupstate_analyze                                          12232
    groupstate_update                                           86242
    instance_name                                                   0
    node_name                                       cluster-01
    process_name                                                    -
    refreshclient_create                                         5241
    refreshclient_delete                                         5241
    refreshserver_create                                         5244
    refreshserver_delete                                         5244
```

The statistics capture gives a nice summary of the percentages of remote file and directory placement in the FlexGroup volume when it spans multiple nodes. (In the following example, the values are 14% remote directories and 1% remote files.).

```
    remote_dirs                                                    14
    remote_files                                                    1
```

Qtree Statistics

Starting in ONTAP 9.5, qtree statistics were made available for FlexGroup volumes. These statistics provide granular performance information about FlexGroup volumes and their qtrees. The following example shows a statistics capture for a FlexGroup volume running a large NFS workload.


```
cluster::> statistics qtree show -interval 5  -iterations 1 -max 25 -vserver DEMO -volume
flexgroup_local

cluster : 11/7/2018 15:19:15

                                                NFS CIFS Internal *Total
                Qtree Vserver          Volume   Ops  Ops      Ops    Ops
--------------------- ------- --------------- ----- ---- -------- ------
DEMO:flexgroup_local/    DEMO flexgroup_local 22396    0        0  22396
DEMO:flexgroup_local/qtree
                         DEMO flexgroup_local     0    0        0
```

Protocol Statistics

It's also possible to get a glimpse of how individual NAS protocols are influencing performance. Simply use the `statistics start` command to include NFS or SMB performance counters in the capture. You'll get more options with **diag privileges**.

```
cluster::*> statistics start -object nfs
    nfs_credstore                       nfs_exports_access_cache
    nfs_exports_cache                   nfs_exports_match
    nfs_file_session_cache              nfs_file_session_cache:constituent
    nfs_generic                         nfs_idle_conn
    nfs_idle_total_conn                 nfs_qtree_export
    nfs_server_byname                   nfserr
    nfsv3                               nfsv3:constituent
    nfsv3:cpu                           nfsv3:node
    nfsv4                               nfsv4:constituent
    nfsv4:cpu                           nfsv4:node
    nfsv4_1                             nfsv4_1:constituent
    nfsv4_1:cpu                         nfsv4_1:node
    nfsv4_1_diag                        nfsv4_1_error
    nfsv4_diag                          nfsv4_error
    nfsv4_spinnp_errors

cluster::*> statistics start -object smb
    smb1          smb1:node     smb1:vserver  smb1_ctx      smb1_ctx:node
    smb2          smb2:node     smb2:vserver  smb2_ctx      smb2_ctx:node

cluster::*> statistics start -object cifs
    cifs                         cifs:node
    cifs:vserver                 cifs_cap
    cifs_cap:constituent         cifs_client
    cifs_client:constituent      cifs_ctx
    cifs_ctx:node                cifs_shadowcopy
    cifs_unsupp_ioctl            cifs_unsupp_ioctl:constituent
    cifs_watch
```

flexgroup show

During FlexGroup I/O, you can also view the member constituent usage and balance through the nodeshell command `flexgroup show`. The command also provides other information that can be useful. Be sure to capture this command output if you run into a FlexGroup issue and need to open a support case.

```
cluster::*> node run * flexgroup show flexgroup4TB_thin
* next snapshot cleanup due in 8760 msec
* next refresh message due in 762 msec (last to member 0x80F03776)
* spinnp version negotiated as 8.5, capability 0x57F
* Ref count is 20
* ShouldEnforceQuotas false
* IsAnyMemberInNvfailedState false
* reaction -3.0, workload +0.0, activity level 0, cv 0%
Idx   Member L            Used      Avail Urgc Targ          Probabilities   D-Ingest   D-Alloc
F-Ingest   F-Alloc
--- -------- - --------------- ---------- ---- ----  ---------------------  --------- ---------
--------- ---------
```



```
  1     1723 L      15715   0%  127499546   0%  12%  [100% 100%  87%   0%]     0+   0    0+   0
325+  22  325+  22
  2     1724 L      15789   0%  127499546   0%  12%  [100% 100%  87%   0%]     0+   0    0+   0
327+  18  327+  18
  3     1725 L      15592   0%  127499546   0%  12%  [100% 100%  87%   0%]     0+   0    0+   0
320+  23  320+  23
  4     1726 L      15714   0%  127499546   0%  12%  [100% 100%  87%   0%]     0+   0    0+   0
321+  20  321+  20
  5     1727 L      15900   0%  127499546   0%  12%  [100% 100%  87%   0%]     0+   0    0+   0
320+  19  320+  19
  6     1728 L      15844   0%  127499546   0%  12%  [100% 100%  87%   0%]     0+   0    0+   0
318+  22  318+  22
  7     1729 L      15428   0%  127499546   0%  12%  [100% 100%  87%   0%]     0+   0    0+   0
318+  23  318+  23
  8     1730 L      15741   0%  127499546   0%  12%  [100% 100%  87%   0%]     0+   0    0+   0
329+  19  329+  19
```

Output Breakdown for the flexgroup show Command

The `flexgroup show` command has a series of values that might not be intuitive at first glance. Table 15 describes those values and how to interpret them. [TR-4557](TR-4557) covers some of these terms and concepts in more detail.

Table 15) `flexgroup show` output column definitions.

Column	Definition
`Idx`	Index number of the member volume.
`Member`	DSID of the FlexGroup member.
`L`	Local or remote to the node.
`Used`	Number and overall percentage of 4K blocks used.
`Urgc`	Urgency: Probability of a file or directory creation being allocated to a remote member volume to avoid premature `ENOSPC` in a member volume. This value increases according to how close to 100% used a volume's capacity is.
`Targ`	Target: Percentage of what new content should be placed on a member volume as related to its peers. The total summation of all target percentages equal ~100%.
`Probabilities`	The likelihood that a member volume is avoided for use. This number increases according to how full a member volume becomes in relation to other member volumes (tolerance).
`D-Ingest` and `D-Alloc`	Directory ingest and directory allocation, respectively; how many directories have been allocated to a local member volume.
`F-Ingest` and `F-Alloc`	File ingest and file allocation, respectively; how many files have been allocated to a local member volume.

You should run the `flexgroup show` command during a period of I/O activity on a FlexGroup volume. This command gives the following useful information:

- How evenly the traffic is distributed across members
- How evenly distributed the space is on members
- How likely a member volume is to be used for ingesting
- The ratio of directory to file creation in a workload
- The member volume's node locality

Note: The output of `flexgroup show` is also captured in Perfstat, under `wafl_susp -w`.

Monitoring Performance by Using Perfstat (ONTAP 9.4 and Earlier)

Perfstat is a tool in the NetApp Support site ToolChest that can capture real-time performance statistics for benchmarking or to troubleshoot current performance issues. These statistics are captured at specified intervals and are crucial for NetApp Support to resolve performance cases.

Perfstat in ONTAP 9.1 and later supports the capture of FlexGroup statistics. When Perfstat runs, it collects data in raw text format and can be reviewed with any text-editing software. Perfstat even has a Windows GUI version that you can download.

When Perfstat finishes running, the tool zips the contents up for submittal to NetApp Support. The tool also creates a folder with the output files that can be read with a text editor. However, this version is in plaintext and isn't easy to read. If you are a NetApp original equipment manufacturer (OEM) partner or an internal employee, you can use the LaTeX tool to view these files.

For more information, see the KB article on how to use Perfstat.

Note: Perfstat is deprecated as of ONTAP 9.5 and later. For later ONTAP versions, use the performance archiver.

Performance Archiver

Starting in ONTAP 9.5, performance data is captured for support issues through the performance archiver, which runs by default in ONTAP. These statistics are captured at all times and can be sent to NetApp using the `autosupport invoke-performance-archive` command. This command can specify date ranges, so when a problem is found, only pertinent data needs to be sent. Additionally, the performance archiver helps remove the need to try to "catch the problem in the act," because the data has already been gathered during an issue.

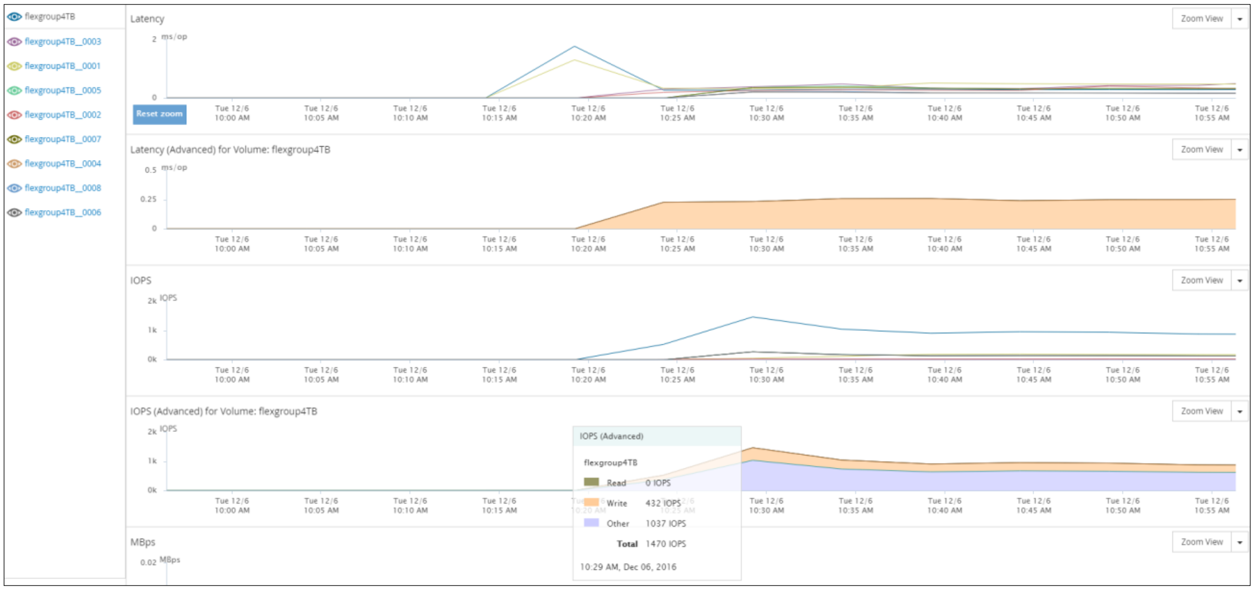
Monitoring Performance (Active IQ Performance Manager or Unified Manager)

A more palatable and widely available tool for monitoring the performance of a FlexGroup volume is NetApp Active IQ Performance Manager. This tool is available as a free .ova file or as a Linux installation from the NetApp download page of the Support site.

Active IQ Performance Manager offers both real-time and historical performance information for a cluster and can be integrated with Active IQ Unified Manager to provide a single management point. Active IQ Performance Manager can give granular performance views for the entire FlexGroup volume or for individual member constituent FlexVol volumes.

Figure 99 is a capture of a simple file creation script on a single Linux VM, so the performance benefits of FlexGroup are not seen here. However, the figure does provide a sense of what Performance Manager can deliver.

Figure 99) Active IQ Performance Manager graphs.



10.4 Failure Scenarios

This section covers some failure scenarios and how a FlexGroup reacts.

Storage Failovers

FlexGroup volumes are built on FlexVol volumes, so storage failover operations perform the same as for a FlexVol volume. Takeovers have no noticeable disruption. Nondisruptive upgrades, head swaps, rolling upgrades, and so on all perform normally. Givebacks of stateful protocols, such as SMB, are slightly disruptive, because of the transfer of locking states.

One caveat is that if an aggregate is not “at home” (not on the node that owns it, such as in a partial giveback state or if an aggregate has been relocated), FlexGroup volumes cannot be created until the aggregates are at home.

Network Failures

If a network connection that is accessing a FlexGroup volume happens to have an interruption or failure, the behavior for a FlexGroup volume mirrors that of a FlexVol volume. The cluster attempts to migrate the data LIF to a port or node that can access the network successfully. Clients experience a brief disruption, as expected with network issues.

Snapshot Failures

If a FlexGroup Snapshot copy fails, ONTAP considers that Snapshot copy to be “partial” and invalidates it for NetApp SnapRestore operations. The Snapshot set is cleaned up by ONTAP.

Hardware Failures

Disk failures operate the same as with a FlexVol; ONTAP fails the disk and selects a spare to use in a rebuild operation. If more disks in an aggregate fail than are allowed in a RAID configuration, the aggregate is considered offline and the member volumes that live on the offline aggregate are inaccessible. In a FlexGroup volume that spans multiple aggregates, access to other member volumes are fenced off to prevent data inconsistencies until the hardware issue has been addressed and the other member volumes are back online.

Node failures result in a storage failover event and enable the FlexGroup volume to continue operations normally. If two nodes fail in the same HA pair, then the FlexGroup volume has member volumes that are considered to be offline, and data access is fenced off until the nodes are repaired and back in working order.

10.5 Nondisruptive Volume Move Considerations

ONTAP enables you to perform nondisruptive volume moves between aggregates or nodes in the same cluster. This feature provides flexibility when you're dealing with maintenance windows or attempting to balance performance or capacity allocation in a cluster.

FlexGroup volumes also support this feature, but with even more granularity; you can move each member volume in a FlexGroup volume by using this functionality. Storage administrators therefore have a way to move workloads around in a cluster if capacity or performance concerns arise. With the ability of Active IQ Performance Manager to review individual member FlexVol volumes, you can quickly identify and resolve issues.

Note: Keep in mind that, although volume moves are nondisruptive, the amount of time that they take depends on the volume size and on the overall load on the node.

When to Use Nondisruptive Volume Moves

Nondisruptive volume moves can come in handy in the following scenarios for FlexGroup:

- The member volume is nearing capacity, and no physical storage is available on the current node.
- The member volume shares an aggregate with other FlexVol volumes and is being affected by the FlexVol volume's performance or capacity.
- A member volume is overworked in a FlexGroup volume and needs more resources.
- You want to migrate FlexGroup volumes from spinning disk to SSD for performance or from SSD to spinning disk for archiving.
- New cluster nodes or data aggregates are added.
- You are performing a head swap or other planned maintenance operations (to provide for the least amount of downtime).

Using Nondisruptive Volume Moves

Nondisruptive `volume move` for a FlexGroup member volume is available at the **admin privilege** level of the command line. Although you can use the ONTAP System Manager GUI to move FlexVol volumes, you currently can't use it to move for FlexGroup member volumes.

To move a FlexGroup member volume, complete the following steps:

1. Identify the volume that needs to be moved. Use Active IQ Performance Manager to determine this information.
2. From the command line, run the volume move start command. This command can be run at the admin privilege level.

```
cluster::> volume move start -vserver SVM -destination-aggregate aggr1_node2 -volume
flexgroup4TB__000
    flexgroup4TB__0001 flexgroup4TB__0002 flexgroup4TB__0003
    flexgroup4TB__0004 flexgroup4TB__0005 flexgroup4TB__0006
    flexgroup4TB__0007 flexgroup4TB__0008

cluster::> volume move start -vserver SVM -volume flexgroup4TB__0003 -destination-aggregate
aggr1_node2
[Job 2603] Job is queued: Move "flexgroup4TB__0003" in Vserver "SVM" to aggregate "aggr1_node2".
Use the "volume move show -vserver SVM -volume flexgroup4TB__0003" command to view the status of
this operation.
```



```
cluster::> volume move show
Vserver   Volume     State    Move Phase Percent-Complete Time-To-Complete
--------- ---------- -------- ---------- ---------------- ----------------
SVM       flexgroup4TB__0003
                     healthy  replicating
                                         45%              Tue Dec 06 13:43:01 2016
```

Auto Balance Aggregate

In Data ONTAP 8.3, the Auto Balance Aggregate feature was introduced. This feature provides ONTAP recommended nondisruptive volume moves when system performance or capacity reaches a point specified by the storage administrator. This feature is not currently supported with FlexGroup volumes.

10.6 Adding Capacity to a FlexGroup Volume

A FlexGroup volume can grow to immense capacities, but as data grows, even a massive container such as a FlexGroup volume might require more capacity.

In addition, FlexGroup performance can be affected as the capacity of member volumes becomes closer to full. See section 10.2, "Monitoring FlexGroup Capacity."

Recommendations for Adding Capacity

There are two main ways to add capacity to a FlexGroup volume:

- Grow existing members by using the `volume size` command.
- Add new members by using `volume expand`.

Each method has considerations that are detailed below. The preferred method of adding capacity to an existing FlexGroup volume is to grow the FlexGroup volume. This method preserves the existing ingest heuristics and keeps performance consistent across the FlexGroup volume, rather than favoring newer, empty member volumes. If this approach is not possible because of physical aggregate limitations or the member FlexVol volumes hitting the 100TB limit, then you should add new member volumes instead. Additionally, if the FlexGroup volume is adding capacity by way of new nodes or aggregates, it might make more sense to first use `volume move` to balance the member volumes across the nodes and then either grow them or add new member volumes in the same multiples per node.

Best Practice 22: FlexGroup Capacity General Recommendations: Increasing Volume Size
• If possible, increase capacity through `volume size` or resize from ONTAP System Manager rather than adding new members; this approach preserves the FlexGroup ingest balance. • If adding new cluster nodes, use `volume move` and `volume size` to rebalance the member volumes and increase the FlexGroup volume size. • Don't run `volume size` commands on FlexGroup member volumes individually without the guidance of NetApp Support. Run `volume size` only on the FlexGroup volume itself. • Use capacity monitoring to keep track of how full member volumes are becoming. • Use thin provisioning to set higher virtual caps of space on volumes without affecting total space allocation. • Do not overprovision a FlexGroup volume on a physical aggregate. By not overprovisioning, you will avoid headaches later as the FlexGroup volume grows closer to full and remediation steps are needed. • If you're adding new nodes or aggregates to use with the FlexGroup volume, use `volume move` to rebalance the members across the new hardware and then either grow FlexGroup by using `volume size` or add new members with `volume expand` in the same multiples per node. For example, if each node has four members, add four new members per node.

Note: There currently is no way to reduce capacity by removing member volumes. Only shrinking a volume is allowed.

Best Practice 23: FlexGroup Capacity General Recommendations—Adding New Members
• If you must add new members, be sure to add them in the same multiples as the existing FlexGroup volume. (That is, if the existing FlexGroup volume has 16 member volumes, eight per node, add 16 new members, eight per node, to promote consistent performance.) • If you add new nodes to an existing cluster and add new members to those nodes, be sure to maintain a consistent number of member volumes per node as in the existing FlexGroup volume. • Adding new members to a FlexGroup volume changes the ingest heuristics to favor the new, empty member volumes and can affect overall system performance while the new members catch up to the existing members. Add member volumes in multiples, preferably equal to the working set of member volumes. (For example, if you have eight member volumes, add eight new member volumes when adding members.) Adding new members adds capacity, and more available inodes. Be sure to consider the maximum file count and the 64-bit file ID guidance in the section "64-Bit File Identifiers.". • When you add new members to a FlexGroup volume, the existing Snapshot copies and SnapMirror relationships are no longer valid for volume-level SnapRestore operations. Previous versions and client-driven restores for files and folders are still allowed. For more information, see section 11, "FlexGroup Data Protection Best Practices."

Note: There is no way to rebalance existing data across new members in ONTAP. For rebalance suggestions, see the section "Rebalancing the Content in a FlexGroup Volume."

Growing the Volume Versus Adding New Members

If a FlexGroup volume requires capacity or increased file count, there are two main approaches.

Growing the FlexGroup Volume (volume size)

You can add capacity to existing member volumes by growing the total size of the FlexGroup volume. You do this in the CLI by using the `volume size` command, or in ONTAP System Manager. The added size is divided evenly across the member volumes in the FlexGroup volume. For instance, if you add 8TB to a FlexGroup volume with eight member volumes, each member volume grows by 1TB. Therefore, when you add space, it's important to know how many member volumes are in the FlexGroup volume. To find this number, use `volume show -name [flexgroup] -is-constituent true` from the CLI.

In deciding whether to grow the FlexGroup volume or add member volumes, consider your desired result and intent. Grow the FlexGroup volume if:

- You simply want more capacity on existing nodes/aggregates.
- You don't want to increase the total volume count in your cluster.
- Your FlexGroup member volumes are nowhere near the 100TB limit.
- You are not at the 2 billion file limit for the member volumes.
- You have available physical space where the member volumes currently live.
- You want to preserve the data balance across the member volumes.

These scenarios are by no means exhaustive; there might be other instances in which you want to grow a FlexGroup volume instead of adding new member volumes. If you are unsure, contact flexgroups-info@netapp.com.

Adding Member Volumes (volume expand)

Another way to add capacity or file count to a FlexGroup volume is by adding member volumes. Currently, NetApp officially supports 200-member FlexGroup volumes, which offers a maximum of 20PB capacity and 400 billion files. If you need more capacity than that, contact your NetApp sales representative to begin a qualification process for larger FlexGroup volumes.

To add more member volumes, you must currently use the `volume expand` command in the CLI. This command adds new, empty member volumes of exactly the **same size** as the existing FlexGroup member volumes. The number of new member volumes is determined by the `aggr-multiplier` and `aggr-list` options in the `volume expand` command. For an example of this command, see the end of this document in Command Examples.

Adding new member volumes is not the preferred way to add capacity in most cases, because the new member volumes will be empty and can throw off the ingest balance of new requests. If you are adding new member volumes, be sure to add them in multiples—preferably the same number as the volumes that already exist in the system.

However, in some use cases, adding member volumes is the best way to add capacity to a FlexGroup volume. For example:

- The FlexGroup member volumes are already at or near 100TB.
- New nodes or aggregates are being added to the cluster.
- More maxfiles are needed and the member volumes are already at their maximum values for their sizes.
- The cluster capacity is at a level where member volumes cannot be grown, and other nodes in the cluster have only enough space for member volumes of the same capacity.

These scenarios are by no means exhaustive; there might be other instances where you want to add members to a FlexGroup volume instead of growing the volume. If you are unsure, contact flexgroups-info@netapp.com.

Volume Size

To grow the volume capacity as necessary, you can run the `volume size` command on the FlexGroup volume. This command is available at the admin privilege level. This action affects the ingest heuristics favorably because the member volumes have more available free space and allocate files remotely less often.

When you use this command, the member FlexVol volumes are each increased by the total capacity/total number of member volumes. For example, if a FlexGroup volume has eight member volumes and is grown by 80TB, then each member volume increases by 10TB automatically by ONTAP.

It is important to consider these individual increases when the total FlexGroup size increase is factored in. It’s easy to forget that you are dealing with multiple FlexVol volumes when managing a FlexGroup volume.

Currently, FlexGroup volumes can be increased only with the command line. For an example of this output, see “Command Examples.”

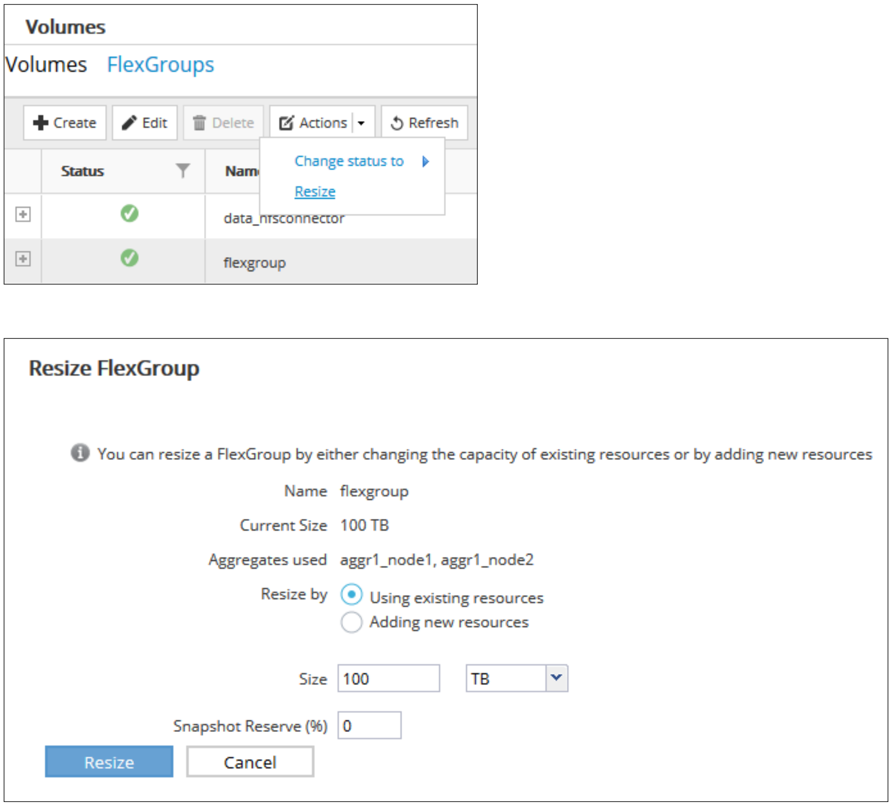
FlexGroup volumes added manual volume shrink support in ONTAP 9.6 and autoshrink functionality in ONTAP 9.3.

```
cluster::*> vol size -vserver SVM -volume flexgroup -new-size 1t
vol size: Error setting size of volume "SVM:flexgroup". The specified size is not valid because
decreasing the size of a FlexGroup is not supported. Current size: 80TB (87960930222080B).
Requested size: 1TB (1099511627776B).
```

Resizing a Volume from ONTAP System Manager

In addition to using the command line, you can also resize a volume from ONTAP System Manager. From the FlexGroup portion of the Volumes page, click the volume to select it. Then click Actions > Resize.

Figure 100) ONTAP System Manager—FlexGroup resize.



Volume Autosize (Autogrow/Autoshrink)

In ONTAP 9.3, support for volume autogrow was added for FlexGroup volumes. This support enables a storage administrator to set an autogrow policy for the FlexGroup volume that allows ONTAP to increase the FlexVol size to a predefined threshold when a volume approaches capacity. This ability is especially useful in a FlexGroup volume, because volume autogrow can help prevent member volumes from filling prematurely and causing premature out-of-space scenarios in the entire FlexGroup volume. Applying volume autogrow to a FlexGroup volume is done in the same way as with a FlexVol volume. Autoshrink is also supported with autosize. ONTAP 9.6 introduced elastic sizing, which operates independently of volume autogrow.

Autosize Interaction with Elastic Sizing

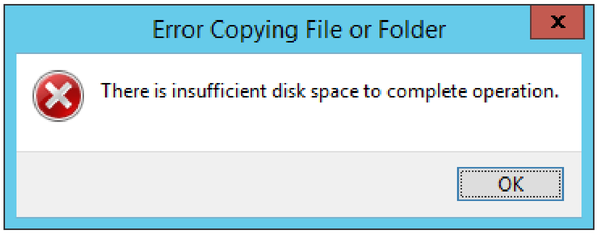
Starting in ONTAP 9.6, elastic sizing provides a way for file writes to complete in nearly filled member volumes by borrowing space from other member volumes. This takes place without growing the total size of the FlexGroup volume. As space is freed up in the filled member volume, elastic sizing begins to normalize the member volume sizes back to their original capacities.

Volume autosize, on the other hand, adds space to the total size of the FlexGroup volume by automatically growing a member volume when it reaches a space threshold.

When autosize is enabled for a volume, elastic sizing no longer takes effect for that volume. If you want to use elastic sizing for a volume, disable autosize.

How Volume Autogrow Works in a FlexGroup Volume

When a FlexGroup volume has a member that cannot honor a write, ONTAP returns an Insufficient Space (`ENOSPC`) error to the client.



Note: Starting in ONTAP 9.6, this error is usually mitigated by elastic sizing.

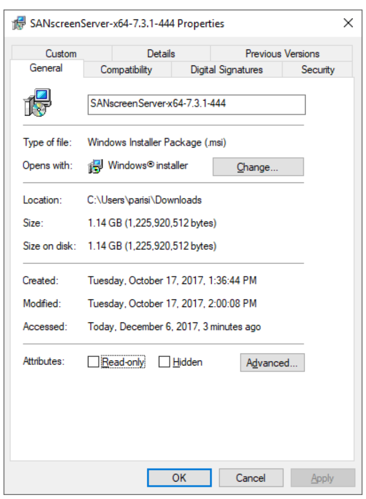
In the preceding scenario, all the member volumes in the FlexGroup volume have roughly 915MB available:

```
cluster::*> vol show -vserver DEMO -volume fgautogrow* -sort-by used -fields available
vserver volume           available
------- ---------------- ---------
DEMO    fgautogrow__0001 915.6MB
DEMO    fgautogrow__0002 915.6MB
DEMO    fgautogrow__0003 915.6MB
DEMO    fgautogrow__0004 915.6MB
DEMO    fgautogrow__0005 915.6MB
DEMO    fgautogrow__0006 915.6MB
DEMO    fgautogrow__0007 915.6MB
DEMO    fgautogrow__0008 915.6MB
```

Volume autosize is off.

```
cluster::*> vol autosize -vserver DEMO -volume fgautogrow
Volume autosize is currently OFF for volume "DEMO:fgautogrow".
```

The file being created is a 1.1GB file.



Naturally, this is the appropriate behavior. However, suppose other member volumes have enough free space, but another member volume does not. Occasionally, writes might land on member volumes that have less free space, although ONTAP does its best to avoid this scenario. For example, the member volume `fgautogrow__0004` has 915MB free, while other member volumes have 1.6GB available. In this case, if the write landed on `fgautogrow__0004`, an Insufficient Space error would occur.

```
cluster::*> vol show -vserver DEMO -volume fgautogrow* -sort-by used -fields available
vserver volume           available
------- ---------------- ---------
DEMO    fgautogrow__0001   1.6GB
DEMO    fgautogrow__0002   1.6GB
```



```
DEMO    fgautogrow__0003   1.6GB
DEMO    fgautogrow__0004 915.6MB
DEMO    fgautogrow__0005   1.6GB
DEMO    fgautogrow__0006   1.6GB
DEMO    fgautogrow__0007   1.6GB
DEMO    fgautogrow__0008   1.6GB
```

Before ONTAP 9.3, those scenarios required a storage administrator to manually intervene and grow the member volumes or delete data to restore proper functionality. Starting in ONTAP 9.3, `volume autosize` can be used to grow the member volumes automatically when the attempted writes cannot be honored.

How to Enable Volume Autosize

```
cluster::*> volume autosize -vserver DEMO -volume fgautogrow -maximum-size 100g -grow-threshold-
percent 80 -mode grow

cluster::*> vol autosize -vserver DEMO -volume fgautogrow
Volume autosize is currently ON for volume "DEMO:fgautogrow".
The volume is set to grow to a maximum of 100g when the volume-used space is above 80%.
Volume autosize for volume 'DEMO:fgautogrow' is currently in mode grow.
```

The member volumes are all 1GB in size (not recommended), so any write larger than 1GB is expected to fail if `volume autosize` is not enabled.

```
cluster::*> vol show -vserver DEMO -volume fgautogrow* -sort-by used -fields available
vserver volume           size
------- ---------------- ----
DEMO    fgautogrow__0001 1GB
DEMO    fgautogrow__0002 1GB
DEMO    fgautogrow__0003 1GB
DEMO    fgautogrow__0004 1GB
DEMO    fgautogrow__0005 1GB
DEMO    fgautogrow__0006 1GB
DEMO    fgautogrow__0007 1GB
DEMO    fgautogrow__0008 1GB
```

1GB is not a recommended size for member volumes; the minimum member volume size should be no less than 100GB. ONTAP will programmatically prevent creation of FlexGroup volumes that have member volumes smaller than 100GB with REST APIs and will warn you in the CLI:

```
cluster::*> vol create -vserver DEMO -volume smallFG -aggr-list aggr1_node1 -aggr-list-multiplier
4 -size 200g

Notice: The FlexGroup "smallFG" will be created with the following number of constituents of size
50GB: 4.

Warning: The constituent size is smaller than the recommended minimum constituent size of 100GB.
It is recommended that the size of a 4 constituent FlexGroup be at least 400GB, or the
performance of the FlexGroup will be less than optimal.
Do you want to continue? {y|n}: y
```

With `volume autosize`, however, the write succeeds because the member volume in which the write lands grows to the appropriate size to honor the write. In this case, the file is written to member volume `fgautogrow__0003`.

```
cluster::*> vol show -vserver DEMO -volume fgautogrow* -sort-by used -fields available,size
vserver volume           size   available used
------- ---------------- ------ --------- -------
DEMO    fgautogrow__0004 1GB    915.6MB   57.23MB
DEMO    fgautogrow__0005 1GB    915.6MB   57.23MB
DEMO    fgautogrow__0006 1GB    915.6MB   57.23MB
DEMO    fgautogrow__0007 1GB    915.6MB   57.23MB
DEMO    fgautogrow__0008 1GB    915.6MB   57.23MB
DEMO    fgautogrow__0002 1GB    915.5MB   57.26MB
DEMO    fgautogrow__0001 1GB    915.5MB   57.27MB
```



```
DEMO     fgautogrow__0003 1.60GB 498.7MB   1.03GB
```

When this happens, an event is triggered in the event management system and can be seen with `event log show`.

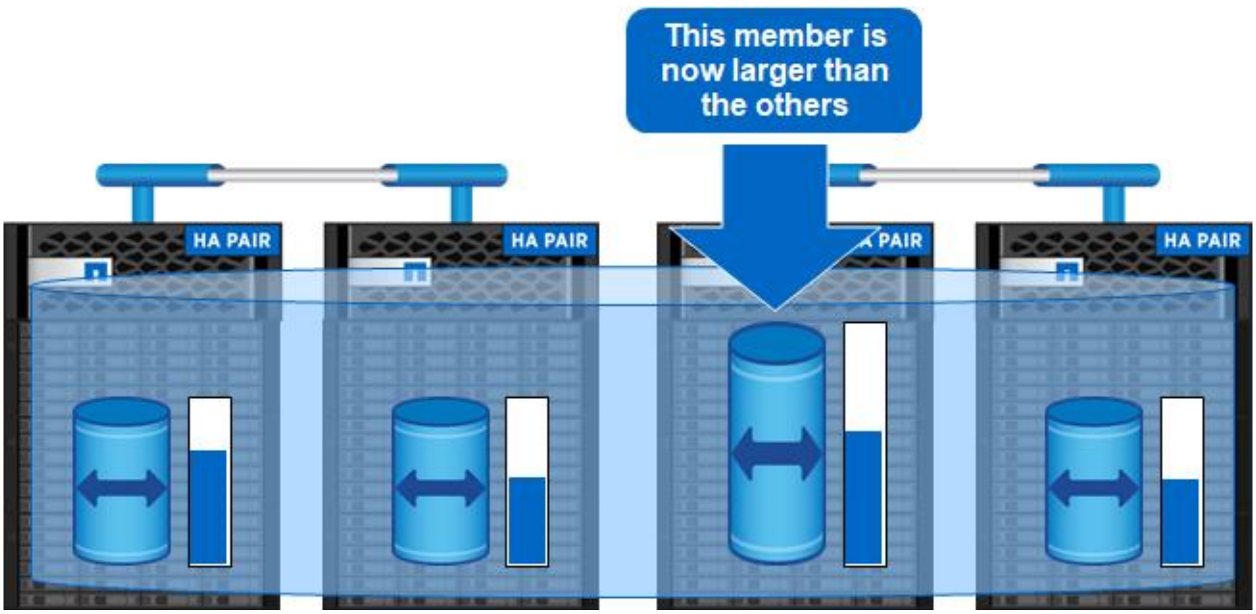
```
INFORMATIONAL wafl.vol.autoSize.done: Volume Autosize: Automatic grow of volume
'fgautogrow__0003@vserver:7e3cc08e-d9b3-11e6-85e2-00a0986b1210' by 611MB complete.
```

This event can be monitored with SNMP, by sending alerts through event destinations, or with Active IQ Unified Manager.

```
cluster::> event route show -message-name wafl.vol.autoSize.done -instance

                                     Message Name: wafl.vol.autoSize.done
                                         Severity: INFORMATIONAL
                                Corrective Action: (NONE)
                                      Description: This message occurs on successful autosize of volume.
                                Supports SNMP trap: true
                                     Destinations: -
            Number of Drops Between Transmissions: 0
Dropping Interval (Seconds) Between Transmissions: 0
```

After a member volume has been grown through autogrow, there is an imbalance of member volume available size/allocation.

Figure 101) Member volume size allocation after a `volume autosize` operation.



Best Practice 24: Actions to Take After Volume Autosize Operations
After volume autogrow takes place on a member volume, there is a discrepancy in overall member volume size in the FlexGroup volume that could affect ingest algorithms. After an event management system event for the autosize operation has been issued, grow other member volumes to the same size to help preserve the integrity of the ingest operations and load balancing in the FlexGroup volume. ONTAP 9.6 introduced elastic sizing. Elastic sizing helps mitigate `ENOSPC` errors but should not be viewed as a substitute for volume autogrow.

Volume Expand

You can grow FlexGroup volumes nondisruptively, and you can also add more capacity dynamically by using the `volume expand` command, which is available at the admin privilege level. This command provides more FlexVol member volumes in the FlexGroup volume when one of the following occurs:

- The existing member volumes have reached their maximum capacities
- The physical limits of the node capacities have been reached
- New nodes are added to the cluster

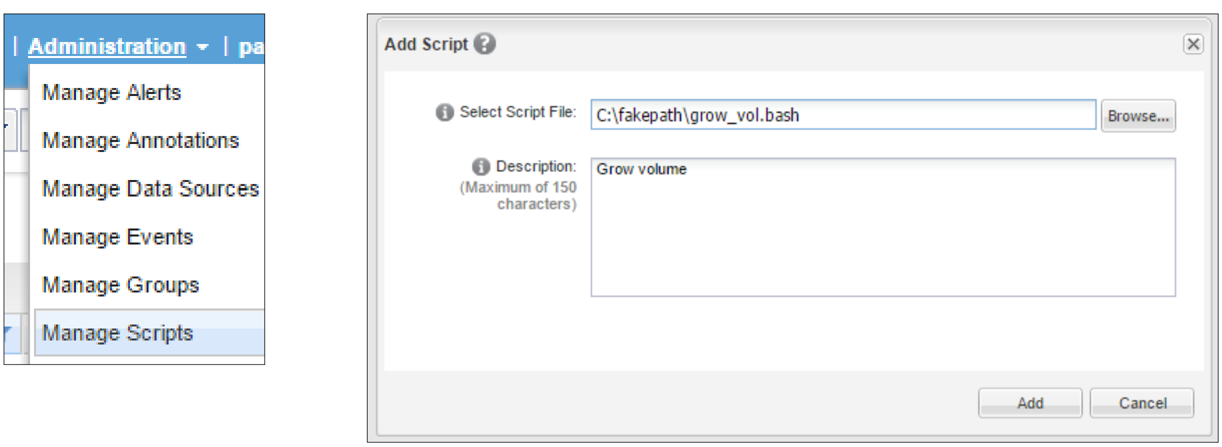
As mentioned earlier, it is preferable to grow the member volumes and to apply `volume move` in the usual way. However, when you need to use `volume expand`, be sure to follow the guidance listed in "Best Practice 23: FlexGroup Capacity General Recommendations—Adding New Members." For an example, see the section "Command Examples."

Using Active IQ Unified Manager to Automate Volume Increases

Active IQ Unified Manager alerting allows you to run a script when an alert is triggered. In response to capacity warnings, storage administrators can run a simple script to increase volume sizes automatically in lieu of the volume autogrow feature. They can also upgrade to ONTAP 9.3 or later and implement volume autogrow functionality. Using volume autogrow is the preferred method for automating volume size increases.

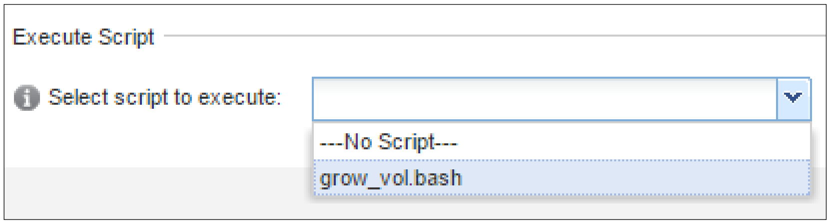
Select Administration > Manage Scripts to open the appropriate windows. The supported file formats for scripts are Perl, Shell (.sh), Windows PowerShell, and .bat files.

Figure 102) Managing scripts in Active IQ Unified Manager.

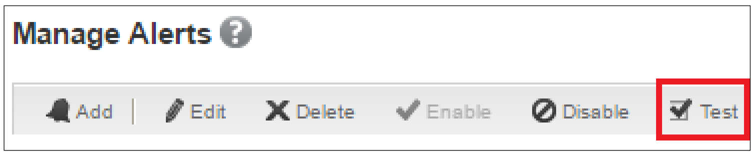


When the script is added to Active IQ, modify the alert to start the script when it is triggered. You can manage alerts through the Administration menu that is shown in Figure 103.

Figure 103) Managing and testing scripts.



After the alert is modified, you can test it in the Manage Alerts view with the Test button.



Rebalancing the Content in a FlexGroup Volume

In rare cases, a FlexGroup workload might have an imbalance of capacity in the member volumes. This scenario can occur especially when you're dealing with a mix of small files and larger files. If a client writes a large file (for example, 1TB) and then writes several smaller files, the smaller files might avoid the member volume where the 1TB file was written. Over time, the volumes might become unbalanced and potentially affect how the FlexGroup volume operates. This scenario might also occur in workloads that are prone to bursts of utilization, in which files might be allocated remotely more often because the periods of inactivity have thrown off the ingest heuristics. Generally speaking, the imbalance does not impact end user access negatively. If issues are present, open a technical support case to isolate and remediate the issue.

Currently, ONTAP has no method to rebalance the files that have already been ingested. The only way to effectively rebalance is to copy the data from a FlexGroup volume to a new, empty FlexGroup volume all at once. This process is disruptive, because clients and applications have to point to the new FlexGroup volume after the data has been moved. Also, this process is performed at a file level, so it could take a considerable amount of time. Rebalancing the files should be considered only if the imbalance of volumes creates an issue that affects production. Often, imbalances will be imperceptible to client activity. Also, ingest heuristics improvements in ONTAP 9.5P4 and later are designed to help offset impact from imbalanced FlexGroup volumes.

If rebalance is necessary, consider using the NetApp XCP Migration Tool to speed up the file copy process.

10.7 Resolving Member Volume Capacity Issues

If a FlexGroup member volume fills to capacity before other member volumes are filled, the FlexGroup volume could return `ENOSPC` (out of space) errors to clients for the whole FlexGroup volume. This issue is rare, because the ingest heuristics are designed to aggressively avoid this scenario. Also, the best practices for capacity monitoring and initial configuration that this report recommends should give storage administrators enough timely warning about impending capacity issues. However, for large files or files that grow over time, it's possible for this issue to occur. ONTAP 9.3 provides volume autogrow functionality, which is one method of preventing member volumes from running out of space. ONTAP 9.6 and later versions offer elastic sizing support, which helps you avoid "out of space" errors without having to increase the overall capacity footprint of the FlexGroup volume.

If a client receives an `ENOSPC` (out of space) error, you should take the following steps.

Verify That the FlexGroup Has Available Capacity

If you receive an error that you are out of space, you might actually be out of space. Be sure to check the overall FlexGroup capacity to see if there is enough free space for the attempted write.

Verify That the FlexGroup Is Not Hitting Maxdirsize or Maxfiles Limits

When a FlexGroup volume (or member volume) hits the assigned maxdirsize or maxfiles limits, the client error will be the same as if the volume is actually out of available capacity. Use `event log show`, Active IQ Unified Manager, System Manager, or other logging tools to see if the following events were logged:

```
fg.inodes.member.full
fg.inodes.member.nearlyFull
callhome.no.inodes
wafl.dir.size.warning
wafl.dir.size.max
wafl.dir.size.max.warning
```

When the above errors are present:

- If the volume is out of inodes, increase maxfiles by 10% at a time.
- If the volume has reached the `maxdirsize` value, either delete files from the directory or contact support for guidance on increasing `maxdirsize`.

For guidance on how to monitor for these specific events, see the section on Inode Monitoring in this document.

Verify Which Member Volume Is Running Out of Space

To verify which member volume is running out of space, run `volume show` at the **diag privilege** level or `volume show-space` at the **admin privilege** level with filters to make the output more readable. In addition, sort the values by `percent-used` and `available` for a faster look at which volume is closer to full. The volumes are sorted in ascending order by the amount of space available. In the following example, we're looking for member volumes with a noticeable discrepancy. The discrepancy is due to a single file being written to the FlexGroup volume that is nearly the same size (1TB) as the member volume (1.25TB).

```
cluster::*> volume show -vserver SVM -volume uneven* -fields size -sort-by size
vserver volume          size
------- --------------- ------
SVM     uneven_fg__0001 1.25TB
SVM     uneven_fg__0002 1.25TB
SVM     uneven_fg__0003 1.25TB
SVM     uneven_fg__0004 1.25TB
SVM     uneven_fg__0005 1.25TB
SVM     uneven_fg__0006 1.25TB
SVM     uneven_fg__0007 1.25TB
SVM     uneven_fg__0008 1.25TB
SVM     uneven_fg       10TB
9 entries were displayed.


# dd if=/dev/zero of=/uneven/1TB.file count=262144000 bs=4096
# ls -lah
total 1004G
drwxr-xr-x   2 root root  4.0K Dec  7 10:27 .
dr-xr-xr-x. 26 root root  4.0K Dec  7 10:22 ..
-rw-r--r--   1 root root 1000G Dec  7 13:00 1TB.file
drwxrwxrwx   7 root root  4.0K Dec  7 14:05 .snapshot
```

```
cluster::*> volume show -vserver SVM -volume uneven* -fields available -sort-by percent-
used,available,size
```



```
vserver volume          size   available percent-used
------- --------------- ------ --------- ------------
SVM     uneven_fg__0002 1.25TB 1.19TB    5%
SVM     uneven_fg__0004 1.25TB 1.19TB    5%
SVM     uneven_fg__0006 1.25TB 1.19TB    5%
SVM     uneven_fg__0008 1.25TB 1.19TB    5%
SVM     uneven_fg__0003 1.25TB 1.19TB    5%
SVM     uneven_fg__0005 1.25TB 1.19TB    5%
SVM     uneven_fg__0007 1.25TB 1.19TB    5%
SVM     uneven_fg       10TB   8.52TB    14%
SVM     uneven_fg__0001 1.25TB 211.7GB   83%
9 entries were displayed.
```

The output shows that the file was placed on the member volume `uneven_fg__0001`.

When this scenario occurs, the FlexGroup volume starts to favor other member volumes when ingesting new data. This situation can be seen through the `flexgroup show` output. The target percentage for that volume is 0%. That value means that the volume is the least likely to have files placed on it, but the rest of the member volumes have a 14% chance of file placement. The 100% probabilities show that the member volume has a 100% chance that the system avoids it when ingesting new data.

```
Idx   Member L            Used       Avail Urgc Targ          Probabilities
--- -------- - --------------- ---------- ---- ----  ---------------------
  1     1736 R  263264604  82%  318759808   0%   0%  [100% 100% 100% 100%]
  2     1737 L       7367   0%  318759808   0%  14%  [100% 100%  85%  85%]
  3     1738 R       7368   0%  318759808   0%  14%  [100% 100%  85%  85%]
  4     1739 L       7367   0%  318759808   0%  14%  [100% 100%  85%  85%]
  5     1740 R       7368   0%  318759808   0%  14%  [100% 100%  85%  85%]
  6     1741 L       7367   0%  318759808   0%  14%  [100% 100%  85%  85%]
  7     1742 R       7368   0%  318759808   0%  14%  [100% 100%  85%  85%]
  8     1743 L       7367   0%  318759808   0%  14%  [100% 100%  85%  85%]
```

A similar discrepancy could be seen if a volume had eight members, each with the same size file and one file that is deleted from the FlexGroup volume. In that case, new files would heavily favor the least utilized FlexVol member; therefore, performance and capacity distribution could suffer negative impacts.

In the following example, one volume is almost empty in the FlexGroup volume. This scenario could occur if you wrote seven new 1TB files to this FlexGroup volume, or if you deleted a 1TB file from the FlexGroup volume. Thus, the target percentage (the likelihood of data being placed on that volume) is nearly 100%, and the probabilities of remote placement by the other members is nearly 100%.

```
Idx   Member L            Used       Avail Urgc Targ          Probabilities
--- -------- - --------------- ---------- ---- ----  ---------------------
  1     1736 L  263264699  82%  318759808   0%   0%  [100% 100%  99%  99%]
  2     1737 R  263264678  82%  318759808   0%   0%  [100% 100%  99%  99%]
  3     1738 L  263264699  82%  318759808   0%   0%  [100% 100%  99%  99%]
  4     1739 R  263264688  82%  318759808   0%   0%  [100% 100%  99%  99%]
  5     1740 L  263264736  82%  318759808   0%   0%  [100% 100% 100% 100%]
  6     1741 R  263264688  82%  318759808   0%   0%  [100% 100%  99%  99%]
  7     1742 L  263264716  82%  318759808   0%   0%  [100% 100%  99%  99%]
  8     1743 R       7447   0%  318759808   0%  99%  [100% 100%   0%   0%]
```

Keep in mind that 1TB is not the magic-bullet file size that can cause issues; large file sizes are a matter of percentage of the available space. For more information, see the section that defines what a large file means in relation to a FlexGroup volume.

Decide How to Resolve the Issue

If a member volume is approaching 100% capacity, you have a few options.

Delete Files

This option is less reliable because the FlexGroup data layout isn't easily ascertained. Therefore, it might not be easy to determine whether deleting a dataset would help alleviate space issues. If you need assistance in locating files that are appropriate for deleting from problematic member volumes, contact NetApp Support.

Delete Snapshot Copies

This approach would free up some space in the overall FlexGroup volume, but only if the Snapshot allocated space exceeds the default 5% space reservation for Snapshot copies. To verify, use the `snapshot-space-used` field.

```
cluster::*> volume show -fields snapshot-space-used -vserver SVM -volume uneven_fg* -sort-by
snapshot-space-used
vserver volume          snapshot-space-used
------- --------------- -------------------
SVM     uneven_fg       0%
SVM     uneven_fg__0001 0%
SVM     uneven_fg__0002 0%
SVM     uneven_fg__0003 0%
SVM     uneven_fg__0004 0%
SVM     uneven_fg__0005 0%
SVM     uneven_fg__0006 0%
SVM     uneven_fg__0007 0%
SVM     uneven_fg__0008 0%
```

Increase the FlexGroup Total Size

Increasing the overall FlexGroup volume size would be the fastest and easiest way to fix issues with space allocation on member volumes. When taking this step, aim for a size that gets the FlexGroup volume below the 80% threshold for the largest member volume.

Enable Volume Autogrow

Enabling volume autogrow removes the need to perform any manual intervention if a member volume fills to capacity.

Upgrade to ONTAP 9.6 or Later

ONTAP 9.6 and later versions offer support for elastic sizing, which can help avoid “out of space” errors.

Add Disk or Shelf Capacity

If the FlexGroup volume cannot be grown because of physical limitations, adding more storage can provide the cushion that is needed to resolve capacity issues.

Increase the Size of an Individual Member Volume

In rare cases, there might not be space to increase the entire FlexGroup volume. However, there might be a way to move member volumes around in a cluster temporarily to free up enough space to increase a single member volume. You should take this step only as a last resort and only with the guidance of NetApp Support.

Increase the Maxfiles or Maxdirsize

If the capacity issue is not size but number of inodes, then either increase the `maxfiles` value on the FlexGroup volume or contact NetApp Support to increase the `maxdirsize` value.

Create and Mount a New FlexGroup Volume to the Existing FlexGroup Volume

In some situations, more space is needed, but the existing FlexGroup members are already at the 100TB capacity limit. In that case, mount a new FlexGroup volume to the existing one and point users to the new folder location while the data on the existing FlexGroup volume is rebalanced.

Fence Data Access to Read-Only on the FlexGroup Volume Until the Data Is Rebalanced

If the existing FlexGroup volume must remain intact, as a last resort, create read-only export policies and rules that prevent users from writing to the FlexGroup volume until the data can be rebalanced through file copy.

Real World Example: FlexGroup Out of Space

In the following example, a FlexGroup user was able to isolate a space issue by using a combination of Active IQ Performance Manager and NetApp AutoSupport®. The initial error seen was the following:

```
Volume flexgroup__0006@vserver:18c1fb66-57c9-11e7-a316-00a0989e94da is full (using or reserving
98% of space and 10% of inodes)
```

The error can be seen in AutoSupport, by using [Active IQ](), or by using `event log show` from the cluster CLI. It can also be configured to send notifications when the alert is triggered, as is described in the section "Capacity Monitoring and Alerting with the Command Line."

With AutoSupport, we were able to determine that member volume `flexgroup__006` hit a space issue between Sunday and Wednesday, when the alert was triggered. The `df` section showed this on Sunday:

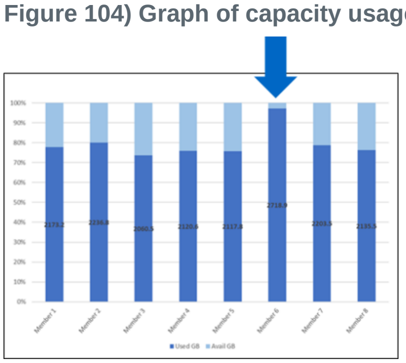
```
/vol/flexgroup__0002/ 2932657360 2547943464  384713896      87%  /vol/flexgroup__0002/
/vol/flexgroup__0002/.snapshot  154350384  239276116          0     155%
/vol/flexgroup__0002/.snapshot
/vol/flexgroup__0004/ 2932657360 2528758844  403898516      86%  /vol/flexgroup__0004/
/vol/flexgroup__0004/.snapshot  154350384  363345924          0     235%
/vol/flexgroup__0004/.snapshot
/vol/flexgroup__0006/ 2932657360 2509467192  423190168      86%  /vol/flexgroup__0006/
/vol/flexgroup__0006/.snapshot  154350384  194732648          0     126%
/vol/flexgroup__0006/.snapshot
/vol/flexgroup__0008/ 2932657360 2511307952  421349408      86%  /vol/flexgroup__0008/
/vol/flexgroup__0008/.snapshot  154350384  267119112          0     173%
/vol/flexgroup__0008/.snapshot
```

The `df` section then showed this on Wednesday:

```
/vol/flexgroup__0002/ 2932657360 2345427276  587230084      80%  /vol/flexgroup__0002/
/vol/flexgroup__0002/.snapshot  154350384   27389872  126960512      18%
/vol/flexgroup__0002/.snapshot
/vol/flexgroup__0004/ 2932657360 2223625248  709032112      76%  /vol/flexgroup__0004/
/vol/flexgroup__0004/.snapshot  154350384     122452  154227932       0%
/vol/flexgroup__0004/.snapshot
/vol/flexgroup__0006/ 2932657360 2850978768   81678592      97%  /vol/flexgroup__0006/
/vol/flexgroup__0006/.snapshot  154350384      66640  154283744       0%
/vol/flexgroup__0006/.snapshot
/vol/flexgroup__0008/ 2932657360 2239204232  693453128      76%  /vol/flexgroup__0008/
/vol/flexgroup__0008/.snapshot  154350384   22684864  131665520      15%
/vol/flexgroup__0008/.snapshot
```

When the member volumes were graphed out, the discrepancy stood out decisively.

Figure 104) Graph of capacity usage of an uneven FlexGroup volume.



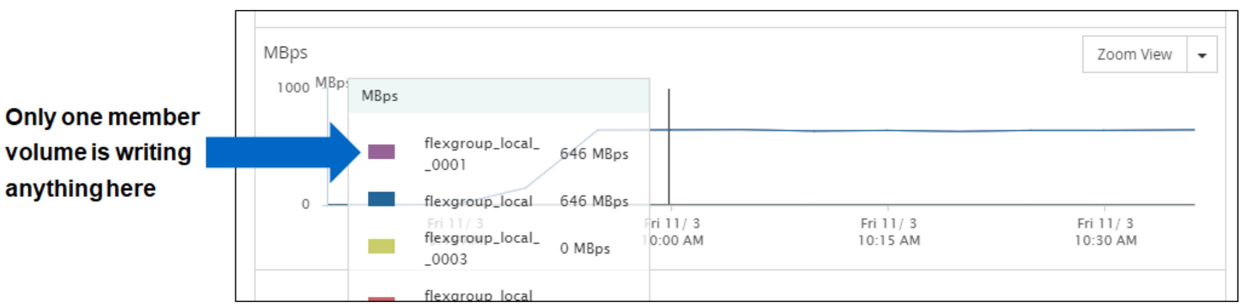
When we look at `flexgroup show` from AutoSupport, we can see that the member volume's urgency, probabilities, and target percentages are skewed to have the system avoid that member volume. So, we were able to make the educated assumption that the cause of the issue was not the creation of a new file, because the FlexGroup volume would try to direct new files to other member volumes.

Figure 105) `flexGroup show` output of unbalanced FlexGroup volume.

```
Member L            Used       Avail Urgc Targ          Probabilities
------ - --------------- ---------- ---- ----  ---------------------
  1221 L  563237873  76%  732234177   0%  14%  [100% 100%   71%    0%]
  1222 R  585459745  79%  732267266   0%  11%  [100% 100%   76%    0%]
  1223 L  539118731  73%  732140959    %  16%  [100% 100%   67%     %]
  1224 R  555272523  75%  732530542    %  14%  [100% 100%   70%     %]
  1225 L  554478833  75%  732588073    %  14%  [100% 100%   70%     %]
  1226 R  710538192  96%  732936978  86%   0%  [100% 100%  100%  100%]
  1227 L  572658179  78%  732395549   0%  13%  [100% 100%   73%   27%]
  1228 R  558708439  76%  732071719   0%  14%  [100% 100%   70%    0%]
```

Because we knew that the issue was not likely to be new file creation, we ascertained that the problem had to do with a file that grew over time. But we had a 3-day window in which the problem was seen in AutoSupport. How could we have reduced the time to resolution here? The answer is simple: Active IQ Performance Manager. Performance Manager collects data over time and saves it so that we can identify workload patterns and see when they occur. From Performance Manager, we saw evidence of a series of writes that happened in the morning. In this example, we simulated the workload on a lab system to show how that would look in Performance Manager.

Figure 106) Active IQ Performance Manager example—file that grows over time.



After narrowing down the time period of the issue and isolating the problem to a single member volume, we were able to figure out that the problem was due to an unexpected change in the workload. The FlexGroup volume had been performing normally over time, showing an almost perfectly even allocation of capacity since it was created. However, a client zipped up several files in the FlexGroup volume, causing a small file to grow in a single member volume, which ultimately filled up the volume. Therefore, we deleted that file to remediate the problem. To prevent this problem from occurring again, we set up a separate FlexVol volume, mounted it to the FlexGroup volume, and used it as a dumping ground for zip files as needed.

10.8 Applying Storage Efficiencies to a FlexGroup Volume

FlexGroup supports storage efficiency features such as the following:

- Thin provisioning
- Deduplication (aggregate and volume level, inline, and postprocess)

- Inline data compaction
- Data compression (inline and postprocess)
- Automatic deduplication scheduling

These features are effective at the member volume level individually (aside from aggregate-level deduplication), meaning that volume-level storage efficiencies like deduplication are only applied for each member volume. For example, if you have identical files in two different member volumes, then they are not deduplicated with the volume-level deduplication. However, policies are applied at the FlexGroup level for simpler management of storage efficiencies. Table 16 shows which ONTAP versions support FlexGroup-level management of these features and shows which ONTAP versions require more granular management of the efficiencies per member volume.

Best Practice 25: FlexGroup Capacity General Recommendations—Applying Storage Efficiencies
If it's necessary to enable storage efficiencies in a FlexGroup volume on a per-member volume basis, make sure that all member volumes have identical storage efficiency settings. To view storage efficiency settings, run the following commands with diag privileges:

```
volume efficiency show -volume flexgroup* -fields data-compaction,compression,inline-
compression,inline-dedupe
volume show -volume flexgroup* -fields is-sis-state-enabled
```

For more information about these features, see TR-4557, NetApp ONTAP FlexGroup Volumes: A Technical Overview and TR-4476: NetApp Data Compression, Deduplication, and Compaction.

Table 16) Storage efficiency guidance for FlexGroup in ONTAP versions.

	9.1RC1	9.1RC2	9.1GA and Later
Thin provisioning	FlexGroup level	FlexGroup level	FlexGroup level
Inline deduplication	FlexVol member	FlexGroup level	FlexGroup level
Postprocess deduplication	FlexVol member	FlexGroup level	FlexGroup level
Inline data compaction	FlexVol member	FlexGroup level	FlexGroup level
Inline data compression	FlexVol member	FlexGroup level	FlexGroup level
Postprocess data compression	FlexVol member	FlexGroup level	FlexGroup level
Cross-volume deduplication (9.2 and later)	N/A	N/A	FlexGroup level

Keep in mind the following ONTAP defined caveats:

- Inline data compaction can be used only on thin-provisioned volumes.
- Inline deduplication is allowed only on NetApp AFF or hybrid aggregate configurations.
- Cross-volume deduplication will see storage efficiencies on a per-aggregate basis.
- In general, FlexGroup volumes will not see the same storage efficiency ratios as FlexVol volumes, because of how storage efficiencies work and how FlexGroup volumes can span an entire cluster.

Applying Storage Efficiencies per FlexGroup Member Volume (ONTAP 9.1RC1)

If you're using ONTAP 9.1RC1* and a FlexGroup volume does not have support to enable storage efficiencies at the FlexGroup level, use the following command to enable it on every FlexGroup member.

```
cluster::*> volume efficiency on -vserver SVM -volume flexgroup4*
Efficiency for volume "flexgroup4TB__0001" of Vserver "SVM" is enabled.
Efficiency for volume "flexgroup4TB__0002" of Vserver "SVM" is enabled.
```



```
Efficiency for volume "flexgroup4TB__0003" of Vserver "SVM" is enabled.
Efficiency for volume "flexgroup4TB__0004" of Vserver "SVM" is enabled.
Efficiency for volume "flexgroup4TB__0005" of Vserver "SVM" is enabled.
Efficiency for volume "flexgroup4TB__0006" of Vserver "SVM" is enabled.
Efficiency for volume "flexgroup4TB__0007" of Vserver "SVM" is enabled.
Efficiency for volume "flexgroup4TB__0008" of Vserver "SVM" is enabled.
8 entries were acted on.

cluster::*> volume efficiency modify -vserver SVM -volume flexgroup4* -compression true -data-
compaction true -inline-compression true -inline-dedupe true
8 entries were modified.

cluster::*> volume efficiency show -vserver SVM -volume flexgroup4* -fields data-
compaction,compression,inline-compression,inline-dedupe
vserver volume             compression inline-compression inline-dedupe data-compaction
------- ------------------ ----------- ------------------ ------------- ---------------
SVM     flexgroup4TB__0001 true        true               true          true
SVM     flexgroup4TB__0002 true        true               true          true
SVM     flexgroup4TB__0003 true        true               true          true
SVM     flexgroup4TB__0004 true        true               true          true
SVM     flexgroup4TB__0005 true        true               true          true
SVM     flexgroup4TB__0006 true        true               true          true
SVM     flexgroup4TB__0007 true        true               true          true
SVM     flexgroup4TB__0008 true        true               true          true
```

Note: If you're using ONTAP 9.1RC1, your best approach is to upgrade the ONTAP version.

10.9 Applying Storage Efficiencies to Aggregates That Own FlexGroup Volumes

ONTAP 9.2 introduced a feature called aggregate inline deduplication. This feature offers deduplication at the aggregate level, meaning that files that are duplicates in different FlexVol volumes on the same aggregate receive the benefits of being deduplicated, providing even more storage efficiencies in ONTAP. The feature is applied at the aggregate level through the normal volume efficiency commands with the option `-cross-volume-inline-dedupe` and is independent of the FlexGroup volume.

The feature is available only for AFF systems and is disabled by default.

```
[-cross-volume-inline-dedupe {true|false}] - Cross Volume Inline Deduplication
This option is used to enable and disable cross volume inline deduplication. The default value is
false.
```

Aggregate inline deduplication offers extra benefit to FlexGroup volumes, because it can perform data reduction across multiple member volumes for large datasets. Before aggregate inline deduplication, all storage efficiencies available in ONTAP were effective only at the member volume level. For more information about aggregate inline deduplication, see TR-4476: NetApp Data Compression, Deduplication, and Data Compaction.

11 FlexGroup Data Protection Best Practices

For the FlexGroup data protection best practices, see TR-4678, Data Protection and Backup: NetApp FlexGroup Volumes.

12 Migrating to NetApp ONTAP FlexGroup

One challenge of having many files or a massive amount of capacity is deciding how to effectively move the data as quickly and as nondisruptively as possible. This challenge is greatest in high-file-count, high-metadata-operation workloads. Copies of data at the file level require file-system crawls of the attributes and the file lists, which can greatly affect the time that it takes to copy files from one location to another. That duration does not account for other aspects such as network latency, WANs, system performance bottlenecks, or other things that can make a data migration painful.

With NetApp ONTAP FlexGroup, the benefits of performance, scale, and manageability are apparent.

Data migrations can take three general forms when dealing with FlexGroup:

- Migrating from non-NetApp (third-party) storage to FlexGroup
- Migrating from NetApp Data ONTAP operating in 7-Mode to FlexGroup
- Migrating from NetApp FlexVol volumes, SAN LUNs, or Infinite Volume in ONTAP to FlexGroup

Data migrations to FlexGroup volumes are the best way to migrate. FlexGroup volume migrations currently cannot be performed with the following methods:

- FlexVol to FlexGroup volume move
- NetApp SnapMirror or SnapVault between FlexVol and FlexGroup
- 7-Mode Transition Tool (CBT and CFT)

The following sections cover these use cases and how to approach them.

12.1 NDMP

In ONTAP 9.7 and later, FlexGroup volumes now support NDMP operations. These include the `ndmpcopy` command, which can be used to migrate data from a FlexVol to a FlexGroup volume. For information about setting up `ndmpcopy`, see: https://kb.netapp.com/app/answers/answer_view/a_id/1032750.

In the following example, `ndmpcopy` was used to migrate around five million folders and files from a FlexVol to a FlexGroup volume. The process took around 51 minutes:

```
cluster::*> system node run -node ontap9-tme-8040-01 ndmpcopy -sa ndmpuser:AcDjtsU827tputjN -da
ndmpuser:AcDjtsU827tputjN 10.x.x.x:/DEMO/flexvol/nfs 10.x.x.x:/DEMO/flexgroup_16/ndmpcopy
Ndmpcopy: Starting copy [ 2 ] ...
Ndmpcopy: 10.x.x.x: Notify: Connection established
Ndmpcopy: 10.x.x.x: Notify: Connection established
Ndmpcopy: 10.x.x.x: Connect: Authentication successful
Ndmpcopy: 10.x.x.x: Connect: Authentication successful
Ndmpcopy: 10.x.x.x: Log: Session identifier: 12584
Ndmpcopy: 10.x.x.x: Log: Session identifier: 12589
Ndmpcopy: 10.x.x.x: Log: Session identifier for Restore : 12589
Ndmpcopy: 10.x.x.x: Log: Session identifier for Backup : 12584
Ndmpcopy: 10.x.x.x: Log: DUMP: creating "/DEMO/flexvol/../snapshot_for_backup.1" snapshot.
Ndmpcopy: 10.x.x.x: Log: DUMP: Using subtree dump
Ndmpcopy: 10.x.x.x: Log: DUMP: Using snapshot_for_backup.1 snapshot
Ndmpcopy: 10.x.x.x: Log: DUMP: Date of this level 0 dump snapshot: Thu Jan  9 11:53:18 2020.
Ndmpcopy: 10.x.x.x: Log: DUMP: Date of last level 0 dump: the epoch.
Ndmpcopy: 10.x.x.x: Log: DUMP: Dumping /DEMO/flexvol/nfs to NDMP connection
Ndmpcopy: 10.x.x.x: Log: DUMP: mapping (Pass I)[regular files]
Ndmpcopy: 10.x.x.x: Log: DUMP: Reference time for next incremental dump is : Fri Jun 21 16:41:27
2019
Ndmpcopy: 10.x.x.x: Log: DUMP: mapping (Pass II)[directories]
Ndmpcopy: 10.x.x.x: Log: DUMP: estimated 12524018 KB.
Ndmpcopy: 10.x.x.x: Log: DUMP: dumping (Pass III) [directories]
Ndmpcopy: 10.x.x.x: Log: RESTORE: Thu Jan  9 12:05:07 2020: Begin level 0 restore
Ndmpcopy: 10.x.x.x: Log: RESTORE: Thu Jan  9 12:05:09 2020: Reading directories from the backup
Ndmpcopy: 10.x.x.x: Log: DUMP: dumping (Pass IV) [regular files]
Ndmpcopy: 10.x.x.x: Log: RESTORE: Thu Jan  9 12:09:37 2020: Creating files and directories.
Ndmpcopy: 10.x.x.x: Log: RESTORE: Thu Jan  9 12:10:04 2020 : We have processed 58223 files and
directories.
…
Ndmpcopy: 10.x.x.x: Log: RESTORE: Thu Jan  9 12:35:04 2020 : We have processed 4814787 files and
directories.
Ndmpcopy: 10.x.x.x: Log: RESTORE: Thu Jan  9 12:38:41 2020: Writing data to files.
Ndmpcopy: 10.x.x.x: Log: DUMP: Thu Jan  9 12:38:41 2020 : We have written 1597813 KB.
Ndmpcopy: 10.x.x.x: Log: RESTORE: Thu Jan  9 12:40:04 2020 : We have read 4215061 KB from the
backup.
Ndmpcopy: 10.x.x.x: Log: DUMP: Thu Jan  9 12:43:41 2020 : We have written 10995860 KB.
Ndmpcopy: 10.x.x.x: Log: ACL_START is '11842836480'
Ndmpcopy: 10.x.x.x: Log: RESTORE: Thu Jan  9 12:44:00 2020: Restoring NT ACLs.
```



```
Ndmpcopy: 10.x.x.x: Log: DUMP: dumping (Pass V) [ACLs]
Ndmpcopy: 10.x.x.x: Log: DUMP: Debug: 11566072 KB
Ndmpcopy: 10.x.x.x: Log: DUMP: DUMP IS DONE
Ndmpcopy: 10.x.x.x: Log: DUMP: Deleting "/DEMO/flexvol/../snapshot_for_backup.1" snapshot.
Ndmpcopy: 10.x.x.x: Log: DUMP_DATE is '5856116983'
Ndmpcopy: 10.x.x.x: Notify: dump successful
Ndmpcopy: 10.x.x.x: Log: RESTORE: RESTORE IS DONE
Ndmpcopy: 10.x.x.x: Notify: restore successful
Ndmpcopy: Transfer successful [ 0 hours, 50 minutes, 53 seconds ]
Ndmpcopy: Done
```

The same dataset using `cp` over NFS took 316 minutes—six times as long as `ndmpcopy`:

```
# time cp -R /flexvol/nfs/* /flexgroup/nfscp/

real    316m26.531s
user    0m35.327s
sys     14m8.927s
```

Using the NetApp XCP Migration Tool, that dataset took just under 20 minutes—or around 60% faster than `ndmpcopy`:

```
#  xcp copy 10.193.67.219:/flexvol/nfs 10.193.67.219:/flexgroup_16/xcp
Sending statistics...
5.49M scanned, 5.49M copied, 5.49M indexed, 5.60 GiB in (4.81 MiB/s), 4.55 GiB out (3.91 MiB/s),
19m52s.
```

Note: This XCP copy was done on a VM with a 1GB network and not much RAM or CPU; more robust servers will perform even better.

12.2 FlexVol to FlexGroup Conversion

In ONTAP 9.7, it is now possible to convert a single FlexVol to a FlexGroup volume with a single member volume, in place, with less than 40 seconds disruption. This is regardless of how much data capacity or number of files reside in the volume. There is no need to remount clients, copy data, or make any other modifications that could create a maintenance window. After the FlexVol volume is converted to a FlexGroup volume, you can add new member volumes to expand the capacity.

Why to Convert a FlexVol Volume to a FlexGroup Volume

FlexGroup volumes offer a few advantages over FlexVol volumes, such as:

- Ability to expand beyond 100TB and two billion files in a single volume
- Ability to scale out capacity or performance nondisruptively
- Multi-threaded performance for high-ingest workloads
- Simplification of volume management and deployment

For example, perhaps you have a workload that is growing rapidly and you don’t want to have to migrate the data, but still want to provide more capacity. Or perhaps a workload’s performance isn’t good enough on a FlexVol volume, so you want to provide better performance handling with a FlexGroup volume. Converting can help here.

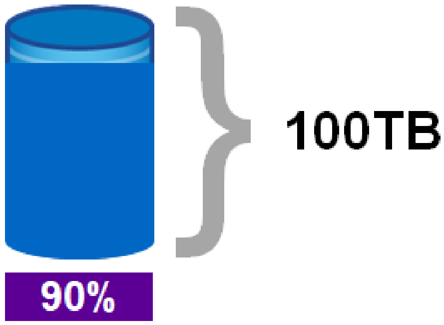
When Not to Convert a FlexVol Volume

Converting a FlexVol volume to a FlexGroup volume might not always be the best option. If you require FlexVol features that aren't available in FlexGroup volumes, then you should hold off. For example, SVM-DR and cascading SnapMirror relationships aren't supported in ONTAP 9.7, so if you need them, you should stay with FlexVol volumes.

Also, if you have a FlexVol volume that's already very large (80–100TB) and already very full (80–90%), you might want to copy the data rather than convert, because the converted FlexGroup volume would have a very large, very full member volume. This could create performance issues and doesn't fully resolve your capacity issues, particularly if that dataset contains files that grow over time.

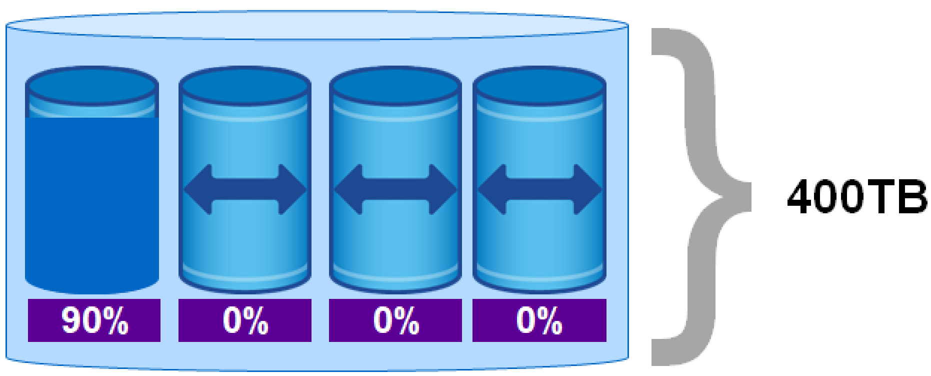
For example, if you have a FlexVol volume that is 100TB in capacity with 90TB used, then it would look like this:

Figure 107) FlexVol volume that is nearly full and at maximum capacity.



If you were to convert this 90% full volume to a FlexGroup volume, you would have a 90% full member volume. If you add new member volumes, they would be 100TB each and 0% full, so they would take on a majority of new workloads. The data would not rebalance and if the original files grew over time, you could still run out of space with nowhere to go (because 100TB is the maximum member volume size).

Figure 108) Converted FlexVol volume that now has a member volume that is at maximum capacity.



Things That Would Block a Conversion

ONTAP blocks conversion of a FlexVol for the following reasons:

- The ONTAP version isn't 9.7 on all nodes.
- ONTAP upgrade issues are preventing conversion.
- A FlexVol volume was transitioned from 7-Mode using 7MTT.
- Something is enabled on the volume that isn't supported with FlexGroup yet (SAN LUNs, Windows NFS, SMB1, part of a fan-out/cascade SnapMirror, SVM-DR, Snapshot naming/autodelete, `vmalign` set, SnapLock, space SLO, logical space enforcement/reporting, and so on)
- NetApp FlexClone volumes are present (the volume being converted can't be a parent or a clone).
- The volume is a NetApp FlexCache origin volume.
- NetApp Snapshot copies with Snap IDs greater than 255.
- Storage efficiencies are enabled (can be reenabled after).
- The volume is a source of a SnapMirror relationship, and the destination has not been converted yet.
- The volume is part of an active (not quiesced) SnapMirror relationship.

- Quotas are enabled (they must be disabled first, then reenabled after).
- Volume names are longer than 197 characters.
- ONTAP processes are running (mirrors, jobs, wafliron, NDMP backup, inode conversion in process, and so on).
- Storage Virtual Machine (SVM) root volume.
- Volume is too full.

You can check for upgrade issues with the following commands:

```
cluster::*> upgrade-revert show
cluster::*> system node image show-update-progress -node *
```

You can check for transitioned volumes with the following commands:

```
cluster::*> volume show -is-transitioned true
There are no entries matching your query.
```

You can check for Snapshot copies with Snap IDs greater than 255 with the following command:

```
cluster::*> volume snapshot show -vserver DEMO -volume testvol -logical-snap-id >255 -fields
logical-snap-id
```

How It Works

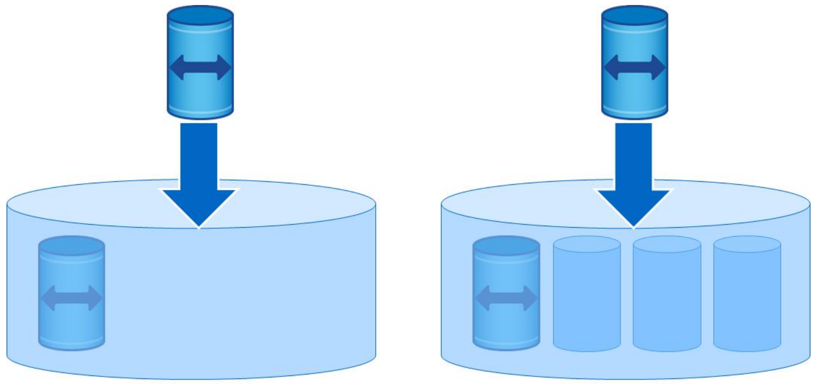
To convert a FlexVol volume to a FlexGroup volume in ONTAP 9.7, you run a single, simple command at the advanced privilege level:

```
cluster::*> volume conversion start ?
-vserver <vserver name> *Vserver Name
[-volume] <volume name> *Volume Name
[ -check-only [true] ] *Validate the Conversion Only
[ -foreground [true] ] *Foreground Process (default: true)
```

When you run this command, it takes a single FlexVol volume and converts it into a FlexGroup volume with one member. You can even run a validation of the conversion before you do the real thing.

The process is 1:1, so you can't currently convert multiple FlexVol volumes into a single FlexGroup volume. When the conversion is done, you have a single-member FlexGroup volume to which you can then add more member volumes of the same size to increase capacity and performance.

Figure 109) Converting a FlexVol volume to a FlexGroup and adding member volumes.



Other Considerations/Caveats

Although the actual conversion process is simple, there are some things to consider before converting. Most of these considerations will go away with future ONTAP releases as support is added for features, but it's still prudent to identify them here.

After the initial conversion is performed, ONTAP unmounts the volume internally and remounts it to get the new FlexGroup information into the appropriate places. Clients won't have to remount or reconnect, but they will see a disruption that last less than 1 minute while this takes place. See the section "Sample Conversion" for details. Data doesn't change at all; file handles all stay the same.

- **FabricPool** doesn't need anything. It just works. No need to rehydrate data on premises.
- **Snapshot copies** remain available for clients to access data from, but you won't be able to use them to restore the volume through `snaprestore` commands. Those Snapshot copies are marked as "pre-conversion."
- **SnapMirror relationships** pick up where they left off without rebaselining, provided the source and destination volumes have both been converted. But there are no SnapMirror restores of the volume—just file retrieval from clients. SnapMirror destinations need to be converted first.
- **FlexClone volumes** need to be deleted or split from the volume to be converted.
- **Storage efficiencies** need to be disabled during the conversion, but your space savings are preserved after the conversion.
- **FlexCache** instances with an origin volume being converted will need to be deleted.
- **Space guarantees** can affect how large a FlexGroup volume can become if they're volume guarantees. New member volumes need to be the same size as the existing members, so you need adequate space to honor them.
- **Quotas** are supported in FlexGroup volumes but are done a bit differently than in FlexVol volumes. So, while the conversion is being performed, quotas must be disabled (`quota off`) and then reenabled later (`quota on`).

Conversion to FlexGroup volumes is a one-way street after you expand it, so be sure you're ready to make the jump. If anything goes wrong during the conversion process, there is a rescue method that NetApp Support can help you use so that your data is safe even if you run into an issue.

When you expand the FlexGroup volume to add new member volumes, they are the same size as the converted member volume, so be sure there is adequate space available. Additionally, the existing data that resides in the original volume remains in that member volume. Data does not redistribute. Instead, the FlexGroup volume favors newly added member volumes for new files.

Are You Nervous About Converting?

If you don't feel comfortable about converting your production FlexVol volume to a FlexGroup volume right away, you have options.

First, ONTAP allows you to run a check on the conversion command with `–check-only true` that tells you what prerequisites you might be missing.

```
cluster::*> volume conversion start -vserver DEMO -volume flexvol -foreground true -check-only
true
Error: command failed: Cannot convert volume "flexvol" in Vserver "DEMO" to a FlexGroup. Correct
the following issues and retry the command:
* The volume has Snapshot copies with IDs greater than 255. Use the (privilege: advanced) "volume
snapshot show -vserver DEMO -volume flexvol -logical-snap-id >255 -fields logical-snap-id"
command to list the Snapshot copies with IDs greater than 255 then delete them using the
"snapshot delete -vserver DEMO -volume flexvol" command.
* Quotas are enabled. Use the 'volume quota off -vserver DEMO -volume flexvol' command to disable
quotas.
```



```
* Cannot convert because the source "flexvol" of a SnapMirror relationship is source to more than
one SnapMirror relationship. Delete other Snapmirror relationships, and then try the conversion
of the source "flexvol" volume.
* Only volumes with logical space reporting disabled can be converted. Use the 'volume modify -
vserver DEMO -volume flexvol -is-space-reporting-logical false' command to disable logical space
reporting.
```

Also, remember that ONTAP can create multiple SVMs, which can be fenced off from network access. You can use this approach to test things such as volume conversion. The only trick is getting a copy of that data over—but it's really not that tricky.

Option 1: SnapMirror

You can use SnapMirror to replicate your “to be converted” volume to the same SVM or a new SVM. Then, break the mirror and delete the relationship. Now you have a sandbox copy of your volume, complete with Snapshot copies, to test out conversion, expansion, and performance.

Option 2: FlexClone and Volume Rehost

If you don't have SnapMirror or you want to try a method that is less taxing on your network, you can use a combination of FlexClone (an instant copy of your volume backed by a Snapshot copy) and `volume rehost` (an instant move of the volume from one SVM to another). Keep in mind that FlexClone copies cannot be rehosted, but you can split the clone and then rehost.

Essentially, the process is as follows:

1. Use flexclone create.
2. Use flexclone split.
3. Issue volume rehost to the new SVM (or convert on the existing SVM).

Sample Conversion

In this sample conversion, before we converted a volume, we added around 300,000 files to help determine how long the process might take with many files present.

```
cluster::*> df -i lotsafiles
Filesystem iused ifree %iused Mounted on Vserver
/vol/lotsafiles/ 330197 20920929 1% /lotsafiles DEMO

cluster::*> volume show lots*
Vserver   Volume       Aggregate    State      Type       Size  Available Used%
--------- ------------ ------------ ---------- ---- ---------- ---------- -----
DEMO      lotsafiles   aggr1_node1  online     RW         10TB     7.33TB    0%
```

First, let's try out the validation.

```
cluster::*> volume conversion start -vserver DEMO -volume lotsafiles -foreground true -check-only
true
Error: command failed: Cannot convert volume "lotsafiles" in Vserver "DEMO" to a FlexGroup.
Correct the following issues and retry the command:
* SMB1 is enabled on Vserver "DEMO". Use the 'vserver cifs options modify -smb1-enabled false -
vserver DEMO' command to disable SMB1.
* The volume contains LUNs. Use the "lun delete -vserver DEMO -volume lotsafiles -lun *" command
to remove the LUNs, or use the "lun move start" command to relocate the LUNs to other
FlexVols.
* NFSv3 MS-DOS client support is enabled on Vserver "DEMO". Use the "vserver nfs modify -vserver
DEMO -v3-ms-dos-client disabled" command to disable NFSv3 MS-DOS client support on the
Vserver. Note that disabling this support will disable access for all NFSv3 MS-DOS clients
connected to Vserver "DEMO".
```

As you can see, there are some blockers, such as SMB1 and the LUN we created (to intentionally break conversion). So, we clear them with the recommendations and run the validation again. We see some caveats:


```
cluster::*> volume conversion start -vserver DEMO -volume lotsafiles -foreground true -check-only
true
Conversion of volume "lotsafiles" in Vserver "DEMO" to a FlexGroup can proceed with the following
warnings:
* After the volume is converted to a FlexGroup, it will not be possible to change it back to a
flexible volume.
* Converting flexible volume "lotsafiles" in Vserver "DEMO" to a FlexGroup will cause the state
of all Snapshot copies from the volume to be set to "pre-conversion". Pre-conversion Snapshot
copies cannot be restored.
```

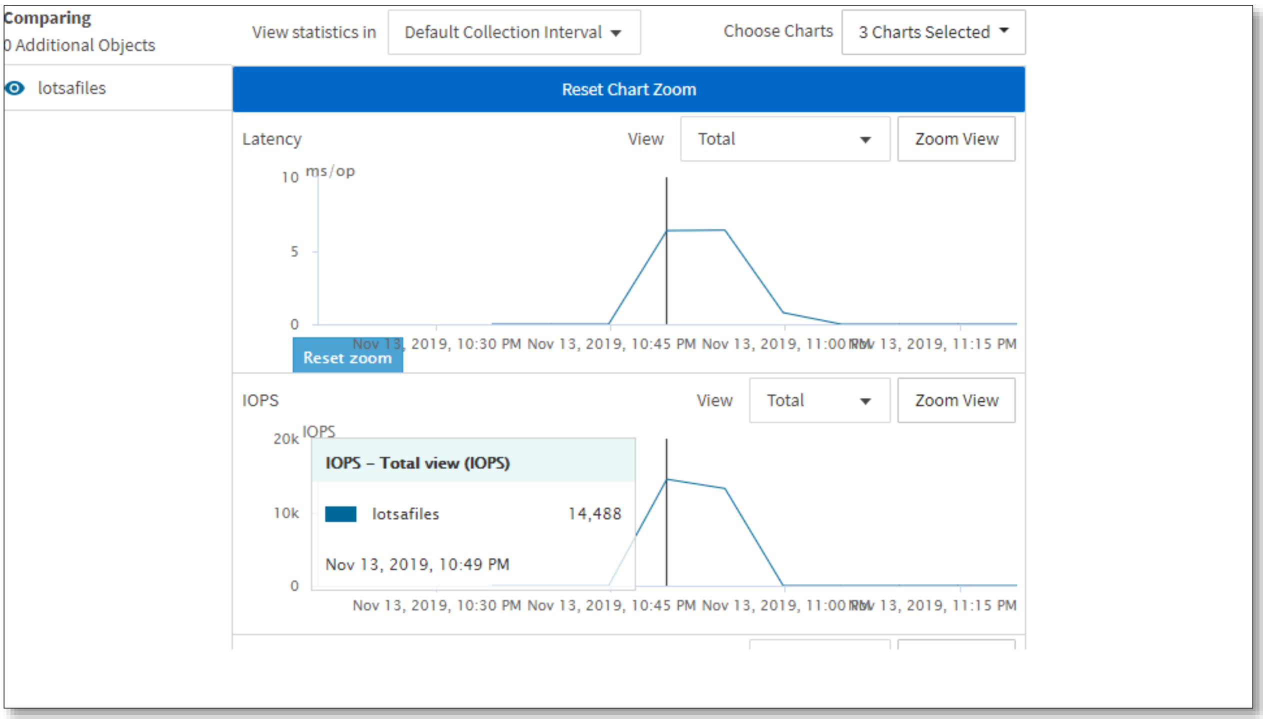
Now, let's convert. But, first, we start a script that takes a while to complete, while also using NetApp Active IQ Performance Manager to monitor performance during the conversion.

The conversion of the volume takes less than 1 minute, and the only disruption is a slight drop in IOPS.

```
cluster::*> volume conversion start -vserver DEMO -volume lotsafiles -foreground true

Warning: After the volume is converted to a FlexGroup, it will not be possible to change it back
to a flexible volume.
Do you want to continue? {y|n}: y
Warning: Converting flexible volume "lotsafiles" in Vserver "DEMO" to a FlexGroup will cause the
state of all Snapshot copies from the volume to be set to "pre-conversion". Pre-conversion
Snapshot copies cannot be restored.
Do you want to continue? {y|n}: y
[Job 23671] Job succeeded: success
cluster::*> statistics show-periodic
cpu cpu total fcache total total data data data cluster cluster cluster disk disk pkts pkts
avg busy ops nfs-ops cifs-ops ops spin-ops recv sent busy recv sent busy recv sent read write
recv sent
---- ---- -------- -------- -------- -------- -------- -------- -------- ---- -------- -------- -
------ -------- -------- -------- -------- -------- --------
34% 44% 14978 14968 10 0 14978 14.7MB 15.4MB 0% 3.21MB 3.84MB 0% 11.5MB 11.6MB 4.43MB 1.50MB
49208 55026
40% 45% 14929 14929 0 0 14929 15.2MB 15.7MB 0% 3.21MB 3.84MB 0% 12.0MB 11.9MB 3.93MB 641KB 49983
55712
36% 44% 15020 15020 0 0 15019 14.8MB 15.4MB 0% 3.24MB 3.87MB 0% 11.5MB 11.5MB 3.91MB 23.9KB 49838
55806
30% 39% 15704 15694 10 0 15704 15.0MB 15.7MB 0% 3.29MB 3.95MB 0% 11.8MB 11.8MB 2.12MB 4.99MB
50936 57112
32% 43% 14352 14352 0 0 14352 14.7MB 15.3MB 0% 3.33MB 3.97MB 0% 11.3MB 11.3MB 4.19MB 27.3MB 49736
55707
37% 44% 14807 14797 10 0 14807 14.5MB 15.0MB 0% 3.09MB 3.68MB 0% 11.4MB 11.4MB 4.34MB 2.79MB
48352 53616
39% 43% 15075 15075 0 0 15076 14.9MB 15.6MB 0% 3.24MB 3.86MB 0% 11.7MB 11.7MB 3.48MB 696KB 50124
55971
32% 42% 14998 14998 0 0 14997 15.1MB 15.8MB 0% 3.23MB 3.87MB 0% 11.9MB 11.9MB 3.68MB 815KB 49606
55692
38% 43% 15038 15025 13 0 15036 14.7MB 15.2MB 0% 3.27MB 3.92MB 0% 11.4MB 11.3MB 3.46MB 15.8KB
50256 56150
43% 44% 15132 15132 0 0 15133 15.0MB 15.7MB 0% 3.22MB 3.87MB 0% 11.8MB 11.8MB 1.93MB 15.9KB 50030
55938
34% 42% 15828 15817 10 0 15827 15.8MB 16.5MB 0% 3.39MB 4.10MB 0% 12.4MB 12.3MB 4.02MB 21.6MB
52142 58771
28% 39% 11807 11807 0 0 11807 12.3MB 13.1MB 0% 2.55MB 3.07MB 0% 9.80MB 9.99MB 6.76MB 27.9MB 38752
43748
33% 42% 15108 15108 0 0 15107 15.1MB 15.5MB 0% 3.32MB 3.91MB 0% 11.7MB 11.6MB 3.50MB 1.17MB 50903
56143
32% 42% 16143 16133 10 0 16143 15.1MB 15.8MB 0% 3.28MB 3.95MB 0% 11.8MB 11.8MB 3.78MB 9.00MB
50922 57403
24% 34% 8843 8843 0 0 8861 14.2MB 14.9MB 0% 3.70MB 4.44MB 0% 10.5MB 10.5MB 8.46MB 10.7MB 46174
53157
27% 37% 10949 10949 0 0 11177 9.91MB 10.2MB 0% 2.45MB 2.84MB 0% 7.46MB 7.40MB 5.55MB 1.67MB 31764
35032
28% 38% 12580 12567 13 0 12579 13.3MB 13.8MB 0% 2.76MB 3.26MB 0% 10.5MB 10.6MB 3.92MB 19.9KB
44119 48488
30% 40% 14300 14300 0 0 14298 14.2MB 14.7MB 0% 3.09MB 3.68MB 0% 11.1MB 11.1MB 2.66MB 600KB 47282
52789
31% 41% 14514 14503 10 0 14514 14.3MB 14.9MB 0% 3.15MB 3.75MB 0% 11.2MB 11.2MB 3.65MB 728KB 48093
53532
```

```
31% 42% 14626 14626 0 0 14626 14.3MB 14.9MB 0% 3.16MB 3.77MB 0% 11.1MB 11.1MB 4.84MB 1.14MB 47936
53645
cluster: cluster.cluster: 11/13/2019 22:44:39
cpu cpu total fcache total total data data data cluster cluster cluster disk disk pkts pkts
avg busy ops nfs-ops cifs-ops ops spin-ops recv sent busy recv sent busy recv sent read write
recv sent
---- ---- -------- -------- -------- -------- -------- -------- -------- ---- -------- -------- -
------ -------- -------- -------- -------- -------- --------
30% 39% 15356 15349 7 0 15370 15.3MB 15.8MB 0% 3.29MB 3.94MB 0% 12.0MB 11.8MB 3.18MB 6.90MB 50493
56425
32% 42% 14156 14146 10 0 14156 14.6MB 15.3MB 0% 3.09MB 3.68MB 0% 11.5MB 11.7MB 5.49MB 16.3MB
48159 53678
```

This is what the performance looked like from Active IQ:



And now we have a single member FlexGroup volume.

```
cluster::*> volume show lots*
Vserver Volume Aggregate State Type Size Available Used%
--------- ------------ ------------ ---------- ---- ---------- ---------- -----
DEMO lotsafiles - online RW 10TB 7.33TB 0%
DEMO lotsafiles__0001 aggr1_node1 online RW 10TB 7.33TB 0%
2 entries were displayed.
```

And our Snapshot copies are still there, but are marked as “pre-conversion.”

```
cluster::> set diag
cluster::*> snapshot show -vserver DEMO -volume lotsafiles -fields is-convert-recovery,state
vserver volume snapshot state is-convert-recovery
------- ---------- -------- -------------- -------------------
DEMO lotsafiles base pre-conversion false
DEMO lotsafiles hourly.2019-11-13_1705
pre-conversion false
DEMO lotsafiles hourly.2019-11-13_1805
pre-conversion false
DEMO lotsafiles hourly.2019-11-13_1905
pre-conversion false
DEMO lotsafiles hourly.2019-11-13_2005
pre-conversion false
```



```
DEMO lotsafiles hourly.2019-11-13_2105
pre-conversion false
DEMO lotsafiles hourly.2019-11-13_2205
pre-conversion false
DEMO lotsafiles clone_clone.2019-11-13_223144.0
pre-conversion false
DEMO lotsafiles convert.2019-11-13_224411
pre-conversion true
9 entries were displayed.
```

When a Snapshot copy is in “pre-conversion” state, using it for SnapRestore operation fails.

```
cluster::*> snapshot restore -vserver DEMO -volume lotsafiles -snapshot convert.2019-11-13_224411

Error: command failed: Promoting a pre-conversion Snapshot copy is not supported.
```

However, we can still obtain files from the client using the Snapshot copies.

```
[root@centos7 scripts]# cd /lotsafiles/.snapshot/convert.2019-11-13_224411/pre-convert/
[root@centos7 pre-convert]# ls
topdir_0 topdir_14 topdir_2 topdir_25 topdir_30 topdir_36 topdir_41 topdir_47 topdir_52 topdir_58
topdir_63 topdir_69 topdir_74 topdir_8 topdir_85 topdir_90 topdir_96
topdir_1 topdir_15 topdir_20 topdir_26 topdir_31 topdir_37 topdir_42 topdir_48 topdir_53
topdir_59 topdir_64 topdir_7 topdir_75 topdir_80 topdir_86 topdir_91 topdir_97
topdir_10 topdir_16 topdir_21 topdir_27 topdir_32 topdir_38 topdir_43 topdir_49 topdir_54
topdir_6 topdir_65 topdir_70 topdir_76 topdir_81 topdir_87 topdir_92 topdir_98
topdir_11 topdir_17 topdir_22 topdir_28 topdir_33 topdir_39 topdir_44 topdir_5 topdir_55
topdir_60 topdir_66 topdir_71 topdir_77 topdir_82 topdir_88 topdir_93 topdir_99
topdir_12 topdir_18 topdir_23 topdir_29 topdir_34 topdir_4 topdir_45 topdir_50 topdir_56
topdir_61 topdir_67 topdir_72 topdir_78 topdir_83 topdir_89 topdir_94
topdir_13 topdir_19 topdir_24 topdir_3 topdir_35 topdir_40 topdir_46 topdir_51 topdir_57
topdir_62 topdir_68 topdir_73 topdir_79 topdir_84 topdir_9 topdir_95
```

Growing the newly converted FlexGroup volume is simple. We can add more member volumes by using `volume expand`.

```
cluster::*> volume expand -vserver DEMO -volume lotsafiles -aggr-list aggr1_node1,aggr1_node2 -
aggr-list-multiplier 2

Warning: The following number of constituents of size 10TB will be added to FlexGroup
"lotsafiles": 4. Expanding the FlexGroup will cause the state of all Snapshot copies to be set to
"partial".
Partial Snapshot copies cannot be restored.
Do you want to continue? {y|n}: y

Warning: FlexGroup "lotsafiles" is a converted flexible volume. If this volume is expanded, it
will no longer be able to be converted back to being a flexible volume.
Do you want to continue? {y|n}: y
[Job 23676] Job succeeded: Successful
```

But remember, the data doesn't redistribute. The original member volume keeps the files in place.

```
cluster::*> df -i lots*
Filesystem iused ifree %iused Mounted on Vserver
/vol/lotsafiles/ 3630682 102624948 3% /lotsafiles DEMO
/vol/lotsafiles__0001/ 3630298 17620828 17% /lotsafiles DEMO
/vol/lotsafiles__0002/ 96 21251030 0% --- DEMO
/vol/lotsafiles__0003/ 96 21251030 0% --- DEMO
/vol/lotsafiles__0004/ 96 21251030 0% --- DEMO
/vol/lotsafiles__0005/ 96 21251030 0% --- DEMO
6 entries were displayed.

cluster::*> df -h lots*
Filesystem total used avail capacity Mounted on Vserver
/vol/lotsafiles/ 47TB 2735MB 14TB 0% /lotsafiles DEMO
/vol/lotsafiles/.snapshot
2560GB 49MB 2559GB 0% /lotsafiles/.snapshot DEMO
/vol/lotsafiles__0001/ 9728GB 2505MB 7505GB 0% /lotsafiles DEMO
/vol/lotsafiles__0001/.snapshot
```



```
512GB 49MB 511GB 0% /lotsafiles/.snapshot DEMO
/vol/lotsafiles__0002/ 9728GB 57MB 7505GB 0% --- DEMO
/vol/lotsafiles__0002/.snapshot
512GB 0B 512GB 0% --- DEMO
/vol/lotsafiles__0003/ 9728GB 57MB 7766GB 0% --- DEMO
/vol/lotsafiles__0003/.snapshot
512GB 0B 512GB 0% --- DEMO
/vol/lotsafiles__0004/ 9728GB 57MB 7505GB 0% --- DEMO
/vol/lotsafiles__0004/.snapshot
512GB 0B 512GB 0% --- DEMO
/vol/lotsafiles__0005/ 9728GB 57MB 7766GB 0% --- DEMO
/vol/lotsafiles__0005/.snapshot
512GB 0B 512GB 0% --- DEMO
12 entries were displayed.
```

Converting a FlexVol Volume in a SnapMirror Relationship

You can also convert FlexVol volumes that are part of existing SnapMirror relationships without disruption.

1. Here is a volume in a SnapMirror relationship.

```
cluster::*> snapmirror show -destination-path data_dst -fields state
source-path destination-path state
----------- ---------------- ------------
DEMO:data   DEMO:data_dst    Snapmirrored
```

2. If you try to convert the source, you get an error.

```
cluster::*> vol conversion start -vserver DEMO -volume data -check-only true

Error: command failed: Cannot convert volume "data" in Vserver "DEMO" to a FlexGroup. Correct the
following issues and retry the command:
       * Cannot convert source volume "data" because destination volume "data_dst" of the
SnapMirror relationship with "data" as the source is not converted.  First check if the source
can be converted to a FlexGroup volume using "vol conversion start -volume data -convert-to
flexgroup -check-only true". If the conversion of the source can proceed then first convert the
destination and then convert the source.
```

3. So, you would need to convert the destination first. To do that, you must quiesce the SnapMirror relationship.

```
cluster::*> vol conversion start -vserver DEMO -volume data_dst -check-only true

Error: command failed: Cannot convert volume "data_dst" in Vserver "DEMO" to a FlexGroup. Correct
the following issues and retry the command:
* The relationship was not quiesced. Quiesce SnapMirror relationship using "snapmirror quiesce -
destination-path data_dst" and then try the conversion.
```

4. Now you can convert the volume.

```
cluster::*> snapmirror quiesce -destination-path DEMO:data_dst
Operation succeeded: snapmirror quiesce for destination "DEMO:data_dst".

cluster::*> vol conversion start -vserver DEMO -volume data_dst -check-only true
Conversion of volume "data_dst" in Vserver "DEMO" to a FlexGroup can proceed with the following
warnings:
* After the volume is converted to a FlexGroup, it will not be possible to change it back to a
flexible volume.
* Converting flexible volume "data_dst" in Vserver "DEMO" to a FlexGroup will cause the state of
all Snapshot copies from the volume to be set to "pre-conversion". Pre-conversion Snapshot copies
cannot be restored.
```

5. When you convert the volume, the system lets you know your next steps.

```
cluster::*> vol conversion start -vserver DEMO -volume data_dst

Warning: After the volume is converted to a FlexGroup, it will not be possible to change it back
to a flexible volume.
Do you want to continue? {y|n}: y
```



```
Warning: Converting flexible volume "data_dst" in Vserver "DEMO" to a FlexGroup will cause the
state of all Snapshot copies from the volume to be set to "pre-conversion". Pre-conversion
Snapshot copies cannot be restored.
Do you want to continue? {y|n}: y
[Job 23710] Job succeeded: SnapMirror destination volume "data_dst" has been successfully
converted to a FlexGroup volume. You must now convert the relationship's source volume,
"DEMO:data", to a FlexGroup. Then, re-establish the SnapMirror relationship using the "snapmirror
resync" command.
```

6. Now convert the source volume.

```
cluster::*> vol conversion start -vserver DEMO -volume data

Warning: After the volume is converted to a FlexGroup, it will not be possible to change it back
to a flexible volume.
Do you want to continue? {y|n}: y
Warning: Converting flexible volume "data" in Vserver "DEMO" to a FlexGroup will cause the state
of all Snapshot copies from the volume to be set to "pre-conversion". Pre-conversion Snapshot
copies cannot be restored.
Do you want to continue? {y|n}: y
[Job 23712] Job succeeded: success
```

7. Resync the mirror.

```
cluster::*> snapmirror resync -destination-path DEMO:data_dst
Operation is queued: snapmirror resync to destination "DEMO:data_dst".

cluster::*> snapmirror show -destination-path DEMO:data_dst -fields state
source-path destination-path state
----------- ---------------- ------------
DEMO:data DEMO:data_dst Snapmirrored
```

Conversion works fine, but the most important part of a SnapMirror relationship is the restore. So you must see if you can access files from the destination volume's Snapshot copy.

1. First, mount the source and destination and compare `ls` output.

```
# mount -o nfsvers=3 DEMO:/data_dst /dst
# mount -o nfsvers=3 DEMO:/data /data
```

This is what's in the source volume.

```
# ls -lah /data
total 14G
drwxrwxrwx 6 root root 4.0K Nov 14 11:57 .
dr-xr-xr-x. 54 root root 4.0K Nov 15 10:08 ..
drwxrwxrwx 2 root root 4.0K Sep 14 2018 cifslink
drwxr-xr-x 12 root root 4.0K Nov 16 2018 nas
-rwxrwxrwx 1 prof1 ProfGroup 0 Oct 3 14:32 newfile
drwxrwxrwx 5 root root 4.0K Nov 15 10:06 .snapshot
lrwxrwxrwx 1 root root 23 Sep 14 2018 symlink -> /shared/unix/linkedfile
drwxrwxrwx 2 root bin 4.0K Jan 31 2019 test
drwxrwxrwx 3 root root 4.0K Sep 14 2018 unix
-rwxrwxrwx 1 newuser1 ProfGroup 0 Jan 14 2019 userfile
-rwxrwxrwx 1 root root 6.7G Nov 14 11:58 Windows2.iso
-rwxrwxrwx 1 root root 6.7G Nov 14 11:37 Windows.iso
```

The destination volume matches exactly, as it should.

```
# ls -lah /dst
total 14G
drwxrwxrwx 6 root root 4.0K Nov 14 11:57 .
dr-xr-xr-x. 54 root root 4.0K Nov 15 10:08 ..
drwxrwxrwx 2 root root 4.0K Sep 14 2018 cifslink
dr-xr-xr-x 2 root root 0 Nov 15 2018 nas
-rwxrwxrwx 1 prof1 ProfGroup 0 Oct 3 14:32 newfile
drwxrwxrwx 4 root root 4.0K Nov 15 10:05 .snapshot
lrwxrwxrwx 1 root root 23 Sep 14 2018 symlink -> /shared/unix/linkedfile
drwxrwxrwx 2 root bin 4.0K Jan 31 2019 test
drwxrwxrwx 3 root root 4.0K Sep 14 2018 unix
-rwxrwxrwx 1 newuser1 ProfGroup 0 Jan 14 2019 userfile
```



```
-rwxrwxrwx 1 root root 6.7G Nov 14 11:58 Windows2.iso
-rwxrwxrwx 1 root root 6.7G Nov 14 11:37 Windows.iso
```

2. If you `ls` to the Snapshot copy in the destination volume, you see the expected files.

```
# ls -lah /dst/.snapshot/snapmirror.7e3cc08e-d9b3-11e6-85e2-00a0986b1210_2163227795.2019-11-
15_100555/
total 14G
drwxrwxrwx 6 root root 4.0K Nov 14 11:57 .
drwxrwxrwx 4 root root 4.0K Nov 15 10:05 ..
drwxrwxrwx 2 root root 4.0K Sep 14 2018 cifslink
dr-xr-xr-x 2 root root 0 Nov 15 2018 nas
-rwxrwxrwx 1 prof1 ProfGroup 0 Oct 3 14:32 newfile
lrwxrwxrwx 1 root root 23 Sep 14 2018 symlink -> /shared/unix/linkedfile
drwxrwxrwx 2 root bin 4.0K Jan 31 2019 test
drwxrwxrwx 3 root root 4.0K Sep 14 2018 unix
-rwxrwxrwx 1 newuser1 ProfGroup 0 Jan 14 2019 userfile
-rwxrwxrwx 1 root root 6.7G Nov 14 11:58 Windows2.iso
-rwxrwxrwx 1 root root 6.7G Nov 14 11:37 Windows.iso
```

3. Now expand the FlexGroup source to provide more capacity.

```
cluster::*> volume expand -vserver DEMO -volume data -aggr-list aggr1_node1,aggr1_node2 -aggr-
list-multiplier

Warning: The following number of constituents of size 30TB will be added to FlexGroup "data": 4.
Expanding the FlexGroup will cause the state of all Snapshot copies to be set to "partial".
Partial Snapshot copies cannot be restored.
Do you want to continue? {y|n}: y
[Job 23720] Job succeeded: Successful
```

The source volume now has five member volumes. The destination volume has only one.

```
cluster::*> vol show -vserver DEMO -volume data*
Vserver Volume Aggregate State Type Size Available Used%
--------- ------------ ------------ ---------- ---- ---------- ---------- -----
DEMO data - online RW 150TB 14.89TB 0%
DEMO data__0001 aggr1_node2 online RW 30TB 7.57TB 0%
DEMO data__0002 aggr1_node1 online RW 30TB 7.32TB 0%
DEMO data__0003 aggr1_node2 online RW 30TB 7.57TB 0%
DEMO data__0004 aggr1_node1 online RW 30TB 7.32TB 0%
DEMO data__0005 aggr1_node2 online RW 30TB 7.57TB 0%
DEMO data_dst - online DP 30TB 7.32TB 0%
DEMO data_dst__0001
aggr1_node1 online DP 30TB 7.32TB 0%
8 entries were displayed.
```

4. Update the mirror, and ONTAP fixes it for you.

```
cluster::*> snapmirror update -destination-path DEMO:data_dst
Operation is queued: snapmirror update of destination "DEMO:data_dst".
```

The update initially fails with the following error message:

```
Last Transfer Error: A SnapMirror transfer for the relationship with destination FlexGroup
"DEMO:data_dst" was aborted because the source FlexGroup was expanded. A SnapMirror AutoExpand
job with id "23727" was created to expand the destination FlexGroup and to trigger a SnapMirror
transfer for the SnapMirror relationship. After the SnapMirror transfer is successful, the
"healthy" field of the SnapMirror relationship will be set to "true". The job can be monitored
using either the "job show -id 23727" or "job history show -id 23727" commands.
```

5. The job expands the volume, and then you can update again.

```
cluster::*> job show -id 23727
Owning
Job ID Name Vserver Node State
------ -------------------- ---------- -------------- ----------
23727 Snapmirror Expand cluster
node1
Success
Description: SnapMirror FG Expand data_dst
```



```
cluster::*> snapmirror show -destination-path DEMO:data_dst -fields state
source-path destination-path state
----------- ---------------- ------------
DEMO:data DEMO:data_dst Snapmirrored
```

Now both FlexGroup volumes have the same number of member volumes.

```
cluster::*> vol show -vserver DEMO -volume data*
Vserver Volume Aggregate State Type Size Available Used%
--------- ------------ ------------ ---------- ---- ---------- ---------- -----
DEMO data - online RW 150TB 14.88TB 0%
DEMO data__0001 aggr1_node2 online RW 30TB 7.57TB 0%
DEMO data__0002 aggr1_node1 online RW 30TB 7.32TB 0%
DEMO data__0003 aggr1_node2 online RW 30TB 7.57TB 0%
DEMO data__0004 aggr1_node1 online RW 30TB 7.32TB 0%
DEMO data__0005 aggr1_node2 online RW 30TB 7.57TB 0%
DEMO data_dst - online DP 150TB 14.88TB 0%
DEMO data_dst__0001
aggr1_node1 online DP 30TB 7.32TB 0%
DEMO data_dst__0002
aggr1_node1 online DP 30TB 7.32TB 0%
DEMO data_dst__0003
aggr1_node2 online DP 30TB 7.57TB 0%
DEMO data_dst__0004
aggr1_node1 online DP 30TB 7.32TB 0%
DEMO data_dst__0005
aggr1_node2 online DP 30TB 7.57TB 0%
12 entries were displayed.
```

Does a High File Count Impact the Convert Process?

Short answer: No!

The sample conversion just given showed a volume with 300,000 files. 300,000 files in a volume isn't a true high file count. In this example, we convert a FlexVol volume with 500 million files to a FlexGroup volume.

```
cluster::*> vol show -vserver DEMO -volume fvconvert -fields files,files-used,is-flexgroup
vserver volume    files      files-used is-flexgroup
------- --------- ---------- ---------- ------------
DEMO    fvconvert 2040109451 502631608  false
```

Because it took so long to create that many files, we created a [FlexClone volume]() of it and split it. This approach lets us keep the origin volume intact and test without risk.

In this example, the cloning process took about 30 minutes:

```
cluster::*> vol clone split start -vserver DEMO -flexclone fvconvert -foreground true

Warning: Are you sure you want to split clone volume fvconvert in Vserver DEMO ? {y|n}: y
[Job 24230] 0% inodes processed.

cluster::*> job history show -id 24230 -fields starttime,endtime
node               record  vserver         endtime        starttime
------------------ ------- --------------- -------------- --------------
node1              2832338 cluster         12/09 10:27:08 12/09 09:58:16
```

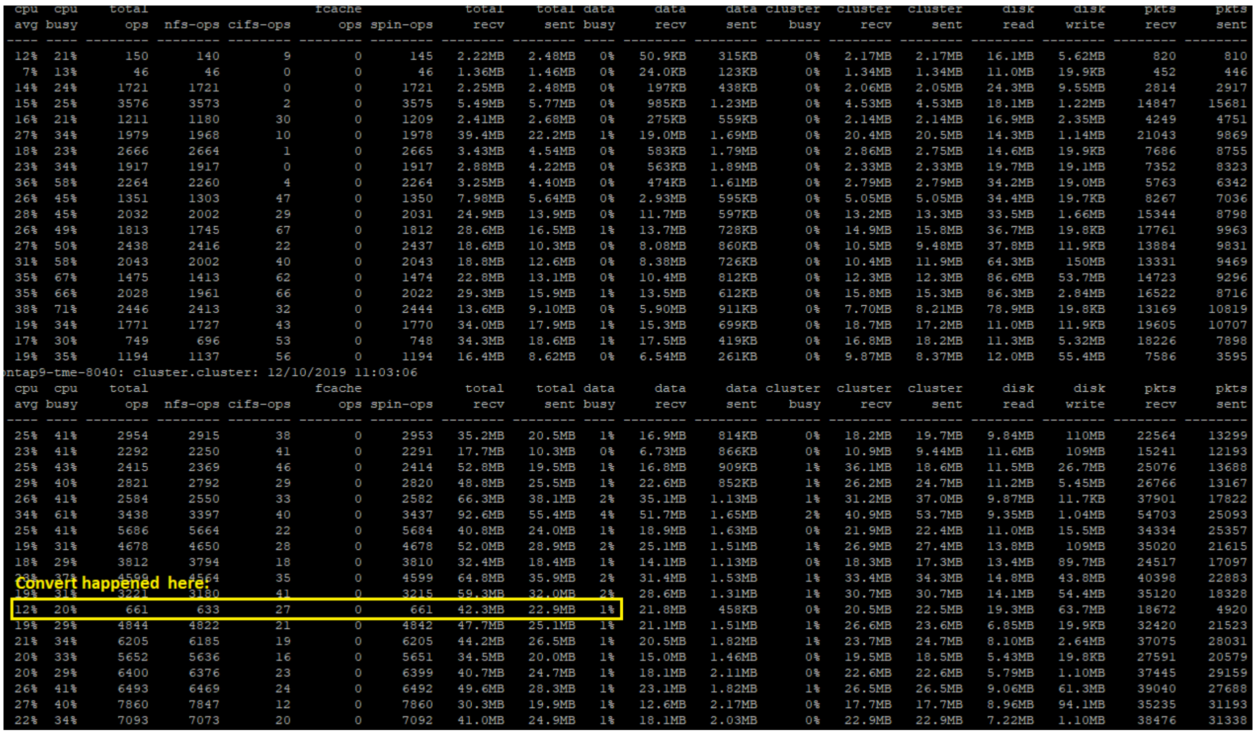
After the clone split, we ran the check. We had to run `volume clone sharing-by-split undo` to get rid of shared FlexClone blocks, which took some time, but then the check produced this output:

```
cluster::*> volume conversion start -vserver DEMO -volume fvconvert -foreground true -check-only
true
Conversion of volume "fvconvert" in Vserver "DEMO" to a FlexGroup can proceed with the following
warnings:
* After the volume is converted to a FlexGroup, it will not be possible to change it back to a
flexible volume.
```


We then ran the script that we ran earlier to generate load and watched the statistics on the cluster to see if we hit any outage. Again, the conversion took seconds (with 500 million files) and there was just a small, barely noticeable delay.

```
cluster::*> volume conversion start -vserver DEMO -volume fvconvert -foreground true

Warning: After the volume is converted to a FlexGroup, it will not be possible to change it back
to a flexible volume.
Do you want to continue? {y|n}: y
[Job 24259] Job succeeded: success
```

Figure 110) Sample statistics from conversion process.



```
 cpu  cpu    total                             fcache             total    total data     data     data cluster  cluster  cluster     disk     disk     pkts     pkts
 avg busy      ops  nfs-ops cifs-ops      ops spin-ops     recv     sent busy     recv     sent    busy     recv     sent     read    write     recv     sent
---- ---- -------- -------- -------- -------- -------- -------- -------- ---- -------- -------- ------- -------- -------- -------- -------- -------- --------
 12%  21%      150      140        9        0      145   2.22MB   2.48MB   0%   50.9KB    315KB      0%   2.17MB   2.17MB   16.1MB   5.62MB      820      810
  7%  13%       46       46        0        0       46   1.36MB   1.46MB   0%   24.0KB    123KB      0%   1.34MB   1.34MB   11.0MB   19.9KB      452      446
 14%  24%     1721     1721        0        0     1721   2.25MB   2.48MB   0%    197KB    438KB      0%   2.06MB   2.05MB   24.3MB   9.55MB     2814     2917
 15%  25%     3576     3573        2        0     3575   5.49MB   5.77MB   0%    985KB   1.23MB      0%   4.53MB   4.53MB   18.1MB   1.22MB    14847    15681
 16%  21%     1211     1180       30        0     1209   2.41MB   2.68MB   0%    275KB    559KB      0%   2.14MB   2.14MB   16.9MB   2.35MB     4249     4751
 27%  34%     1979     1968       10        0     1978   39.4MB   22.2MB   1%   19.0MB   1.69MB      0%   20.4MB   20.5MB   14.3MB   1.14MB    21043     9869
 18%  23%     2666     2664        1        0     2665   3.43MB   4.54MB   0%    583KB   1.79MB      0%   2.86MB   2.75MB   14.6MB   19.9KB     7686     8755
 23%  34%     1917     1917        0        0     1917   2.88MB   4.22MB   0%    563KB   1.89MB      0%   2.33MB   2.33MB   19.7MB   19.1MB     7352     8323
 36%  58%     2264     2260        4        0     2264   3.25MB   4.40MB   0%    474KB   1.61MB      0%   2.79MB   2.79MB   34.2MB   19.0MB     5763     6342
 26%  45%     1351     1303       47        0     1350   7.98MB   5.64MB   0%   2.93MB    595KB      0%   5.05MB   5.05MB   34.4MB   19.7KB     8267     7036
 28%  45%     2032     2002       29        0     2031   24.9MB   13.9MB   0%   11.7MB    597KB      0%   13.2MB   13.3MB   33.5MB   1.66MB    15344     8798
 26%  49%     1813     1745       67        0     1812   28.6MB   16.5MB   1%   13.7MB    728KB      0%   14.9MB   15.8MB   36.7MB   19.8KB    17761     9963
 27%  50%     2438     2416       22        0     2437   18.6MB   10.3MB   0%   8.08MB    860KB      0%   10.5MB   9.48MB   37.8MB   11.9KB    13884     9831
 31%  58%     2043     2002       40        0     2043   18.8MB   12.6MB   0%   8.38MB    726KB      0%   10.4MB   11.9MB   64.3MB    150MB    13331     9469
 35%  67%     1475     1413       62        0     1474   22.8MB   13.1MB   0%   10.4MB    812KB      0%   12.3MB   12.3MB   86.6MB   53.7MB    14723     9296
 35%  66%     2028     1961       66        0     2022   29.3MB   15.9MB   1%   13.5MB    612KB      0%   15.8MB   15.3MB   86.3MB   2.84MB    16522     8716
 38%  71%     2446     2413       32        0     2444   13.6MB   9.10MB   0%   5.90MB    911KB      0%   7.70MB   8.21MB   78.9MB   19.8KB    13169    10819
 19%  34%     1771     1727       43        0     1770   34.0MB   17.9MB   1%   15.3MB    699KB      0%   18.7MB   17.2MB   11.0MB   11.9KB    19605    10707
 17%  30%      749      696       53        0      748   34.3MB   18.6MB   1%   17.5MB    419KB      0%   16.8MB   18.2MB   11.3MB   5.32MB    18226     7898
 19%  35%     1194     1137       56        0     1194   16.4MB   8.62MB   0%   6.54MB    261KB      0%   9.87MB   8.37MB   12.0MB   55.4MB     7586     3595
ontap9-tme-8040: cluster.cluster: 12/10/2019 11:03:06
 cpu  cpu    total                             fcache             total    total data     data     data cluster  cluster  cluster     disk     disk     pkts     pkts
 avg busy      ops  nfs-ops cifs-ops      ops spin-ops     recv     sent busy     recv     sent    busy     recv     sent     read    write     recv     sent
---- ---- -------- -------- -------- -------- -------- -------- -------- ---- -------- -------- ------- -------- -------- -------- -------- -------- --------
 25%  41%     2954     2915       38        0     2953   35.2MB   20.5MB   1%   16.9MB    814KB      0%   18.2MB   19.7MB   9.84MB    110MB    22564    13299
 23%  41%     2292     2250       41        0     2291   17.7MB   10.3MB   0%   6.73MB    866KB      0%   10.9MB   9.44MB   11.6MB    109MB    15241    12193
 25%  43%     2415     2369       46        0     2414   52.8MB   19.5MB   1%   16.8MB    909KB      1%   36.1MB   18.6MB   11.5MB   26.7MB    25076    13688
 29%  40%     2821     2792       29        0     2820   48.8MB   25.5MB   1%   22.6MB    852KB      1%   26.2MB   24.7MB   11.2MB   5.45MB    26766    13167
 26%  41%     2584     2550       33        0     2582   66.3MB   38.1MB   2%   35.1MB   1.13MB      1%   31.2MB   37.0MB   9.87MB   11.7KB    37901    17822
 34%  61%     3438     3397       40        0     3437   92.6MB   55.4MB   4%   51.7MB   1.65MB      2%   40.9MB   53.7MB   9.35MB   1.04MB    54703    25093
 25%  41%     5686     5664       22        0     5684   40.8MB   24.0MB   1%   18.9MB   1.63MB      0%   21.9MB   22.4MB   11.0MB   15.5MB    34334    25357
 19%  31%     4678     4650       28        0     4678   52.0MB   28.9MB   2%   25.1MB   1.51MB      1%   26.9MB   27.4MB   13.8MB    109MB    35020    21615
 18%  29%     3812     3794       18        0     3810   32.4MB   18.4MB   1%   14.1MB   1.13MB      0%   18.3MB   17.3MB   13.4MB   89.7MB    24517    17097
 23%  37%     4599     4564       35        0     4599   64.8MB   35.9MB   2%   31.4MB   1.53MB      1%   33.4MB   34.3MB   14.8MB   43.8MB    40398    22883
 19%  31%     3221     3180       41        0     3215   59.3MB   32.0MB   2%   28.6MB   1.31MB      1%   30.7MB   30.7MB   14.1MB   54.4MB    35120    18328
 12%  20%      661      633       27        0      661   42.3MB   22.9MB   1%   21.8MB    458KB      0%   20.5MB   22.5MB   19.3MB   63.7MB    18672     4920
 19%  29%     4844     4822       21        0     4842   47.7MB   25.1MB   1%   21.1MB   1.51MB      1%   26.6MB   23.6MB   6.85MB   19.9KB    32420    21523
 21%  34%     6205     6185       19        0     6205   44.2MB   26.5MB   1%   20.5MB   1.82MB      1%   23.7MB   24.7MB   8.10MB   2.64MB    37075    28031
 20%  33%     5652     5636       16        0     5651   34.5MB   20.0MB   1%   15.0MB   1.46MB      0%   19.5MB   18.5MB   5.43MB   19.8KB    27591    20579
 20%  29%     6400     6376       23        0     6399   40.7MB   24.7MB   1%   18.1MB   2.11MB      0%   22.6MB   22.6MB   5.79MB   1.10MB    37445    29159
 26%  41%     6493     6469       24        0     6492   49.6MB   28.3MB   1%   23.1MB   1.82MB      1%   26.5MB   26.5MB   9.06MB   61.3MB    39040    27688
 27%  40%     7860     7847       12        0     7860   30.3MB   19.9MB   1%   12.6MB   2.17MB      0%   17.7MB   17.7MB   8.96MB   94.1MB    35235    31193
 22%  34%     7093     7073       20        0     7092   41.0MB   24.9MB   1%   18.1MB   2.03MB      0%   22.9MB   22.9MB   7.22MB   1.10MB    38476    31338
```


Then, as the script was running, we added new member volumes to the FlexGroup volume. Again, there was no disruption.

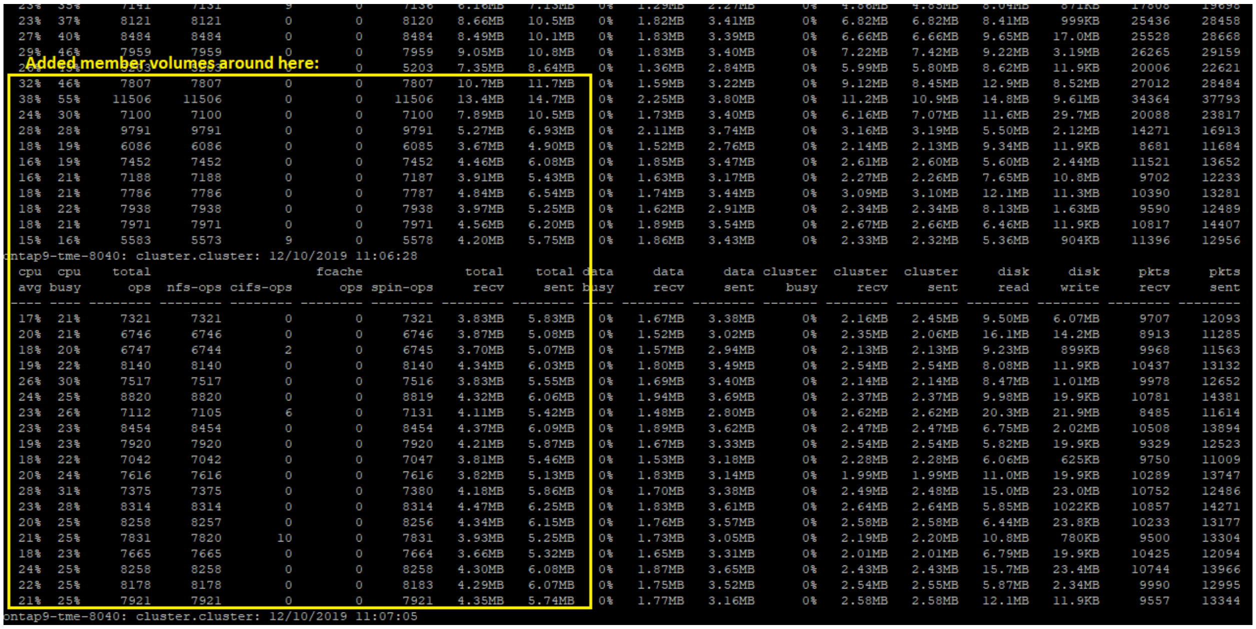
```
cluster::*> volume expand -vserver DEMO -volume fvconvert -aggr-list aggr1_node1 -aggr-list-
multiplier 3 -foreground true

Warning: The following number of constituents of size 40TB will be added to FlexGroup
"fvconvert": 3.
Do you want to continue? {y|n}: y
[Job 24261] Job succeeded: Successful
```

Then we added four more member volumes:

```
cluster::*> volume expand -vserver DEMO -volume fvconvert -aggr-list aggr1_node2 -aggr-list-
multiplier 4

Warning: The following number of constituents of size 40TB will be added to FlexGroup
"fvconvert": 4.
Do you want to continue? {y|n}: y
[Job 24264] Job succeeded: Successful
```

As an added bonus, we started to see more total IOPS for the workload. The job itself took much less time overall than when we ran it on a FlexVol volume, because the FlexGroup volume’s parallel ingest started to help the script run faster.

Figure 111) Sample statistics during conversion process – Adding member volumes.



You can view a video of this capture.

We also captured the time of completion for each job.

```
This was the job on the FlexVol before it was converted:
# python file-create.py /fvconvert/files
Starting overall work: 2019-12-09 10:32:21.966337
End overall work: 2019-12-09 12:11:15.990707
total time: 5934.024611
```

Converting the FlexVol volume to a FlexGroup volume (with added member volumes) saved some time:

```
# python file-create.py /fvconvert/files2
Starting overall work: 2019-12-10 11:02:28.621532
End overall work: 2019-12-10 12:22:48.523772
total time: 4819.95753193
```

That's savings of about 1100 seconds, or 18 minutes—which saved us around 20% of the total time of completion.

The following output shows the file distribution before the script run. Note that the first member volume has the most files, because it was previously the FlexVol volume.

```
cluster::*> volume show -vserver DEMO -volume fvconvert* -fields files,files-used
vserver volume         files      files-used
------- -------------- ---------- ----------
DEMO    fvconvert__0001 2040109451 502848737
DEMO    fvconvert__0002 2040109451 12747
DEMO    fvconvert__0003 2040109451 12749
DEMO    fvconvert__0004 2040109451 12751
```

At the end of the job, we can see that the files spread out fairly evenly.

```
cluster::*> volume show -vserver DEMO -volume fvconvert* -fields files,files-used
vserver volume         files      files-used
------- -------------- ---------- ----------
DEMO    fvconvert__0001 2040109451 506770209
DEMO    fvconvert__0002 2040109451 3345330
DEMO    fvconvert__0003 2040109451 3345330
DEMO    fvconvert__0004 2040109451 3345319
DEMO    fvconvert__0005 2040109451 3331657
```



```
DEMO     fvconvert__0006 2040109451 3331635
DEMO     fvconvert__0007 2040109451 3331657
DEMO     fvconvert__0008 2040109451 3331657
```

We ran the script again on the newly converted FlexGroup volume. This time, we wanted to see how much faster the job ran and how the files distributed on the emptier FlexVol member volumes.

Remember, when we started out, the newer member volumes all had less than 1% of files used (3.3 million of two billion possible files). The member volume that was converted from a FlexVol volume was using 25% of the total files (500 million of two billion).

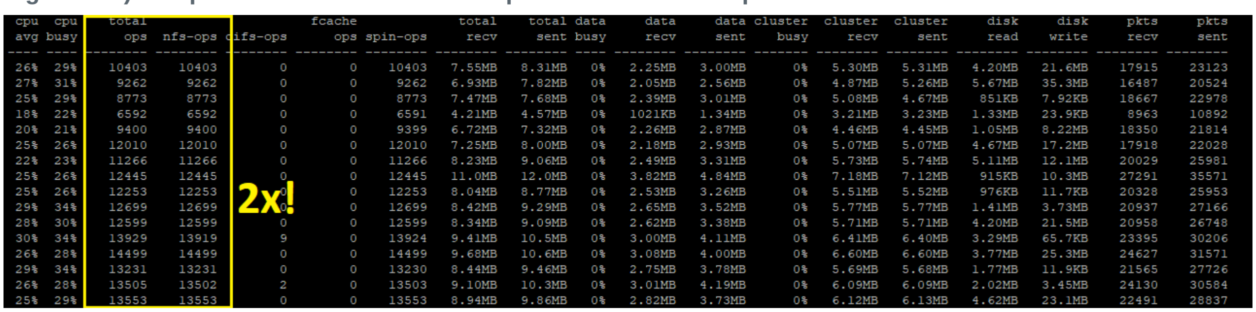
After the job ran, we saw a file count delta of about 3.2 million on the original member volume and of about 3.58 million on all the other members. We're still balancing across all member volumes, but favoring the less full ones for new file and folder creations.

```
cluster::*> volume show -vserver DEMO -volume fvconvert* -fields files,files-used
vserver volume          files      files-used
------- --------------  ---------- ----------
DEMO     fvconvert__0001 2040109451 509958440
DEMO     fvconvert__0002 2040109451 6808792
DEMO     fvconvert__0003 2040109451 6809225
DEMO     fvconvert__0004 2040109451 6806843
DEMO     fvconvert__0005 2040109451 6798959
DEMO     fvconvert__0006 2040109451 6800054
DEMO     fvconvert__0007 2040109451 6849375
DEMO     fvconvert__0008 2040109451 6801600
```

With the new FlexGroup volume, converted from a FlexVol volume, our job time dropped from 5900 seconds to 4656 seconds. We were also able to push two times the amount of IOPS:

```
# python file-create.py /fvconvert/files3
Starting overall work: 2019-12-10 13:14:26.816860
End overall work: 2019-12-10 14:32:03.565705
total time: 4656.76723099
```

Figure 112) Sample statistics of conversion process – two times performance.



cpu avg	cpu busy	total ops	nfs-ops	cifs-ops	fcache ops	spin-ops	total recv	total sent	data busy	data recv	data sent	cluster busy	cluster recv	cluster sent	disk read	disk write	pkts recv	pkts sent
26%	29%	10403	10403	0	0	10403	7.55MB	8.31MB	0%	2.25MB	3.00MB	0%	5.30MB	5.31MB	4.20MB	21.6MB	17915	23123
27%	31%	9262	9262	0	0	9262	6.93MB	7.82MB	0%	2.05MB	2.56MB	0%	4.87MB	5.26MB	5.67MB	35.3MB	16487	20524
25%	29%	8773	8773	0	0	8773	7.47MB	7.68MB	0%	2.39MB	3.01MB	0%	5.08MB	4.67MB	851KB	7.92KB	18667	22978
18%	22%	6592	6592	0	0	6591	4.21MB	4.57MB	0%	1021KB	1.34MB	0%	3.21MB	3.23MB	1.33MB	23.9KB	8963	10892
20%	21%	9400	9400	0	0	9399	6.72MB	7.32MB	0%	2.26MB	2.87MB	0%	4.46MB	4.45MB	1.05MB	8.22MB	18350	21814
25%	26%	12010	12010	0	0	12010	7.25MB	8.00MB	0%	2.18MB	2.93MB	0%	5.07MB	5.07MB	4.67MB	17.2MB	17918	22028
22%	23%	11266	11266	0	0	11266	8.23MB	9.06MB	0%	2.49MB	3.31MB	0%	5.73MB	5.74MB	5.11MB	12.1MB	20029	25981
25%	26%	12445	12445	0	0	12445	11.0MB	12.0MB	0%	3.82MB	4.84MB	0%	7.18MB	7.12MB	915KB	10.3MB	27291	35571
25%	26%	12253	12253	0	0	12253	8.04MB	8.77MB	0%	2.53MB	3.26MB	0%	5.51MB	5.52MB	976KB	11.7KB	20328	25953
29%	34%	12699	12699	0	0	12699	8.42MB	9.29MB	0%	2.65MB	3.52MB	0%	5.77MB	5.77MB	1.41MB	3.73MB	20937	27166
28%	30%	12599	12599	0	0	12599	8.34MB	9.09MB	0%	2.62MB	3.38MB	0%	5.71MB	5.71MB	4.20MB	21.5MB	20958	26748
30%	34%	13929	13919	9	0	13924	9.41MB	10.5MB	0%	3.00MB	4.11MB	0%	6.41MB	6.40MB	3.29MB	65.7KB	23395	30206
26%	28%	14499	14499	0	0	14499	9.68MB	10.6MB	0%	3.08MB	4.00MB	0%	6.60MB	6.60MB	3.77MB	25.3MB	24627	31571
29%	34%	13231	13231	0	0	13230	8.44MB	9.46MB	0%	2.75MB	3.78MB	0%	5.69MB	5.68MB	1.77MB	11.9KB	21565	27726
26%	28%	13505	13502	2	0	13503	9.10MB	10.3MB	0%	3.01MB	4.19MB	0%	6.09MB	6.09MB	2.02MB	3.45MB	24130	30584
25%	29%	13553	13553	0	0	13553	8.94MB	9.86MB	0%	2.82MB	3.73MB	0%	6.12MB	6.13MB	4.62MB	23.1MB	22491	28837

As you can see, there's an imbalance of files and data in these member volumes (much more in the original FlexVol volume), but performance is still much better than the previous FlexVol performance because work across multiple nodes is more efficient. That’s the power of the FlexGroup volume.

12.3 Migrating from Third-Party Storage to FlexGroup

When migrating from non-NetApp storage (SAN or NAS), the migration path is a file-based copy. Various methods are available to perform this migration; some are free and some are paid through third-party vendors.

For NFSv3-only data, NetApp strongly recommends the NetApp XCP Migration Tool. XCP is a free, license-based tool that can vastly improve the speed of data migration of high-file-count environments. XCP also offers robust reporting capabilities. XCP 1.5 and later versions also offer NFSv4.x and NFSv4.x ACL support, as well as being officially supported by NetApp.

Note: XCP is supported only for migration to a NetApp storage system.

For CIFS/SMB data, XCP for CIFS is available. Robocopy is a free tool, but the speed of transfer depends on using its multithreaded capabilities. Third-party providers can also perform this type of data transfer.

12.4 Migrating from NetApp Data ONTAP Operating in 7-Mode

Migrate data from Data ONTAP operating in 7-Mode to FlexGroup in one of two ways:

- Full migration of 7-Mode systems to ONTAP systems by using the copy-based or copy-free transition methodology. When using copy-free transition, the process is followed by copy-based migration of data in FlexVol volumes to FlexGroup volumes.
- Copy-based transition from a FlexVol or host-based copy from a LUN by using the previously mentioned tools for migrating from non NetApp storage to FlexGroup.

Currently, there is no migration path directly from FlexVol to FlexGroup that does not involve copy-based migrations.

Note: FlexVol to FlexGroup conversion currently cannot be performed on volumes transitioned from 7-Mode.

12.5 Migrating from FlexVol Volumes, SAN LUNs, or Infinite Volume in ONTAP

When migrating from existing ONTAP objects such as FlexVol volumes, SAN-based LUNs, or Infinite Volume, the current migration path is copy-based. The previously mentioned tools for migrating from non-NetApp storage to FlexGroup can also be used for migrating from ONTAP objects. Future releases will provide more options for migrating from FlexVol and Infinite Volume to FlexGroup volumes.

Deprecation of Infinite Volume

Starting in ONTAP 9.4, infinite volumes could no longer be created with admin privileges. This step prepared for the eventual removal of Infinite Volume support in ONTAP 9.5 and later. Starting in ONTAP 9.5, infinite volumes can no longer be created or modified, and an infinite volume cannot have protocol access.

If infinite volumes are present in an ONTAP cluster and you attempt to upgrade, the ONTAP compatibility checker prevents the upgrade from completing and warns of existing infinite volumes. Be sure to use the NetApp Upgrade Advisor when planning your ONTAP upgrade.

12.6 XCP Migration Tool

The NetApp XCP Migration Tool is free and was designed specifically for scoping, migration, and management of large sets of unstructured NAS data. The initial version was NFSv3 only, but a CIFS version is now available. To use the tool, download it and request a free license (for software tracking purposes only).

XCP addresses the challenges that high-file-count environments have with metadata operation and data migration performance by using a multicore, multichannel I/O streaming engine that can process many requests in parallel.

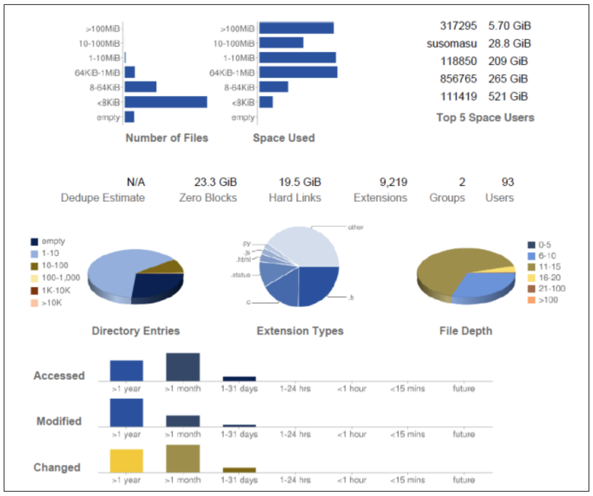
These requests include the following:

- Data migration
- File or directory listings (a high-performance, flexible alternative to `ls`)
- Space reporting (a high-performance, flexible alternative to `du`)

XCP has sometimes reduced the length of data migration by 20 to 30 times for high-file-count environments. In addition, XCP has reduced the file list time for 165 million files from 9 days on a competitor's system to 30 minutes on NetApp technology—a performance improvement of 400 times.

XCP also gives some handy reporting graphs, as shown in Figure 113.

Figure 113) XCP reporting graphs.



For more information, see the official XCP website at http://xcp.netapp.com.

12.7 Using XCP to Scan Files Before Migration

When deploying a FlexGroup volume, evaluate the file system and structure to help you determine initial sizing considerations and the best way to lay out member volumes. In high-file-count environments, this can be time consuming and tedious. XCP allows you to scan files and export to the CSV or XML format to easily review your file system.

The following example shows a FlexGroup volume with over a million files. Ideally, we don't want to spend much time analyzing these files.

```
cluster::> vol show -vserver DEMO -fields files,files-used -volume flexgroup_16
vserver volume       files     files-used
------- ------------ --------- ----------
DEMO    flexgroup_16 318766960 1103355
```

To streamline this process, you can use `xcp scan` to get file information. Here's a sample command:

```
C:\> xcp scan -stats \\demo\flexgroup > C:\destination.csv
```

When you do this, the client scans the files and adds information to a comma-separated values (CSV) document. This document shows information such as the following:

- Maximum and average values for size, depth of directory, and dirsize

```
== Maximum Values ==
      Size       Depth     Namelen     Dirsize
    340MiB           9          86         500

== Average Values ==
      Size       Depth     Namelen     Dirsize
   1.61KiB           4           6          11
```

- Top file extensions

```
== Top File Extensions ==
```



```
             .docx        .png       .pptx        .pdf        .css       other
   1000038       260         175         128          91          33         219
```

- Number of files, broken down by size ranges

```
== Number of files ==
     empty      <8KiB    8-64KiB 64KiB-1MiB    1-10MiB  10-100MiB    >100MiB
         8    1000215        156         288        265         10          2
```

- Space used by size range

```
== Space used ==
     empty      <8KiB    8-64KiB 64KiB-1MiB    1-10MiB  10-100MiB    >100MiB
         0    28.7MiB    3.94MiB     124MiB     695MiB     272MiB     453MiB
```

- Directory entries, broken down by file counts

```
== Directory entries ==
     empty       1-10     10-100     100-1K     1K-10K       >10k
         7     100118         30        200
```

- Directory depth ranges

```
== Depth ==
       0-5       6-10      11-15      16-20     21-100       >100
   1100966        333
```

- Modified and created date ranges

```
== Modified ==
   >1 year   >1 month  1-31 days   1-24 hrs    <1 hour   <15 mins     future
       579    1100559         11                   150

== Created ==
   >1 year   >1 month  1-31 days   1-24 hrs    <1 hour   <15 mins     future
              1100210                             1089
```

- A summary of the file structure, including total file count, total directories, symlinks, junctions, and total space used

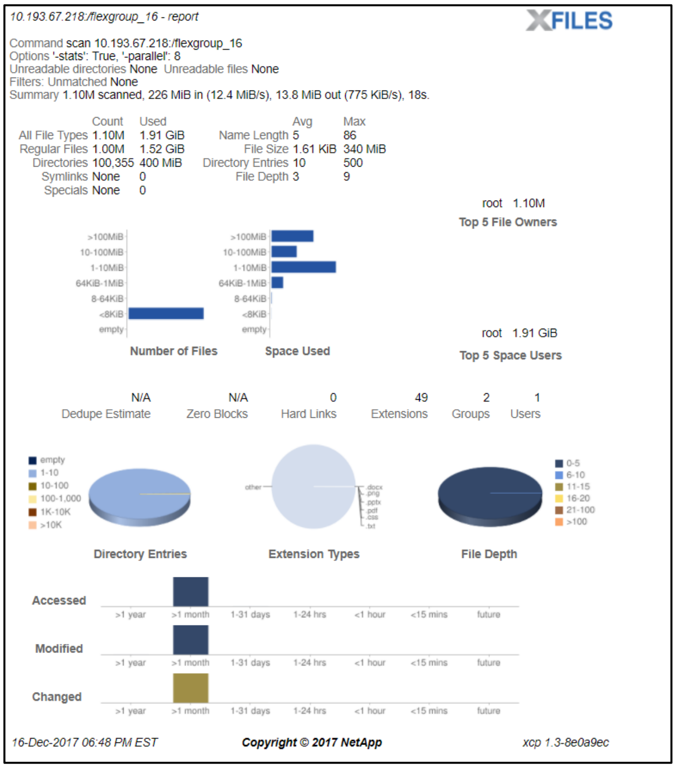
```
Total count: 1101299
Directories: 100355
Regular files: 1000944
Symbolic links:
Junctions:
Special files:
Total space for regular files: 1.54GiB
Total space for directories: 0
Total space used: 1.54GiB
1,101,299 scanned, 0 errors, 26m34s
```

You can also use XCP over NFS to scan CIFS volumes and get more robust reporting and the ability to export to HTML, which presents the data in graphical format.

For example, the following command creates the report shown in Figure 114:

```
xcp scan -stats -html demo:/flexgroup_16 > /flexgroup.html
```


Figure 114) XCP report.



Using XCP to scan file systems provides average file size information, largest file size, capacity and file count measurements for the top five file owners, and much more. These statistics are available only in the NFS version of XCP, but you can still run NFS scans on datasets that only do SMB traffic by setting up a virtual machine that can use NFS.

Using XCP to Run Disk Usage (du) Scans

One common complaint is that in high-file-count environments, running commands like `du` can take an exceedingly long time. For example, this `du` command ran on a FlexGroup volume with 1,101,002 files and folders and took 21 minutes and 22.600 seconds.

With XCP, this command scanned the same dataset in 22.852 seconds with the same client:

```
[root@centos7 ~]# xcp -duk DEMO:/FGlocal 2>/dev/null | egrep -v '.*?/.*?/'
```

13 Event Management System Examples

Example of Maxdirsize Message

```
Message Name: wafl.dir.size.max
Severity: ERROR

Corrective Action: Use the "volume file show-inode" command with the file ID and volume name
information to find the file path. Reduce the number of files in the directory. If not possible,
use the (privilege:advanced) option "volume modify -volume vol_name -maxdir-size new_value" to
increase the maximum number of files per directory. However, doing so could impact system
performance. If you need to increase the maximum directory size, work with technical support.

Description: This message occurs after a directory has reached its maximum directory size
(maxdirsize) limit.
                                Supports SNMP trap: true
```



```
                                  Destinations: -
            Number of Drops Between Transmissions: 0
Dropping Interval (Seconds) Between Transmissions: 0


Message Name: wafl.dir.size.max.warning
Severity: ERROR

Corrective Action: Use the "volume file show-inode" command with the file ID and volume name
information to find the file path. Reduce the number of files in the directory. If not possible,
use the (privilege:advanced) option "volume modify -volume vol_name -maxdir-size new_value" to
increase the maximum number of files per directory. However, doing so could impact system
performance. If you need to increase the maximum directory size, work with technical support.

Description: This message occurs when a directory has reached or surpassed 90% of its current
maximum directory size (maxdirsize) limit, and the current maxdirsize is less than the default
maxdirsize, which is 1% of total system memory.
                               Supports SNMP trap: true
                                     Destinations: -
            Number of Drops Between Transmissions: 0
Dropping Interval (Seconds) Between Transmissions: 0


Message Name: wafl.dir.size.warning
Severity: ERROR

Corrective Action: Use the "volume file show-inode" command with the file ID and volume name
information to find the file path. Reduce the number of files in the directory. If not possible,
use the (privilege:advanced) option "volume modify -volume vol_name -maxdir-size new_value" to
increase the maximum number of files per directory. However, doing so could impact system
performance. If you need to increase the maximum directory size, work with technical support.

Description: This mesaage occurs when a directory surpasses 90% of its current maximum directory
size (maxdirsize) limit.
                               Supports SNMP trap: true
                                     Destinations: -
            Number of Drops Between Transmissions: 0
Dropping Interval (Seconds) Between Transmissions: 0
```


Examples of Capacity-Related Event Management System Messages

```
Message Name: monitor.volume.full
Severity: DEBUG
Corrective Action: (NONE)

Description: This message occurs when one or more file systems are full, typically indicating at
least 98% full. This event is accompanied by global health monitoring messages for the customer.
The space usage is computed based on the active file system size and is computed by subtracting
the value of the "Snapshot Reserve" field from the value of the "Used" field of the "volume show-
space" command. The volume/aggregate can be over 100% full due to space used or reserved by
metadata. A value greater than 100% might cause Snapshot(tm) copy space to become unavailable or
cause the volume to become logically overallocated. See the "vol.log.overalloc" EMS message for
more information.
                               Supports SNMP trap: true
                                     Destinations: -
            Number of Drops Between Transmissions: 0
Dropping Interval (Seconds) Between Transmissions: 0
```

```
Message Name: monitor.volume.nearlyFull
Severity: ALERT
Corrective Action: Create space by increasing the volume or aggregate sizes, or by deleting data
or deleting Snapshot(R) copies. To increase a volume's size, use the "volume size" command. To
delete a volume's Snapshot(R) copies, use the "volume snapshot delete" command. To increase an
aggregate's size, add disks by using the "storage aggregate add-disks" command. Aggregate
Snapshot(R) copies are deleted automatically when the aggregate is full.
                                      Description: This message occurs when one or more file
systems are nearly full, typically indicating at least 95% full. This event is accompanied by
global health monitoring messages for the customer. The space usage is computed based on the
active file system size and is computed by subtracting the value of the "Snapshot Reserve" field
from the value of the "Used" field of the "volume show-space" command.
                               Supports SNMP trap: true
                                     Destinations: -
            Number of Drops Between Transmissions: 0
Dropping Interval (Seconds) Between Transmissions: 0
```

```
Message Name: monitor.volume.ok
Severity: DEBUG
Corrective Action: (UNKNOWN)

Description: The previously-reported volume full condition is fixed. * We log this event, as well
as the other monitor.volume events, at LOG_DEBUG level to avoid spamming the messages file with
events which are already being reported as part of the global health messages.
                               Supports SNMP trap: true
                                     Destinations: -
            Number of Drops Between Transmissions: 0
Dropping Interval (Seconds) Between Transmissions: 0
```



```
Message Name: monitor.volumes.one.ok
Severity: DEBUG
Corrective Action: (NONE)

Description: This message occures when one file system that was nearly full (usually this means
>= 95% full) is now OK. This event and other "monitor.volume" events are logged at LOG_DEBUG
level to avoid spamming the messages file with events that are already being reported as part of
the global health messages. The space usage is computed based on the active file system size and
is computed by subtracting the value of the "Snapshot Reserve" field from the value of the "Used"
field of the "volume show-space" command.
                               Supports SNMP trap: true
                                     Destinations: -
            Number of Drops Between Transmissions: 0
Dropping Interval (Seconds) Between Transmissions: 0
```

```
Message Name: vol.log.overalloc
Severity: ALERT
Corrective Action: Create space by increasing the volume or aggregate size, deleting data,
deleting Snapshot(R) copies, or changing the provisioning from thick to thin. To increase a
volume's size, use the "volume size" command. To delete a volume's Snapshot(R) copies, use the
"volume snapshot delete" command. To change provisioning in a volume, reserved files can be
unreserved by using the "volume file reservation" command. To increase an aggregate's size, add
disks by using the "storage aggregate add-disks" command. Aggregate Snapshot(R) copies are
deleted automatically when the aggregate is full. To change provisioning of a volume in an
aggregate, change the volume guarantee from "volume" to "none" by using the "space-guarantee"
field of the "volume modify" command.

Description: This message occurs when the volume or aggregate allocates more space than it can
honor by way of reservations, or the aggregate has allocated more space than it can honor by way
of guarantees. If the reserved or guaranteed space is consumed, there is insufficient physical
space, which can cause the volume or aggregate to be taken offline.
                               Supports SNMP trap: true
                                     Destinations: -
            Number of Drops Between Transmissions: 0
Dropping Interval (Seconds) Between Transmissions: 0
```


14 Command Examples

FlexGroup Capacity Commands

```
cluster::*> aggr show-space -instance -aggregate aggr1_node1

                          Aggregate Name: aggr1_node1
                       Volume Footprints: 2.05TB
               Volume Footprints Percent: 26%
Total Space for Snapshot Copies in Bytes: 0B
      Space Reserved for Snapshot Copies: 0%
                      Aggregate Metadata: 15.20MB
              Aggregate Metadata Percent: 0%
                              Total Used: 2.05TB
                      Total Used Percent: 26%
                                    Size: 7.86TB
               Snapshot Reserve Unusable: -
       Snapshot Reserve Unusable Percent: -
                Total Physical Used Size: 143.7GB
                Physical Used Percentage: 2%

                          Aggregate Name: aggr1_node2
                       Volume Footprints: 2.02TB
               Volume Footprints Percent: 26%
Total Space for Snapshot Copies in Bytes: 0B
      Space Reserved for Snapshot Copies: 0%
                      Aggregate Metadata: 8.63MB
              Aggregate Metadata Percent: 0%
                              Total Used: 2.02TB
                      Total Used Percent: 26%
                                    Size: 7.86TB
               Snapshot Reserve Unusable: -
       Snapshot Reserve Unusable Percent: -
                Total Physical Used Size: 69.71GB
                Physical Used Percentage: 1%
2 entries were displayed.
```

```
cluster::*> volume show-space -vserver SVM -volume flexgroup__*
      Vserver : SVM
      Volume  : flexgroup__0001

      Feature                                   Used      Used%
      --------------------------------     ----------    ------
      User Data                                57.06MB       0%
      Filesystem Metadata                       3.51MB       0%
      Inodes                                   87.26MB       0%
      Snapshot Reserve                           512GB       5%
      Deduplication                               12KB       0%
      Performance Metadata                        48KB       0%

      Total Used                               512.1GB       5%
      Total Physical Used                      148.3MB       0%


      Vserver : SVM
      Volume  : flexgroup__0002

      Feature                                   Used      Used%
      --------------------------------     ----------    ------
      User Data                                57.03MB       0%
      Filesystem Metadata                       4.66MB       0%
      Inodes                                   83.66MB       0%
      Snapshot Reserve                           512GB       5%
      Deduplication                               20KB       0%
      Performance Metadata                        44KB       0%

      Total Used                               512.1GB       5%
```



```
Total Physical Used                  145.7MB         0%


Vserver : SVM
Volume  : flexgroup__0003

Feature                                   Used      Used%
--------------------------------    ----------     ------
User Data                              57.02MB         0%
Filesystem Metadata                     3.66MB         0%
Inodes                                 84.55MB         0%
Snapshot Reserve                         512GB         5%
Deduplication                             12KB         0%
Performance Metadata                      44KB         0%

Total Used                             512.1GB         5%
Total Physical Used                    145.6MB         0%


Vserver : SVM
Volume  : flexgroup__0004

Feature                                   Used      Used%
--------------------------------    ----------     ------
User Data                              57.19MB         0%
Filesystem Metadata                     8.93MB         0%
Inodes                                 82.09MB         0%
Snapshot Reserve                         512GB         5%
Deduplication                             12KB         0%
Performance Metadata                      44KB         0%

Total Used                             512.1GB         5%
Total Physical Used                    148.5MB         0%


Vserver : SVM
Volume  : flexgroup__0005

Feature                                   Used      Used%
--------------------------------    ----------     ------
User Data                               3.99GB         0%
Filesystem Metadata                     4.88MB         0%
Inodes                                 83.54MB         0%
Snapshot Reserve                         512GB         5%
Deduplication                             12KB         0%
Performance Metadata                      52KB         0%

Total Used                             516.1GB         5%
Total Physical Used                     4.08GB         0%


Vserver : SVM
Volume  : flexgroup__0006

Feature                                   Used      Used%
--------------------------------    ----------     ------
User Data                              57.04MB         0%
Filesystem Metadata                     3.50MB         0%
Inodes                                 87.26MB         0%
Snapshot Reserve                         512GB         5%
Deduplication                             12KB         0%
Performance Metadata                      44KB         0%

Total Used                             512.1GB         5%
Total Physical Used                    148.2MB         0%


Vserver : SVM
Volume  : flexgroup__0007

Feature                                   Used      Used%
```



```
    -------------------------------     ----------     ------
    User Data                              57.02MB         0%
    Filesystem Metadata                     3.50MB         0%
    Inodes                                 85.03MB         0%
    Snapshot Reserve                         512GB         5%
    Deduplication                             12KB         0%
    Performance Metadata                      44KB         0%

    Total Used                             512.1GB         5%
    Total Physical Used                    145.9MB         0%


    Vserver : SVM
    Volume  : flexgroup__0008

    Feature                                   Used      Used%
    -------------------------------     ----------     ------
    User Data                              57.03MB         0%
    Filesystem Metadata                     3.52MB         0%
    Inodes                                 86.12MB         0%
    Snapshot Reserve                         512GB         5%
    Deduplication                             12KB         0%
    Performance Metadata                      44KB         0%

    Total Used                             512.1GB         5%
    Total Physical Used                    147.0MB         0%
```

```
cluster::> vol show -is-constituent true -volume flexgroup__*
Vserver   Volume       Aggregate    State      Type       Size  Available Used%
--------- ------------ ------------ ---------- ---- ---------- ---------- -----
SVM       flexgroup__0001
                       aggr1_node1  online     RW         10TB     5.05TB   49%
SVM       flexgroup__0002
                       aggr1_node2  online     RW         10TB     5.08TB   49%
SVM       flexgroup__0003
                       aggr1_node1  online     RW         10TB     5.05TB   49%
SVM       flexgroup__0004
                       aggr1_node2  online     RW         10TB     5.08TB   49%
SVM       flexgroup__0005
                       aggr1_node1  online     RW         10TB     5.05TB   49%
SVM       flexgroup__0006
                       aggr1_node2  online     RW         10TB     5.08TB   49%
SVM       flexgroup__0007
                       aggr1_node1  online     RW         10TB     5.05TB   49%
SVM       flexgroup__0008
                       aggr1_node2  online     RW         10TB     5.08TB   49%
8 entries were displayed.

cluster::*> storage aggregate show -aggregate aggr1* -fields usedsize,size,percent-used -sort-by
percent-used
aggregate   percent-used size   usedsize
----------- ------------ ------ --------
aggr1_node1 26%          7.86TB 2.05TB
aggr1_node2 26%          7.86TB 2.02TB
2 entries were displayed.
```


Example of statistics show-periodic Command for Entire Cluster

```
cluster::*> statistics show-periodic
cluster: cluster.cluster: 11/30/2016 11:49:46
  cpu  cpu    total                     fcache             total    total data     data     data
cluster  cluster  cluster     disk     disk     pkts     pkts
  avg busy      ops  nfs-ops cifs-ops      ops spin-ops     recv     sent busy     recv     sent
busy     recv     sent     read    write     recv     sent
 ---- ---- -------- -------- -------- -------- -------- -------- -------- ---- -------- --------
------- -------- -------- -------- -------- -------- --------
   5%   5%        0        0        0        0        0   65.3KB   64.4KB   0%   2.22KB   1.13KB
0%   62.7KB   63.2KB    489KB    407KB       91       83
   5%   5%        0        0        0        0        0   62.5KB   61.6KB   0%   1.28KB     767B
0%   61.0KB   60.9KB   23.8KB   23.8KB       64       60
   4%   5%        0        0        0        0        0   62.3KB   61.3KB   0%   1.43KB     708B
0%   60.7KB   60.7KB   15.8KB   15.8KB       69       58
cluster: cluster.cluster: 11/30/2016 11:49:53
  cpu  cpu    total                     fcache             total    total data     data     data
cluster  cluster  cluster     disk     disk     pkts     pkts
  avg busy      ops  nfs-ops cifs-ops      ops spin-ops     recv     sent busy     recv     sent
busy     recv     sent     read    write     recv     sent
 ---- ---- -------- -------- -------- -------- -------- -------- -------- ---- -------- --------
------- -------- -------- -------- -------- -------- --------
Minimums:
   4%   5%        0        0        0        0        0   62.3KB   61.3KB   0%   1.28KB     708B
0%   60.7KB   60.7KB   15.8KB   15.8KB       64       58
Averages for 3 samples:
   4%   5%        0        0        0        0        0   63.4KB   62.4KB   0%   1.64KB     877B
0%   61.5KB   61.6KB    176KB    149KB       74       67
Maximums:
   5%   5%        0        0        0        0        0   65.3KB   64.4KB   0%   2.22KB   1.13KB
0%   62.7KB   63.2KB    489KB    407KB       91       83
```

Real-Time SVM-Level statistics show-periodic for NFSv3 Read and Write Operations

```
cluster::*> statistics show-periodic -instance SVM -interval 2 -iterations 0 -summary true -
vserver SVM -object nfsv3 -counter nfsv3_ops|nfsv3_read_ops|nfsv3_write_ops
cluster: nfsv3.SVM: 11/30/2016 13:29:57
             nfsv3    nfsv3
    nfsv3     read    write    Complete    Number of
      ops      ops      ops Aggregation Constituents
 -------- -------- -------- ----------- ------------
     2360        0      697      Yes      16
     2245        0      652      Yes      16
     2126        0      629      Yes      16
cluster: nfsv3.SVM: 11/30/2016 13:30:04
             nfsv3    nfsv3
    nfsv3     read    write    Complete    Number of
      ops      ops      ops Aggregation Constituents
 -------- -------- -------- ----------- ------------
Minimums:
     2126        0      629       -       -
Averages for 3 samples:
     2243        0      659       -       -
Maximums:
     2360        0      697       -       -
```


Real-Time FlexGroup Local and Remote Statistics

```
cluster::*> statistics show-periodic -instance 0 -interval 2 -iterations 0 -summary true -object
flexgroup -counter
cat1_tld_local|cat1_tld_remote|cat2_hld_local|cat2_hld_remote|cat3_dir_local|cat3_dir_remote|cat4
_fil_local|cat4_fil_remote
cluster: flexgroup.0: 11/30/2016 13:34:55
     cat1     cat1     cat2     cat2     cat3     cat3     cat4     cat4
      tld      tld      hld      hld      dir      dir      fil      fil    Complete    Number of
    local   remote    local   remote    local   remote    local   remote Aggregation Constituents
 -------- -------- -------- -------- -------- -------- -------- -------- ----------- ------------
        1        0       17      113        0        0      619        0       n/a      n/a
        0        1       17      114        0        0      654        0       n/a      n/a
        0        2       17      112        0        0      647        0       n/a      n/a
cluster: flexgroup.0: 11/30/2016 13:35:02
     cat1     cat1     cat2     cat2     cat3     cat3     cat4     cat4
      tld      tld      hld      hld      dir      dir      fil      fil    Complete    Number of
    local   remote    local   remote    local   remote    local   remote Aggregation Constituents
 -------- -------- -------- -------- -------- -------- -------- -------- ----------- ------------
Minimums:
        0        0       17      112        0        0      619        0        -        -
Averages for 3 samples:
        0        1       17      113        0        0      640        0        -        -
Maximums:
        1        2       17      114        0        0      654        0        -        -
```

Example of Creating a FlexGroup Volume and Specifying Fewer Member Volumes Than the Default Value

This command creates a 10TB FlexGroup volume with two 5TB member volumes across two nodes.

```
cluster::> volume create -vserver DEMO -volume flexgroup -aggr-list aggr1_node1,aggr1_node2 -
aggr-list-multiplier 1 -junction-path /flexgroup -size 10t

Warning: The FlexGroup "flexgroup" will be created with the following number of constituents of
size 5TB: 2.

Do you want to continue? {y|n}: y
```

Note: The `-aggr-list` flag must be used to make sure that the volume is a FlexGroup volume.

Sample REST API for Creating a FlexGroup Volume

The following REST API example creates a 2TB, eight-member thin-provisioned FlexGroup volume across a single aggregate.

```
{
  "aggregates": [
    {
      "name": "aggr1_node1"
    }
  ],
  "constituents_per_aggregate": 8,
  "efficiency": {
    "compaction": "inline",
    "compression": "inline",
    "cross_volume_dedupe": "inline",
    "dedupe": "inline"
  },
  "guarantee": {
    "type": "none"
  },
  "name": "RESTAPI_FG",
  "nas": {
    "export_policy": {
      "id": 42949672961,
      "name": "default"
    },
```



```
    "gid": 0,
    "path": "/RESTAPI_FG",
    "security_style": "unix",
    "uid": 0,
    "unix_permissions": 755
  },
  "size": "2T",
  "style": "flexgroup",
  "svm": {
    "name": "DEMO",
    "uuid": "7e3cc08e-d9b3-11e6-85e2-00a0986b1210"
  }
}
```


This is what the FlexGroup looks like after it's created:

```
cluster::*> vol show -vserver DEMO -volume REST*
Vserver   Volume       Aggregate    State      Type       Size  Available Used%
--------- ------------ ------------ ---------- ---- ---------- ---------- -----
DEMO      RESTAPI_FG   -            online     RW          2TB     1.90TB    0%
DEMO      RESTAPI_FG__0001
                       aggr1_node1  online     RW        256GB    243.1GB    0%
DEMO      RESTAPI_FG__0002
                       aggr1_node1  online     RW        256GB    243.1GB    0%
DEMO      RESTAPI_FG__0003
                       aggr1_node1  online     RW        256GB    243.1GB    0%
DEMO      RESTAPI_FG__0004
                       aggr1_node1  online     RW        256GB    243.1GB    0%
DEMO      RESTAPI_FG__0005
                       aggr1_node1  online     RW        256GB    243.1GB    0%
DEMO      RESTAPI_FG__0006
                       aggr1_node1  online     RW        256GB    243.1GB    0%
DEMO      RESTAPI_FG__0007
                       aggr1_node1  online     RW        256GB    243.1GB    0%
DEMO      RESTAPI_FG__0008
                       aggr1_node1  online     RW        256GB    243.1GB    0%
9 entries were displayed.
```

To include more than one aggregate in the list, use this REST API as an example:


```
{
  "aggregates": [
    { "name": "aggr1_node1" },{ "name": "aggr1_node2" }
  ],
  "efficiency": {
    "compaction": "inline",
    "compression": "inline",
    "cross_volume_dedupe": "inline",
    "dedupe": "inline"
  },
  "guarantee": {
    "type": "none"
  },
  "name": "RESTAPI_FG3",
  "nas": {
    "export_policy": {
      "id": 42949672961,
      "name": "default"
    },
    "gid": 0,
    "path": "/RESTAPI_FG3",
    "security_style": "unix",
    "uid": 0,
    "unix_permissions": 755
  },
  "size": "2T",
  "style": "flexgroup",
  "svm": {
    "name": "DEMO",
    "uuid": "7e3cc08e-d9b3-11e6-85e2-00a0986b1210"
```



```
  }
}
```

This is how it looks:

```
cluster::*> vol show -vserver DEMO -volume *FG3*
Vserver   Volume       Aggregate    State      Type       Size  Available Used%
--------- ------------ ------------ ---------- ---- ---------- ---------- -----
DEMO      RESTAPI_FG3  -            online     RW          2TB     1.90TB    0%
DEMO      RESTAPI_FG3__0001
                       aggr1_node1  online     RW        256GB    243.1GB    0%
DEMO      RESTAPI_FG3__0002
                       aggr1_node2  online     RW        256GB    243.1GB    0%
DEMO      RESTAPI_FG3__0003
                       aggr1_node1  online     RW        256GB    243.1GB    0%
DEMO      RESTAPI_FG3__0004
                       aggr1_node2  online     RW        256GB    243.1GB    0%
DEMO      RESTAPI_FG3__0005
                       aggr1_node1  online     RW        256GB    243.1GB    0%
DEMO      RESTAPI_FG3__0006
                       aggr1_node2  online     RW        256GB    243.1GB    0%
DEMO      RESTAPI_FG3__0007
                       aggr1_node1  online     RW        256GB    243.1GB    0%
DEMO      RESTAPI_FG3__0008
                       aggr1_node2  online     RW        256GB    243.1GB    0%
9 entries were displayed.
```

This REST API creates a four-member FlexGroup volume y using the “style” option and does not specify the `constituents_per_aggregate` option.


```
{
  "aggregates": [
    {
      "name": "aggr1_node1"
    }
  ],
  "efficiency": {
    "compaction": "inline",
    "compression": "inline",
    "cross_volume_dedupe": "inline",
    "dedupe": "inline"
  },
  "guarantee": {
    "type": "none"
  },
  "name": "RESTAPI_FG2",
  "nas": {
    "export_policy": {
      "id": 42949672961,
      "name": "default"
    },
    "gid": 0,
    "path": "/RESTAPI_FG2",
    "security_style": "unix",
    "uid": 0,
    "unix_permissions": 755
  },
  "size": "2T",
  "style": "flexgroup",
  "svm": {
    "name": "DEMO",
    "uuid": "7e3cc08e-d9b3-11e6-85e2-00a0986b1210"
  }
}
```


And this is the resulting FlexGroup:

```
cluster::*> vol show -vserver DEMO -volume RESTAPI_FG2*
Vserver   Volume       Aggregate    State      Type       Size  Available Used%
```



```
--------- ------------ ------------ ---------- ---- ---------- ---------- -----
DEMO      RESTAPI_FG2  -            online     RW          2TB     1.90TB    0%
DEMO      RESTAPI_FG2__0001
                       aggr1_node1  online     RW        512GB    486.3GB    0%
DEMO      RESTAPI_FG2__0002
                       aggr1_node1  online     RW        512GB    486.3GB    0%
DEMO      RESTAPI_FG2__0003
                       aggr1_node1  online     RW        512GB    486.3GB    0%
DEMO      RESTAPI_FG2__0004
                       aggr1_node1  online     RW        512GB    486.3GB    0%
5 entries were displayed.
```

Example of Increasing a FlexGroup Volume's Size

```
cluster::*> volume show -vserver SVM -volume flexgroup*
SVM       flexgroup    -            online     RW      70.20TB    10.14TB   85%
SVM       flexgroup__0001
                       aggr1_node1  online     RW         10TB     5.06TB   49%
SVM       flexgroup__0002
                       aggr1_node2  online     RW         10TB     5.08TB   49%
SVM       flexgroup__0003
                       aggr1_node1  online     RW         10TB     5.06TB   49%
SVM       flexgroup__0004
                       aggr1_node2  online     RW         10TB     5.08TB   49%
SVM       flexgroup__0005
                       aggr1_node1  online     RW         10TB     5.06TB   49%
SVM       flexgroup__0006
                       aggr1_node2  online     RW         10TB     5.08TB   49%
SVM       flexgroup__0007
                       aggr1_node1  online     RW         10TB     5.06TB   49%
SVM       flexgroup__0008
                       aggr1_node2  online     RW         10TB     5.08TB   49%

cluster::*> vol size -vserver SVM -volume flexgroup -new-size 100t
vol size: Volume "SVM:flexgroup" size set to 100t.

cluster::*> volume show -vserver SVM -volume flexgroup*
Vserver   Volume       Aggregate    State      Type       Size  Available Used%
--------- ------------ ------------ ---------- ---- ---------- ---------- -----
SVM       flexgroup    -            online     RW        100TB    10.14TB   89%

SVM       flexgroup__0001
                       aggr1_node1  online     RW      12.50TB     5.06TB   59%
SVM       flexgroup__0002
                       aggr1_node2  online     RW      12.50TB     5.08TB   59%
SVM       flexgroup__0003
                       aggr1_node1  online     RW      12.50TB     5.06TB   59%
SVM       flexgroup__0004
                       aggr1_node2  online     RW      12.50TB     5.08TB   59%
SVM       flexgroup__0005
                       aggr1_node1  online     RW      12.50TB     5.06TB   59%
SVM       flexgroup__0006
                       aggr1_node2  online     RW      12.50TB     5.08TB   59%
SVM       flexgroup__0007
                       aggr1_node1  online     RW      12.50TB     5.06TB   59%
SVM       flexgroup__0008
                       aggr1_node2  online     RW      12.50TB     5.08TB   59%
```


Example of Expanding a FlexGroup Volume

```
cluster::*> volume show -vserver SVM -volume flexgroup4*
Vserver   Volume       Aggregate    State      Type       Size  Available Used%
--------- ------------ ------------ ---------- ---- ---------- ---------- -----
SVM       flexgroup4TB -            online     RW          4TB     3.78TB    5%
SVM       flexgroup4TB__0001
                       aggr1_node1  online     RW        512GB    485.5GB    5%
SVM       flexgroup4TB__0002
                       aggr1_node2  online     RW        512GB    481.2GB    6%
SVM       flexgroup4TB__0003
                       aggr1_node1  online     RW        512GB    481.5GB    5%
SVM       flexgroup4TB__0004
                       aggr1_node2  online     RW        512GB    485.5GB    5%
SVM       flexgroup4TB__0005
                       aggr1_node1  online     RW        512GB    485.5GB    5%
SVM       flexgroup4TB__0006
                       aggr1_node2  online     RW        512GB    485.5GB    5%
SVM       flexgroup4TB__0007
                       aggr1_node1  online     RW        512GB    485.5GB    5%
SVM       flexgroup4TB__0008
                       aggr1_node2  online     RW        512GB    485.5GB    5%

cluster::*> volume expand -vserver SVM -volume flexgroup4TB -aggr-list aggr1_node1,aggr1_node2 -
aggr-list-multiplier 4


cluster::*> volume show -vserver SVM -volume flexgroup4*
Vserver   Volume       Aggregate    State      Type       Size  Available Used%
--------- ------------ ------------ ---------- ---- ---------- ---------- -----
SVM       flexgroup4TB -            online     RW          8TB     7.78TB    1%
SVM       flexgroup4TB__0001
                       aggr1_node1  online     RW        512GB    485.5GB    1%
SVM       flexgroup4TB__0002
                       aggr1_node2  online     RW        512GB    481.2GB    1%
SVM       flexgroup4TB__0003
                       aggr1_node1  online     RW        512GB    481.5GB    1%
SVM       flexgroup4TB__0004
                       aggr1_node2  online     RW        512GB    485.5GB    1%
SVM       flexgroup4TB__0005
                       aggr1_node1  online     RW        512GB    485.5GB    1%
SVM       flexgroup4TB__0006
                       aggr1_node2  online     RW        512GB    485.5GB    1%
SVM       flexgroup4TB__0007
                       aggr1_node1  online     RW        512GB    485.5GB    1%
SVM       flexgroup4TB__0008
                       aggr1_node2  online     RW        512GB    485.5GB    1%
SVM       flexgroup4TB__0009
                       aggr1_node1  online     RW        512GB    485.5GB    1%
SVM       flexgroup4TB__0010
                       aggr1_node2  online     RW        512GB    481.2GB    1%
SVM       flexgroup4TB__0011
                       aggr1_node1  online     RW        512GB    481.5GB    1%
SVM       flexgroup4TB__0012
                       aggr1_node2  online     RW        512GB    485.5GB    1%
SVM       flexgroup4TB__00013
                       aggr1_node1  online     RW        512GB    485.5GB    1%
SVM       flexgroup4TB__0014
                       aggr1_node2  online     RW        512GB    485.5GB    1%
SVM       flexgroup4TB__0015
                       aggr1_node1  online     RW        512GB    485.5GB    1%
SVM       flexgroup4TB__0016
                       aggr1_node2  online     RW        512GB    485.5GB    1%
```


Other Command-Line Examples

Creating a FlexGroup Volume by Using `flexgroup deploy`

```
cluster::> flexgroup deploy -size 20PB -space-guarantee volume -vserver SVM -volume flexgroup
```

Using the ONTAP 9.2 `auto-provision-as` Option

```
cluster::> vol create -auto-provision-as flexgroup -vserver SVM -volume flexgroup92
-junction-path /flexgroup92 -size 100t -space-guarantee none -security-style unix
```

Creating a FlexGroup Volume Across Multiple Nodes by Using `volume create`

```
cluster::> volume create -vserver SVM -volume flexgroup -aggr-list aggr1_node1,aggr1_node2 -
policy default -security-style unix -size 20PB -space-guarantee none -junction-path /flexgroup
```

Modifying the FlexGroup Snapshot Policy

```
cluster::> volume modify -vserver SVM -volume flexgroup -snapshot-policy [policyname|none]
```

Applying Storage QoS

```
cluster::> volume modify -vserver DEMO -volume flexgroup -qos-policy-group FlexGroupQoS
```

Applying Volume Autogrow

```
cluster::> volume autosize -vserver DEMO -volume Tech_ONTAP -mode grow -maximum-size 20t -grow-
threshold-percent 80

cluster::> volume autosize -vserver DEMO -volume Tech_ONTAP
Volume autosize is currently ON for volume "DEMO:Tech_ONTAP".
The volume is set to grow to a maximum of 20t when the volume-used space is above 80%.
Volume autosize for volume 'DEMO:Tech_ONTAP' is currently in mode grow.
```

Where to Find Additional Information

To learn more about the information described in this document, refer to the following documents:

Technical Reports

- NetApp Thin Provisioning Deployment and Implementation Guide
  http://www.netapp.com/us/media/tr-3965.pdf
- TR-3982: NetApp Clustered Data ONTAP 8.3.x and 8.2.x
  https://www.netapp.com/us/media/tr-3982.pdf
- TR-4037: Introduction to NetApp Infinite Volume
  https://www.netapp.com/us/media/tr-4037.pdf
- NFS Best Practice and Implementation Guide
  http://www.netapp.com/us/media/tr-4067.pdf
- TR-4379: Name Services Best Practices Guide (before ONTAP 9.3)
  http://www.netapp.com/us/media/tr-4379.pdf
- TR-4668: Name Services Best Practices Guide (ONTAP 9.3 and later)
  http://www.netapp.com/us/media/tr-4668.pdf
- NetApp Data Compression, Deduplication, and Data Compaction
  http://www.netapp.com/us/media/tr-4476.pdf
- NetApp Storage Solutions for Apache Spark
  http://www.netapp.com/us/media/tr-4570.pdf
- NetApp ONTAP FlexGroup Volumes: Top Best Practices
  http://www.netapp.com/us/media/tr-4571-a.pdf

- Electronic Design Automation Best Practices
  http://www.netapp.com/us/media/tr-4617.pdf
- FabricPool Best Practices
  http://www.netapp.com/us/media/tr-4568.pdf
- NetApp FlexCache Volumes in ONTAP 9.7
  https://www.netapp.com/us/media/tr-4743.pdf
- NetApp XCP Best Practices
  https://www.netapp.com/us/media/tr-4808.pdf

Miscellaneous Content

- Tech OnTap Podcast Episode 46: FlexGroups
  https://soundcloud.com/techontap_podcast/episode-46-flexgroups-1
- Tech OnTap Podcast Episode 188: FlexGroup Update
  https://soundcloud.com/techontap_podcast/episode-188-flexgroup-update
- Tech OnTap Podcast Episode 219: FlexVol to FlexGroup Conversion
  https://soundcloud.com/techontap_podcast/episode-219-flexvol-to-flexgroup-conversion
- What's New For FlexGroup Volumes in ONTAP 9.3?
  https://blog.netapp.com/whats-new-for-netapp-flexgroup-volumes-in-ontap-9-3/
- FlexGroup Volumes: An Evolution of NAS
  https://newsroom.netapp.com/blogs/netapp-flexgroup-volumes-an-evolution-of-nas/
- 7 Myths about NetApp ONTAP FlexGroup Volumes
  https://blog.netapp.com/blogs/seven-myths-about-netapp-ontap-flexgroup-volumes/
- Volume Affinities: How ONTAP and CPU Utilization Has Evolved
  https://blog.netapp.com/volume-affinities-how-ontap-and-cpu-utilization-has-evolved/
- FlexGroup lightboard video
  https://www.youtube.com/watch?v=Wp6jEd4VkgI&t=4s

Version History

Version	Date	Document Version History
Version 1.0	January 2017	Initial release
Version 1.0.1	February 2017	Minor revisions
Version 1.1	May 2017	ONTAP 9.2RC1
Version 1.2	December 2017	ONTAP 9.3GA
Version 1.3	May 2018	ONTAP 9.4RC1
Version 1.4	November 2018	ONTAP 9.5
Version 1.5	June 2019	ONTAP 9.6
Version 1.6	January 2020	ONTAP 9.7

Acknowledgments

We would like to thank Chad Morgenstern, Richard Jernigan, Dan Tennant, Ken Cantrell, Sunitha Rao, Shriya Paramkusam, and Mrinal Devadas for their help preparing this report.

Refer to the Interoperability Matrix Tool (IMT) on the NetApp Support site to validate that the exact product and feature versions described in this document are supported for your specific environment. The NetApp IMT defines the product components and versions that can be used to construct configurations that are supported by NetApp. Specific results depend on each customer's installation in accordance with published specifications.

NetApp

www.netapp.com