Abstract

This document provides a brief overview of the new NetApp® ONTAP® FlexGroup feature 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 NetApp FlexGroup volumes, see TR-4557, NetApp FlexGroup: A Technical Overview.
# TABLE OF CONTENTS

1 Data Is Growing ............................................................................................................................................. 8  
1.1 Flexible Volumes: A Tried-and-True Solution .......................................................................................... 8  
1.2 Infinite Volumes: Massive Capacity with Limitations ............................................................................... 8  
1.3 FlexGroup Volumes: An Evolution of NAS .............................................................................................. 9  

2 Terminology ..................................................................................................................................................... 9  
2.1 What Are Large Files? ............................................................................................................................... 10  

3 NetApp ONTAP FlexGroup Advantages ........................................................................................................ 11  

4 Use Cases ........................................................................................................................................................ 12  
4.1 Ideal Use Cases .......................................................................................................................................... 12  
4.2 Nonideal Cases .......................................................................................................................................... 12  

5 Supported Features ....................................................................................................................................... 15  

6 Maximums and Minimums ............................................................................................................................. 18  

7 Deciding Whether NetApp FlexGroup Is the Right Fit ................................................................................ 18  
7.1 Scale-Out Performance ............................................................................................................................... 19  
7.2 Feature Compatibility Limitations ........................................................................................................... 19  
7.3 Simplifying Performance ........................................................................................................................... 20  
7.4 Workloads and Behaviors .......................................................................................................................... 40  

8 Initial FlexGroup Design Considerations ..................................................................................................... 44  
8.1 Cluster Considerations ............................................................................................................................... 44  
8.2 Volume Name Considerations ................................................................................................................... 44  
8.3 Failure Domains ....................................................................................................................................... 44  
8.4 Initial Volume Size Considerations .......................................................................................................... 45  
8.5 Capacity Considerations ............................................................................................................................ 46  
8.6 Elastic Sizing ............................................................................................................................................. 47  
8.7 Aggregate Layout Considerations ............................................................................................................. 49  
8.8 Security and Access Control List Style Considerations ......................................................................... 49  
8.9 Networking Considerations ....................................................................................................................... 52  
8.10 FlexVol Member Volume Layout Considerations .................................................................................. 53  
8.11 General NAS and High File Count Considerations ............................................................................. 63  
8.12 Other Hardware Considerations ............................................................................................................ 85  

9 FlexGroup Sample Designs ........................................................................................................................... 87  

10 FlexGroup Management Best Practices .................................................................................................... 98
10.1 Viewing FlexGroup Volumes ......................................................................................................................... 98
10.2 Monitoring FlexGroup Capacity .................................................................................................................... 103
10.3 Monitoring FlexGroup Performance ............................................................................................................ 124
10.4 Failure Scenarios ........................................................................................................................................ 128
10.5 Nondisruptive Volume Move Considerations .............................................................................................. 129
10.6 Adding Capacity to a FlexGroup Volume .................................................................................................... 130
10.7 Resolving Member Volume Capacity Issues ............................................................................................... 137
10.8 Applying Storage Efficiencies to a FlexGroup Volume ................................................................................ 142
10.9 Applying Storage Efficiencies to Aggregates That Own FlexGroup Volumes ............................................. 143
11 FlexGroup Data Protection Best Practices ...................................................................................................... 144
12 Migrating to NetApp ONTAP FlexGroup ............................................................................................................ 144
  12.1 Migrating from Third-Party Storage to NetApp FlexGroup .......................................................................... 144
  12.2 Migrating from NetApp Data ONTAP Operating in 7-Mode ......................................................................... 144
  12.3 Migrating from FlexVol Volumes, SAN LUNs, or Infinite Volume in ONTAP ............................................... 145
  12.4 XCP Migration Tool ..................................................................................................................................... 145
  12.5 Using XCP to Scan Files Before Migration .................................................................................................. 146
Where to Find Additional Information .................................................................................................. 149
  Technical Reports ............................................................................................................................................... 149
  Miscellaneous Content ....................................................................................................................................... 149
EMS Examples ........................................................................................................................................ 150
  Example of Maxdirsize EMS Message ............................................................................................................... 150
  Examples of Capacity-Related EMS Messages ................................................................................................. 151
Command Examples ........................................................................................................................................ 153
  Other Command-Line Examples ........................................................................................................................ 167
Version History ........................................................................................................................................ 168
Acknowledgments ........................................................................................................................................ 168

LIST OF TABLES
Table 1) General ONTAP feature support ................................................................................................................... 15
Table 2) General NAS protocol version support .......................................................................................................... 16
Table 3) Unsupported SMB 2.x and 3.x features ......................................................................................................... 17
Table 4) How unsupported SMB features behave with NetApp FlexGroup volumes ............................................. 17
Table 5) NetApp FlexGroup maximums ...................................................................................................................... 18
Table 6) NetApp FlexGroup minimums ........................................................................................................................ 18
Table 7) ONTAP volume family comparison................................................................................................................ 20
Table 8) Best practices for aggregate layout with NetApp FlexGroup volumes.................................................................................. 49
Table 9) Inode defaults and maximums according to FlexVol size...................................................................................... 64
Table 10) Inode defaults resulting from FlexGroup member sizes and member volume counts................................. 64
Table 11) High file count/small capacity footprint examples – increasing member volume counts............................ 65
Table 12) Architectural maximum limits for NetApp FlexGroup volumes on FAS9000................................................................. 88
Table 13) Architectural maximum limits for NetApp FlexGroup volumes on AFF A700.......................................................... 89
Table 14) Cluster maximums, various cluster sizes with eight FlexVol members per node........................................ 89
Table 15) flexgroup show output column definitions................................................................................................ 127
Table 16) Storage efficiency guidance for FlexGroup in ONTAP versions............................................................... 142

LIST OF FIGURES
Figure 1) FlexVol design with junctioned architecture for >100TB capacity................................................................. 8
Figure 2) Evolution of NAS file systems in ONTAP......................................................................................................................... 9
Figure 3) What is a large file?............................................................................................................................................... 11
Figure 4) SQL Server Backup environment........................................................................................................................ 13
Figure 5) Throughput and total operations during test runs........................................................................................... 14
Figure 6) CPOC scale-out throughput results...................................................................................................................... 14
Figure 7) NetApp FlexGroup volume.............................................................................................................................. 19
Figure 8) Git benchmark—Linux compile in FlexGroup versus FlexVol................................................................. 21
Figure 9) Git benchmark – gcc compile in FlexGroup versus FlexVol................................................................. 21
Figure 10) NetApp FlexGroup (two-node cluster) versus competitor (14-node cluster): standard NAS workload ................................................................................................................................. 22
Figure 11) Overall response time, SPEC SFS 2014_swbuild submissions................................................................. 23
Figure 12) Throughput, SPEC SFS 2014_swbuild submissions.................................................................................... 23
Figure 13) IOPS, SPEC SFS 2014_swbuild submissions............................................................................................. 24
Figure 14) Latency versus number of builds, SPEC SFS 2014_swbuild submissions......................................................... 24
Figure 15) SPEC SFS 2014_swbuild – concurrent builds.............................................................................................. 25
Figure 16) SPEC SFS 2014_swbuild – latency versus build ops/sec................................................................................... 26
Figure 17) SPEC SFS 2014_swbuild – megabyte per second comparison........................................................................... 26
Figure 18) FlexVol versus FlexGroup—standard NAS benchmark test; NFSv3.......................................................................... 27
Figure 19) Git benchmark: Linux compile completion time in seconds in FlexGroup volume versus FlexVol volume. ........................................................................................................................................... 28
Figure 20) Git benchmark: Linux compile; maximum throughput in FlexGroup volume versus FlexVol volume........ 28
Figure 21) FlexVol volume versus FlexGroup volume: standard NAS benchmark test; SMB 2................................. 29
Figure 22) FlexVol volume versus FlexGroup volume—maximum throughput trends under increasing workload ........................................................................................................................................... 30
Figure 23) FlexVol volume versus FlexGroup volume—maximum throughput trends under increasing workload, detailed........................................................................................................................................... 30
Figure 24) FlexVol volume versus FlexGroup volume—maximum average total IOPS.............................................................. 31
Figure 25) Standard NAS benchmark (EDA) – ONTAP 9.5: one node versus two nodes (operations/sec)...................... 32
Figure 26) Standard NAS benchmark (EDA) – ONTAP 9.4 versus ONTAP 9.5 (operations/sec)......................................... 32
Figure 27) Standard NAS benchmark (EDA) – ONTAP 9.5: one node versus two nodes (MBps)........................................... 33
LIST OF BEST PRACTICES

Best Practice 1: Large File Size Considerations ................................................................. 42
Best Practice 2: Volume autogrow .......................................................................................... 43
Best Practice 3: Member volume size recommendations ...................................................... 46
Best Practice 4: Preferred ONTAP release ........................................................................... 46
Best Practice 5: Aggregate Usage with NetApp FlexGroup ................................................ 49
Best Practice 6: Volume Security Style: Mixed Security Style ........................................... 50
Best Practice 7: Network Design with NetApp FlexGroup ................................................ 52
Best Practice 8: DNS Load Balancing .................................................................................. 53
Best Practice 9: Member Volume Count Recommendations ............................................. 61
Best Practice 10: Deploying FlexGroup Volumes with Existing FlexVol Volumes in Place ................................................................. 63
Best Practice 11: Inode Count in a FlexGroup Volume (prior to ONTAP 9.3) ....................... 63
Best Practice 12: 64-Bit File Identifiers ............................................................................... 67
Best Practice 13: Recommended ONTAP Version for High-File-Count Environments .......... 70
Best Practice 14: Directory Structure Recommendation .................................................... 73
Best Practice 15: Maxdirsize Maximums ............................................................................ 75
Best Practice 16: Avoiding maxdirsize Issues ................................................................... 75
Best Practice 17: Special Character Handling in NetApp FlexGroup Volumes .................... 76
Best Practice 18: UTF-8 or utf8mb4? ............................................................................... 77
Best Practice 19: SMB Change Notification Recommendation ........................................... 77
Best Practice 20: Using Thin Provisioning ......................................................................... 105
Best Practice 21: Volume Space Threshold Recommendations for FlexGroup .................. 107
Best Practice 22: FlexGroup Capacity General Recommendations: Increasing Volume Size .................................................................................... 130
Best Practice 23: FlexGroup Capacity General Recommendations: Adding New Members .................................................................................. 131
Best Practice 24: Actions to Take After Volume Autosize Operations .................................. 135
Best Practice 25: FlexGroup Capacity General Recommendations—Applying Storage Efficiencies ................................................................. 142

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

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 are over. Storage administrators face increasing demands from application owners for large buckets of capacity with enterprise-level performance.

With the advent of big data frameworks such as Hadoop, in which storage needs for a single namespace can extend into the petabyte range (with billions of files), an evolution of NAS file systems is overdue. NetApp ONTAP® FlexGroup is the ideal solution for these architectures.

1.1 Flexible Volumes: A Tried-and-True Solution

The flexible volume (NetApp FlexVol® software) was introduced in NetApp Data ONTAP 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 also 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. 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.

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 SMB 2.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 an innovation in scale-out NAS file systems: NetApp ONTAP FlexGroup volumes. 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.

2 Terminology

Many of the usual NetApp ONTAP terms (such as storage virtual machine, LIF, FlexVol) are covered in TR-3982: NetApp Clustered Data ONTAP 8.3.x and 8.2.x. FlexGroup-specific terminology is covered in the following sections.

FlexGroup Volume

A FlexGroup volume is a single namespace that is made up of multiple constituent member volumes and that is managed and acts like a NetApp FlexVol volume to storage administrators. 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

Automated Incremental Recovery (AIR) is an ONTAP subsystem that repairs FlexGroup inconsistencies dynamically, with no outage or administrator intervention required.

Constituent and Member

Constituent and member are interchangeable terms when referring to FlexGroup. These items are the underlying FlexVol volumes that make up a FlexGroup volume and provide the capacity and performance gains that are seen only with a FlexGroup volume.

Ingest

Ingest is the consumption of data by way of file or folder creations.

Junction Paths

Junction paths are the process that joins multiple FlexVol volumes together to scale out across a cluster, provide multiple volume affinities, and provide capacity beyond the 100TB limit of a FlexVol volume. The use of a junction path in ONTAP is known as “mounting” the volume within the ONTAP namespace.

Large Files

See “What Are Large Files?” in section 2.1.

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 presents 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. As such, 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

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 NetApp 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?

In this document, the term large file is used liberally. Therefore, it’s important to define up front exactly what a “large file” is regarding NetApp FlexGroup volumes.

A FlexGroup volume operates optimally when a workload contains numerous small files or metadata operations, because FlexGroup can maximize the system resources to address those specific workloads. 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 large files or files that grow over time (such as database files).
In a NetApp 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.

Figure 3) What is a large file?

3 NetApp ONTAP FlexGroup Advantages

NetApp ONTAP FlexGroup volumes provide various advantages for different workloads. The advantages include:

Massive Capacity and Predictable Low Latency for High-Metadata Workloads

High-metadata workloads have traditionally presented storage administrators with significant performance and scale problems. Previously, NetApp ONTAP technology did not have a solution for this challenge. Earlier versions were constrained by architectural limitations and the notion of volume affinity: the tendency of ONTAP operations, particularly metadata operations, to operate in a single serial CPU thread.

The NetApp ONTAP FlexGroup feature solves this problem by automatically balancing ingest workloads across multiple constituent NetApp FlexVol members to provide multiple affinities to handle high-metadata workloads.

Efficient Use of All Cluster Hardware

Previously, file systems in ONTAP were tied to a single FlexVol container. Although it was possible to scale volumes across multiple nodes in a cluster, the management overhead was cumbersome, and the process did nothing to increase the total capacity of a single namespace. To achieve this type of scale, volumes could be junctioned to one another. For an example, see Figure 1.

Simple, Easy-to-Manage Architecture and Balancing

With FlexVol volumes, to achieve scale beyond the single node or aggregate that owns the FlexVol volume, several volumes had to be junctioned to one another. This concept required design, architecture, and management overhead that took valuable time away from storage administrators’ day-to-day operations. A FlexGroup volume can provision storage across every node and aggregate in a cluster in less than a minute through the FlexGroup tab in NetApp ONTAP System Manager.
Superior Density for Big Data Workloads

A FlexGroup volume allows you to condense large amounts of data into a smaller data center footprint by way of the superb storage efficiency features of ONTAP, including the following capabilities:

- Thin provisioning
- Data compaction
- Inline data compression
- Inline deduplication (including inline aggregate deduplication in ONTAP 9.2)

In addition, ONTAP supports 30.2TB solid-state drives (SSDs), which can deliver ~724TB of raw capacity in a single 24-drive enclosure. It is possible to get more than a petabyte of raw capacity in just 10U of rack space with NetApp FlexGroup, which cuts costs on cooling, power consumption, and rack rental space.

Superior Resiliency

Another benefit of NetApp FlexGroup is the new recovery infrastructure known as Automated Incremental Recovery. This feature addresses metadata inconsistencies that are found on data access from clients in real time, with little to no disruption of file system access during the recovery. AIR repairs inconsistencies without administrative intervention and without the FlexGroup volume taking itself offline.

4 Use Cases

The NetApp FlexGroup design is most beneficial in specific use cases (electronic design and automation, software development, and so on). Those use cases are listed in Section 4.1 as Ideal Use Cases. In most instances, the use case is only limited to the supported feature set.

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. See the following examples:

- Electronic design automation (EDA)
- Artificial intelligence and machine learning
- Log file repositories
- Software build/test environments (such as Git)
- Seismic, oil, and gas data analysis
- 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 Nonideal Cases

The general guidance for workloads that are currently not recommended to be used on a FlexGroup volume includes:

- Large files; large files that grow over time (such as databases)
- 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

The following sections discuss 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](https://www.youtube.com/watch?v=5t6Q5zY6quo), representatives from NetApp on NetApp describe how FlexGroup volumes are being used to power the NetApp Active IQ infrastructure. The following blog post also covers the solution and gives details and statistics:


**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 using NetApp SnapMirror® for extra data protection.

Each site has a 6-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 ~2.4x the amount of throughput the job needed. At this rate, the backups could complete in approximately 5-6 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.

Data protection

In addition to the performance seen on the FlexGroup 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 6-node cluster was able to push over 8.4GB per second to a FlexGroup volume. In our Customer Proof of Concept 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 2-node AFF A700.

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

Table 1) General ONTAP feature support.

<table>
<thead>
<tr>
<th>Supported Feature</th>
<th>Version of ONTAP First Supported</th>
</tr>
</thead>
<tbody>
<tr>
<td>NetApp Snapshot™ technology</td>
<td>ONTAP 9.0</td>
</tr>
<tr>
<td>NetApp SnapRestore® software (FlexGroup level)</td>
<td>ONTAP 9.0</td>
</tr>
<tr>
<td>Hybrid aggregates</td>
<td>ONTAP 9.0</td>
</tr>
<tr>
<td>Constituent or member volume move</td>
<td>ONTAP 9.0</td>
</tr>
<tr>
<td>Postprocess deduplication</td>
<td>ONTAP 9.0</td>
</tr>
<tr>
<td>NetApp RAID-TEC™ technology</td>
<td>ONTAP 9.0</td>
</tr>
<tr>
<td>Per-aggregate consistency point</td>
<td>ONTAP 9.0</td>
</tr>
<tr>
<td>Sharing NetApp FlexGroup with FlexVol in the same SVM</td>
<td>ONTAP 9.0</td>
</tr>
<tr>
<td>NetApp OnCommand</td>
<td>ONTAP 9.1</td>
</tr>
<tr>
<td>Inline adaptive compression</td>
<td>ONTAP 9.1</td>
</tr>
<tr>
<td>Inline deduplication</td>
<td>ONTAP 9.1</td>
</tr>
<tr>
<td>Inline data compaction</td>
<td>ONTAP 9.1</td>
</tr>
<tr>
<td>Thin provisioning</td>
<td>ONTAP 9.1</td>
</tr>
<tr>
<td>NetApp AFF</td>
<td>ONTAP 9.1</td>
</tr>
<tr>
<td>Quota reporting</td>
<td>ONTAP 9.1</td>
</tr>
<tr>
<td>NetApp SnapMirror® technology</td>
<td>ONTAP 9.1</td>
</tr>
<tr>
<td>User and group quota reporting (no enforcement)</td>
<td>ONTAP 9.1</td>
</tr>
<tr>
<td>Aggregate inline deduplication (cross-volume deduplication)</td>
<td>ONTAP 9.2</td>
</tr>
<tr>
<td>NetApp Volume Encryption (NVE)</td>
<td>ONTAP 9.2</td>
</tr>
<tr>
<td>NetApp SnapVault® technology</td>
<td>ONTAP 9.3</td>
</tr>
<tr>
<td>Qtrees</td>
<td>ONTAP 9.3</td>
</tr>
<tr>
<td>Automated deduplication schedules</td>
<td>ONTAP 9.3</td>
</tr>
<tr>
<td>Version-independent SnapMirror and unified replication</td>
<td>ONTAP 9.3</td>
</tr>
<tr>
<td>Supported Feature</td>
<td>Version of ONTAP First Supported</td>
</tr>
<tr>
<td>----------------------------------------------------------------------------------</td>
<td>-----------------------------------</td>
</tr>
<tr>
<td>Antivirus scanning for SMB</td>
<td>ONTAP 9.3</td>
</tr>
<tr>
<td>Volume autogrow</td>
<td>ONTAP 9.3</td>
</tr>
<tr>
<td>QoS maximums/ceilings</td>
<td>ONTAP 9.3</td>
</tr>
<tr>
<td>FlexGroup expansion without SnapMirror rebaseline</td>
<td>ONTAP 9.3</td>
</tr>
<tr>
<td>Improved ingest heuristics</td>
<td>ONTAP 9.3</td>
</tr>
<tr>
<td>SMB change/notify</td>
<td>ONTAP 9.3</td>
</tr>
<tr>
<td>File audit</td>
<td>ONTAP 9.4</td>
</tr>
<tr>
<td>NetApp FPolicy®</td>
<td>ONTAP 9.4</td>
</tr>
<tr>
<td>Adaptive QoS</td>
<td>ONTAP 9.4</td>
</tr>
<tr>
<td>QoS minimums (AFF only)</td>
<td>ONTAP 9.4</td>
</tr>
<tr>
<td>Relaxed SnapMirror limits</td>
<td>ONTAP 9.4</td>
</tr>
<tr>
<td>SMB 3.x Multichannel</td>
<td>ONTAP 9.4</td>
</tr>
<tr>
<td>FabricPool</td>
<td>ONTAP 9.5</td>
</tr>
<tr>
<td>Quota enforcement</td>
<td>ONTAP 9.5</td>
</tr>
<tr>
<td>Qtree statistics</td>
<td>ONTAP 9.5</td>
</tr>
<tr>
<td>Inherited SMB watches and change notifications</td>
<td>ONTAP 9.5</td>
</tr>
<tr>
<td>SMB copy offload (offloaded data transfer (ODX))</td>
<td>ONTAP 9.5</td>
</tr>
<tr>
<td>Storage-Level Access Guard</td>
<td>ONTAP 9.5</td>
</tr>
<tr>
<td>NetApp FlexCache® (cache only; FlexGroup as origin not supported yet)</td>
<td>ONTAP 9.5</td>
</tr>
<tr>
<td>Volume rename</td>
<td>ONTAP 9.6</td>
</tr>
<tr>
<td>Volume shrink</td>
<td>ONTAP 9.6</td>
</tr>
<tr>
<td>MetroCluster™</td>
<td>ONTAP 9.6</td>
</tr>
<tr>
<td>Elastic sizing</td>
<td>ONTAP 9.6</td>
</tr>
<tr>
<td>Continuously Available shares (SMB)*</td>
<td>ONTAP 9.6</td>
</tr>
<tr>
<td>* SQL Server and Hyper-V workloads only</td>
<td>ONTAP 9.6</td>
</tr>
<tr>
<td>NetApp Aggregate Encryption (NAE)</td>
<td>ONTAP 9.6</td>
</tr>
</tbody>
</table>

Table 2) General NAS protocol version support.

<table>
<thead>
<tr>
<th>Supported NAS Protocol Version</th>
<th>Version of ONTAP First Supported</th>
</tr>
</thead>
<tbody>
<tr>
<td>NFSv3</td>
<td>ONTAP 9.0</td>
</tr>
<tr>
<td>SMB 2.1, SMB 3.x</td>
<td>ONTAP 9.1 RC2</td>
</tr>
</tbody>
</table>
Table 3) Unsupported SMB 2.x and 3.x features.

<table>
<thead>
<tr>
<th>Unsupported SMB 2.x Features</th>
<th>Unsupported SMB 3.x Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>• SMB Remote Volume Shadow Copy Service (VSS)</td>
<td>• SMB transparent failover</td>
</tr>
<tr>
<td></td>
<td>• SMB scale-out</td>
</tr>
<tr>
<td></td>
<td>• Volume Shadow Copy Service (VSS) for SMB file shares</td>
</tr>
<tr>
<td></td>
<td>• SMB directory leasing</td>
</tr>
<tr>
<td></td>
<td>• SMB direct or remote direct memory access (RDMA)</td>
</tr>
<tr>
<td>Note: SMB 3.0 encryption is supported with FlexGroup volumes.</td>
<td></td>
</tr>
</tbody>
</table>

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 NetApp FlexGroup volumes.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Behavior with NetApp FlexGroup volumes</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMBv1.0</td>
<td>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.</td>
</tr>
<tr>
<td>Change notification/SMB watches (before ONTAP 9.3)</td>
<td>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.</td>
</tr>
<tr>
<td>Offloaded Data Transfer (ODX)</td>
<td>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, offloaded data transfer (ODX) works as expected. For more information about ODX, see this TechNet article.</td>
</tr>
<tr>
<td>Remote Volume Shadow Copy Service (VSS)</td>
<td>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 NetApp FlexGroup volumes. For more information about Remote VSS, see this TechNet article.</td>
</tr>
<tr>
<td>Continuously available shares</td>
<td>Continuously available shares are not allowed on NetApp FlexGroup volumes prior to ONTAP 9.6. Attempting to set the share property fails. For more information about continuously available shares, see this TechNet article.</td>
</tr>
</tbody>
</table>
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 whether the maximum is hard-coded/enforced or a recommended/tested value.

Table 5) NetApp FlexGroup maximums.

<table>
<thead>
<tr>
<th>Value</th>
<th>Value</th>
<th>Value Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>FlexGroup volume size</td>
<td>20PB</td>
<td>Tested/recommended</td>
</tr>
<tr>
<td>File count</td>
<td>400 billion</td>
<td>Tested/recommended</td>
</tr>
<tr>
<td>Cluster node count</td>
<td>24 (12 HA pairs)</td>
<td>Hard-coded/enforced</td>
</tr>
<tr>
<td>NetApp FlexVol member volume size</td>
<td>100TB</td>
<td>Hard-coded/enforced</td>
</tr>
<tr>
<td>FlexVol member volume file count</td>
<td>2 billion</td>
<td>Hard-coded/enforced</td>
</tr>
<tr>
<td>NetApp SnapMirror volume count</td>
<td>32 (ONTAP 9.4 and earlier)</td>
<td>Hard-coded/enforced</td>
</tr>
<tr>
<td>SnapMirror volume count (Total)</td>
<td>100 (ONTAP 9.4 and earlier)</td>
<td>Hard-coded/enforced</td>
</tr>
<tr>
<td></td>
<td>6,000 (ONTAP 9.5 and later)</td>
<td>Hard-coded/enforced</td>
</tr>
<tr>
<td>File size</td>
<td>16TB</td>
<td>Hard-coded/enforced</td>
</tr>
<tr>
<td>FlexVol member constituent count</td>
<td>200</td>
<td>Tested/recommended</td>
</tr>
<tr>
<td>Aggregate size/count</td>
<td>Same as NetApp ONTAP limits</td>
<td>Hard-coded/enforced</td>
</tr>
</tbody>
</table>

**Note:** 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.

Table 6) NetApp FlexGroup minimums.

<table>
<thead>
<tr>
<th>Value</th>
<th>Value</th>
<th>Value Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>FlexVol member size</td>
<td>100GB</td>
<td>Tested/recommended</td>
</tr>
<tr>
<td>Data aggregate count</td>
<td>1</td>
<td>Hard-coded/enforced</td>
</tr>
<tr>
<td>SnapMirror schedule</td>
<td>30 minutes</td>
<td>Tested/recommended</td>
</tr>
<tr>
<td>NetApp Snapshot schedule</td>
<td>30 minutes</td>
<td>Tested/recommended</td>
</tr>
</tbody>
</table>

7 Deciding Whether NetApp FlexGroup Is the Right Fit

NetApp 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 NetApp FlexGroup volumes. This section provides information to help decide whether NetApp 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 NetApp 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 FlexVol volumes, which in turn can be stored on any aggregates and can span multiple nodes in your cluster.

Figure 7) NetApp 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 NFS v3, SMB 2.x, and SMB 3. For details about the support of those protocols, see Table 1 through Table 3 in section 5, “Supported Features.”

Other NAS protocols, such as SMBv1 NFS v4.x (including pNFS), are currently not supported. NetApp 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 NetApp FlexGroup volumes are the right fit for an environment by comparing the currently available container types in ONTAP.
### Table 7) ONTAP volume family comparison.

<table>
<thead>
<tr>
<th></th>
<th>FlexVol Volumes</th>
<th>Infinite Volume</th>
<th>FlexGroup Volumes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Client access protocols (current support)</td>
<td>SAN, NAS—all versions</td>
<td>NAS only:</td>
<td>NAS only:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• SMB 1.0</td>
<td>• SMB 2.1/3.x</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• NFSv3</td>
<td>• NFSv3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• NFSv4.0, NFSv4.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• pNFS</td>
<td></td>
</tr>
<tr>
<td>Capacity scaling</td>
<td>• Single FlexVol volume</td>
<td>• 20PB</td>
<td>• 20PB*</td>
</tr>
<tr>
<td></td>
<td>• Can be mounted to other FlexVol volumes</td>
<td>100TB, 2 billion file limit</td>
<td>• 400 billion files*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>*Current tested limits on 10-node cluster; can extend beyond these values</td>
</tr>
<tr>
<td>Metadata scaling</td>
<td>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.</td>
<td>NetApp FlexGroup volumes can use multiple nodes (and their resources) and multiple aggregates. In addition, NetApp FlexGroup can use multiple volume affinities to maximize CPU thread utilization potential.</td>
<td></td>
</tr>
<tr>
<td>ONTAP feature compatibility</td>
<td>Compatible with all ONTAP features</td>
<td>Supports many ONTAP features. For NetApp FlexGroup information, see <a href="#">TR-4557: NetApp FlexGroup—A Technical Overview</a> and the “Supported Features” section of this document. For details about Infinite Volume, see <a href="#">TR-4037: Introduction to NetApp Infinite Volume</a>.</td>
<td></td>
</tr>
<tr>
<td>Throughput scaling</td>
<td>Limited to:</td>
<td>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.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• One node (set of CPU, RAM, network ports, connection limits, and so on.)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• One aggregate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data permanence during ONTAP upgrades and reverts</td>
<td>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., NetApp FlexGroup volumes cannot be retained.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GUI compatibility</td>
<td>All volume families have some level of GUI support in NetApp OnCommand products, including:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• ONTAP System Manager</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• OnCommand Performance Manager</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Active IQ Unified Manager</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• NetApp OnCommand Insight</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### 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 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.

![Git Benchmark](image)

Figure 9) Git benchmark – gcc compile in FlexGroup versus FlexVol.

![Performance Comparison](image)
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 solid-state drives (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) NetApp 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 allow 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, NetApp FlexGroup volumes achieved the lowest ORT ever recorded for a storage system.
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 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.

![Git Operation Completion Times](image)

Figure 20 shows the same workloads but measures the overall throughput in gigabits 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.

![FlexVol Versus FlexGroup Throughput (GBps)](image)

Metadata-intensive workloads, such as those found in common high-performance computing 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 SMB 2.x: ONTAP 9.3

Figure 21 shows the performance of a standard NAS workload benchmark that used SMB 2 on a NetApp FlexGroup volume in a two-node cluster versus a single FlexVol volume.

Figure 21) FlexVol volume versus FlexGroup volume: standard NAS benchmark test; SMB 2.

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, 8 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.

![Total Maximum Throughput, MBps](image)

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.

![Throughput MB/s](image)
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 new 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 fairly substantial performance improvements in ONTAP. This 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) 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).

**Standard NAS Benchmark (EDA)**

ONTAP 9.5: One node vs. Two nodes

![Graph showing latency vs. throughput for ONTAP 9.5 with one node and two nodes.]

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

**Standard NAS Benchmark (Software Build)**

ONTAP 9.4 vs. ONTAP 9.5

![Graph showing latency vs. throughput for ONTAP 9.4 and ONTAP 9.5 with one node and two nodes.]

© 2019 NetApp, Inc. All Rights Reserved.
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 NetApp FlexGroup volume with big data workloads allows all available hardware to be used and providing a way to nondisruptively scale the capacity and performance by adding nodes for the workload as needed.

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

Node 1

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 storage virtual machine (SVM).

There are some workloads for which a FlexGroup volume cannot maintain an ideal load and usage balance. The following subsections cover the workloads that do and do not maintain ideal load and usage balances.

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.

As such, it is highly recommended to 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 a FlexGroup volume 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 Volumes

With a FlexGroup volume, 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 spanning multiple nodes in a cluster.

Figure 34) Storage QoS on NetApp FlexGroup volumes – single-node connection.
Storage QoS Considerations with FlexGroup Volumes

Currently, storage QoS is applied only at the FlexGroup volume level and supports only QoS maximums. QoS minimums, adaptive QoS, file-level QoS, and nested policies are currently not supported with NetApp FlexGroup volumes. Policies are currently applied at the command line only. GUI support for FlexGroup volume QoS will be included in future ONTAP releases.

Quality of Service (QoS) Minimums

ONTAP 9.4 adds 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 Quality of Service (QoS)

ONTAP 9.4 also introduces 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 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. It 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 in this call (and not a FlexVol) are one or a combination of the following values:

- **Aggregates**: If you specify more than one, then the REST API creates a FlexGroup. 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 the “aggregates.name” or “aggregates.uuid” when creating or expanding a FlexGroup. 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 and it creates a FlexGroup if it is specified. If a
volume is being created on multiple aggregates, the system will always create a FlexGroup. 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 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 that has members that are too small) The job string can be accessed through the browser using: `http://[your_cluster_IP_or_name]/api/cluster/jobs/job_uuid`.

This is what a failure message might look like:

```json
{
    "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 below is what a “success” job would look like:

```json
{
    "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 introduces 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](https://www.netapp.com/).
NetApp Cloud Volumes ONTAP

ONTAP 9.6 introduced official support for Cloud Volumes ONTAP (CVO) – NetApp’s ONTAP solution running in the cloud. You can now deploy a FlexGroup volume using Cloud Manager.

FlexGroups running in CVO can use the same feature sets available in the ONTAP version deployed to the CVO instance. Some common use cases seen for CVO and FlexGroups include:

- Data lake for analytics
- EDA repositories for use with Amazon Elastic Compute Cloud (Amazon EC2) compute 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, CVO instances max out at 368TB per instance and FlexGroup volumes cannot span more than one instance. Also, creating a FlexGroup currently requires use of System Manager or CLI. There is no way to create a FlexGroup in Cloud Manager. 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.

The workloads that are driven against a FlexGroup volume determine the degree to which the volume behaves optimally. There are some workloads that a FlexGroup volume responds well to, and others, it might struggle to accommodate. 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.
- **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 processes 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.

**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 FlexGroup 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).

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:** Free space can also be monitored through GUI utilities such as Active IQ Unified Manager, and by generating 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.

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 by an order of magnitude), 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.
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 substantial; it simply represents suboptimal behavior.

**Best Practice 1: Large File Size Considerations**

Before sizing a FlexGroup volume, perform an analysis to determine the largest possible file system 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 new elastic sizing functionality.
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, resulting in Volume Full errors (ENOSPC) even if other constituents 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.

### 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 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 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.
- Enable quota enforcement to limit the capacity in qtrees or by user. (ONTAP 9.5 and later)
- Use the NetApp XCP Migration Tool before deploying 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 so that the largest file size does not exceed 1-5% 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 10-50GB, 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 extreme performance gains over FlexVol volumes that the ingest-heavy workloads see. Sometimes (especially 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, there are benefits to using FlexGroup volumes that are not apparent with 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 the reader has experience with and knowledge about managing ONTAP through the CLI and the GUI and that the reader has admin-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 high-availability pairs). Each HA pair is a homogenous system (that is, two 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 NetApp AFF nodes and six hybrid nodes for storage tiering functionality.

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

- **NetApp 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 NetApp FlexGroup volume.

- **NetApp 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.

- **NetApp FlexGroup volumes can span portions of a cluster.** A NetApp FlexGroup volume can be configured to span any node 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 but 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 introduces support for volume rename. 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. Junction paths and SMB shares can be changed at any time in a FlexGroup volume but doing so causes a disruption for clients while the volume is remounted and clients reconnect.

Volume renames in ONTAP 9.6 are performed identically to how FlexVol volumes are renamed.

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. The HA pair is a failure domain in the cluster.

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, write access is disabled until all those members are repaired. 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 weigh those considerations against the capacity required and performance needed.
8.4 Initial Volume Size Considerations

One common deployment issue that customers run into is under-sizing their FlexGroup volume’s member volume capacity. FlexGroup volumes can be created at almost any capacity, but it’s important to remember that underneath the large container provided by the FlexGroup volumes are several FlexVol member volumes that make up the total size of the FlexGroup. Generally, each node has eight-member volumes by default (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 with eight-member volumes, each member volume is 10TB in size.

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

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
9 entries were displayed.
```

Aggregate Free Space Considerations

When creating a FlexGroup volume, it’s ideal for the aggregate (or aggregates) that the FlexGroup is deployed on 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. (ONTAP 9.5 and earlier)
ONTAP 9.5 and earlier prevents a FlexGroup from filling an aggregate past 97% when using thick provisioning/space guarantees. Attempts to do so 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 no longer checks for deduplication metadata.

Why Is This Important?

Available member volume size in a FlexGroup affects how often files are ingested locally or remotely in a FlexGroup, which in turn can affect performance and capacity distribution in the FlexGroup. File sizes are also important to consider when 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, 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 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 member volume sizes that have no more than 1% to 5% effect from creating a single file. 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 ~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 prior to 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 NetApp WAFL® reserve, and storage efficiencies are factored in. For more information, see the storage limits on
8.6 Elastic Sizing

Files written to a FlexGroup 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, that write might fail due to lack of space in a member volume.

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

- If you 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, if your volume is composed of 4K files but then you zip some up and create a giant single file).

One solution is to grow volumes 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 can be hundreds of terabytes in size, but the underlying member volumes and their free capacities are what determines the space available for individual files. If a 200TB FlexGroup has 20TB remaining (10% of the volume), the amount of space available for a single file to write is not 20TB – instead, it is 20TB/[number of member volumes in a FlexGroup].

In a 2-node cluster, a FlexGroup that spans both nodes is likely to have 16 member volumes. That means if 20TB is available in a FlexGroup, then the member volumes have 1.25TB available. Any single file that exceeds 1.25TB in size could not write to a FlexGroup volume without volume autogrow enabled prior to ONTAP 9.6.

Starting in ONTAP 9.6, a new feature known as “elastic sizing” was introduced to avoid “out of space” errors in the scenario above. 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 volume developers refers to elastic sizing as an “airbag” in that 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 from 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 will get placed in the best available member based on various factors – such as free capacity, inode count, time since last file creation, member volume performance (new in ONTAP 9.6), and so on.

3. When a file is placed, since ONTAP doesn’t know how large a file will get, 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 error to the client that we’ve run out of space, it will query the other member volumes in the FlexGroup to see if there’s any available space to borrow. If there is, ONTAP will add 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 will continue.

5. During the time ONTAP is looking for space to borrow, that file write is paused – this will appear 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 will be 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, as the system will need to query to borrow space more often.

6. The capacity borrowing will maintain the overall size of the FlexGroup – for example, if your FlexGroup is 40TB in size, it will remain 40TB.

Figure 39) File write behavior before elastic sizing.

![Diagram of Elastic Sizing](image1)

Figure 40) File write behavior after elastic sizing.

![Diagram of Elastic Sizing](image2)

After files are deleted/volumes are grown and space is available in that member volume again, ONTAP will re-adjust the member volumes back to their original sizes to maintain an evenness in space.
Ultimately, elastic sizing helps remove the admin overhead of managing space, and worrying so much about the initial sizing/deployment of a FlexGroup. 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 with features like autogrow/shrink, then ONTAP can manage your capacity usually and help avoid emergency space issues.

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.

<table>
<thead>
<tr>
<th>Best Practice 5: Aggregate Usage with NetApp FlexGroup</th>
</tr>
</thead>
<tbody>
<tr>
<td>For consistent performance across the entire FlexGroup volume when using NetApp FlexGroup, make sure that the FlexGroup volume spans only aggregates with the same disk type and RAID group configurations.</td>
</tr>
</tbody>
</table>

Table 8 shows the best practices that NetApp recommends for aggregate layout when you use FlexGroup volumes. Keep in mind that these best practices are not hard requirements. The one-aggregate-per-node recommendation for AFF systems originates from disk cost with RAID-TEC™ and no Advanced Disk Partitioning (ADP), because you don’t want to use up expensive SSD just for parity. However, with Advanced Disk Partitioning (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>
<thead>
<tr>
<th>Table 8) Best practices for aggregate layout with NetApp FlexGroup volumes.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spinning Disk or Hybrid Aggregates</td>
</tr>
<tr>
<td>Two aggregates per node</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

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 and Failing Drives

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 NetApp 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 the volume is created 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:

- UNIX security style needs Windows users to map to valid UNIX users.
- NTFS security style needs Windows users to map to a valid UNIX user and needs UNIX users to map to valid Windows users to authenticate. Authorizations (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 access control lists (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 fpolicy, antivirus, native file auditing and quota enforcement, then use the latest patched release of ONTAP 9.5 or later.

**Best Practice 6: Volume Security Style: Mixed Security Style**

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 NetApp FlexGroup Volume**

NetApp FlexGroup volumes are intended to be managed similarly to FlexVol volumes. Changing the security style of volumes is included in that methodology.

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 OnCommand System Manager.

Workarounds for Lack of NFSv4.x ACL Support

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. However, FlexGroup volumes do not currently support NFSv4.x.

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. 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 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.

## 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 a storage virtual machine (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, Peripheral Component Interconnect (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.

NetApp FlexGroup volumes are not immune to this line of thinking. Although FlexVol volumes can be distributed across multiple nodes in a cluster just like a FlexGroup volume, the network connection can still be a bottleneck.

### Best Practice 7: Network Design with NetApp FlexGroup

When you design a FlexGroup solution, consider the following networking best practices:

- Create at least one data LIF per node per SVM to confirm a path to each node.
- 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.
- 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 a high-availability (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, due to the design of FlexGroup volumes, remote cluster traffic is a certainty when provisioning a FlexGroup volume across nodes, so network connection/data locality considerations are thrown out the window 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 decide 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.

**Network Connection Concurrency**

In addition to the preceding considerations, it’s worth noting that ONTAP has a limit of 128 concurrent operations per TCP connection for NAS operations. This limit means that for every IP address, the system can handle only up to 128 concurrent operations. Therefore, it’s possible that a client would not be able to push the storage system hard enough to reach the full potential of the FlexGroup technology.

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.

**Border Gateway Protocol (BGP): ONTAP 9.5**

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 is striped 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 NetApp FlexGroup volume. Each FlexGroup 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` (new in ONTAP 9.2), `flexgroup deploy`, or OnCommand System Manager, the default number of member FlexVol volumes in a FlexGroup volume depends on the platform and ONTAP release. It can range from eight per node (four per aggregate) or 16 per node (eight per aggregate).

The following 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: The member volume count recommendation per node is “let ONTAP decide” in nearly all use cases. 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 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” in section 8.8.

• When growing or shrinking (shrink starting in ONTAP 9.6) a FlexGroup, 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?

Usually, letting ONTAP choose the member volumes by using the automated commands or GUI operations is a good option when creating a FlexGroup volume. Generally speaking, these scenarios are only for capacity or system limit considerations, and rarely for performance considerations. In other words, you shouldn’t worry about CPU count/volume affinity best practices for FlexGroup creation – let ONTAP do that for you.

However, there might be some use cases where manual creation might be needed.

Concern Regarding Overprovisioning Volume Counts

In ONTAP, FlexVol volumes are currently limited to a maximum of 1000 per node. Because a FlexGroup 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 limits and potentially need to manually create the FlexGroup volumes to modify the default volume counts or aggregate placement.
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.

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, as it is code-limited to best practices of 8 per node. In the following example, my 2-node cluster can only create a FlexGroup with a maximum of 1.56PB of capacity (8 members per node; 16 members total; 100TB per member volume).

Figure 42) Error when creating a FlexGroup beyond the allowed maximum in System Manager.

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

Similar considerations should be made if the file count needs to exceed the maximum files allowed. In our 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 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 across.
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 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 affects the total number of available member volumes. FlexVol member volumes also contend for the FlexGroup volumes participating in SnapMirror relationships. See TR-4678 for the latest details on those limitations.

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 OnCommand System Manager or the `flexgroup deploy/-auto-provision-as` options in the CLI. For more information, see the previous section, “FlexVol Member Volume Layout Considerations.”
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 ~500GB in size.

Figure 44) How capacity is divided among member volumes in a FlexGroup.

In many workloads, the distribution shown in Figure 44 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, then the FlexGroup could report as out of space.

For example, if some files in a workload are 250GB, then each time a file is written to the FlexGroup, 50% of the total capacity of a member volume is filled.
If a second 250GB file attempts to write to a member volume that is not able to honor the write, then 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 help alleviate this scenario with the inclusion of **elastic sizing**.

Features like volume autogrow (available starting in ONTAP 9.3) and **elastic sizing** in ONTAP 9.6 can help offset the impact of this type of workload, but eventually, when physical space is exhausted, more storage must be added. Fortunately, **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 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 FlexGroup 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.**
  Size the total FlexGroup to a value large enough to accommodate member volume sizes that can handle the workload. In our example, the FlexGroup 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, the CLI can be used to create a FlexGroup that is identical in total capacity but contains fewer member volumes that are larger in size. 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 available performance of the FlexGroup volume), but it would allow larger files to be placed.

    Figure 47) Fewer, larger member volumes.

Volume autogrow should be used whenever possible to avoid out-of-space issues, and ideally, ONTAP 9.6 or later should be used to apply the benefits of elastic sizing.

In most cases, letting ONTAP create the FlexGroup volume/set the member volume count via the automated methods (GUI/auto-provision-as via CLI) is the preferred way to create a FlexGroup volume. For guidance on what situations would 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 on 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).

```bash
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 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 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:

```bash
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 (new in ONTAP 9.2), the flexgroup deploy command, or the OnCommand System Manager GUI handle this setup for the storage administrator.

Considerations When Deleting FlexGroup Volumes

In ONTAP 8.3, a new 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 in 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 as well, so space is not freed up until the recovery queue expires or is manually purged.

Deleted volumes can be seen from the command line by specifying -type DEL in Diag Privilege. (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 in Diag Privilege:

```
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 in Advanced Privilege. ONTAP System Manager does not support forced deletions.
Qtrees

ONTAP 9.3 introduces 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 and provides flexibility in data management by enabling unique export policies, unique security styles, and granular statistics. Individual qtrees currently cannot be replicated via 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 some more information regarding qtrees in FlexGroup volumes.

- Qtrees are distributed across a FlexGroup volume the same way as 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.
- ONTAP 9.5 added quota enforcement support. For more information about tree quotas, 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 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 existing 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. 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, 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 in ONTAP versions earlier than 9.6. For more information, see “Aggregate Layout Considerations” earlier in this document.

Figure 48) How FlexVol capacity can affect FlexGroup load distribution.
Note: This is only an issue only if the FlexGroup volume is thin provisioned. Space-guaranteed FlexGroup volumes would not have other volumes eating into the space footprint.

Consider the Performance Impact of the Existing FlexVol Volumes

In addition to the capacity considerations of existing FlexVol volumes, it’s also important to consider the amount of work the existing FlexVol volumes are doing when deploying a FlexGroup volume. 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.

Consider the Volume Count Limits

ONTAP places a volume count of 1000 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 impact 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 ActiveIQ 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 a balanced 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 move to relocate volumes to 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. 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 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 is:

\[ 200 \times 21,251,126 = 4,250,225,200 \]

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 (prior to ONTAP 9.3)

Prior to ONTAP 9.3, the ingest calculations for data that is written into a FlexGroup volume did not consider inode counts when deciding where to place files. 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
Best Practice 11: Inode Count in a FlexGroup Volume (prior to ONTAP 9.3)

Volumes running out of inodes prematurely. If 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 varies depending on workload, but you should not set the value to the maximum at the start. Setting the maxfiles to the largest value gives no room to increase later without having to add member volumes. If possible, upgrade to ONTAP 9.3 to take advantage of the new ingest calculations for high file-count environments.

**Note:** Inodes and maxfiles are interchangeable terms here.

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

<table>
<thead>
<tr>
<th>FlexVol Size</th>
<th>Default Inode Count</th>
<th>Maximum Inode Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>20MB</td>
<td>566</td>
<td>4,855</td>
</tr>
<tr>
<td>1GB</td>
<td>31,122</td>
<td>249,030</td>
</tr>
<tr>
<td>100GB</td>
<td>3,112,959</td>
<td>24,903,679</td>
</tr>
<tr>
<td>1TB</td>
<td>21,251,126</td>
<td>255,013,682</td>
</tr>
<tr>
<td>10TB</td>
<td>21,251,126</td>
<td>2,040,109,451</td>
</tr>
<tr>
<td>100TB</td>
<td>21,251,126</td>
<td>2,040,109,451</td>
</tr>
</tbody>
</table>

**Note:** FlexGroup members 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>
<thead>
<tr>
<th>Member Volume Size</th>
<th>Member Volume Count</th>
<th>Default Inode Count (FlexGroup)</th>
</tr>
</thead>
<tbody>
<tr>
<td>100GB</td>
<td>8</td>
<td>24,903,672</td>
</tr>
<tr>
<td>100GB</td>
<td>16</td>
<td>49,807,344</td>
</tr>
<tr>
<td>1TB</td>
<td>8</td>
<td>170,009,008</td>
</tr>
<tr>
<td>1TB</td>
<td>16</td>
<td>340,018,016</td>
</tr>
<tr>
<td>100TB</td>
<td>8</td>
<td>170,009,008</td>
</tr>
<tr>
<td>100TB</td>
<td>16</td>
<td>340,018,016</td>
</tr>
</tbody>
</table>

**High File Counts, Low Capacity Needs**

ONTAP allocates a default inode and maximum inode count based on volume capacity. In the chart above, member volumes smaller than 10TB will not be able to achieve the 2 billion inodes available to a FlexVol volume. To get 2 billion inodes per member volume, the member volume capacity would need to be 10TB or greater. In a FlexGroup volume with 8 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, as file sizes can be small and won’t require that much capacity. For example, if all files in a workload are 288 bytes in size, then 16 billion files would only use up 4.6TB, which is around 6% of the total capacity needed to contain 16 billion files in ONTAP.

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 impact the free capacity, so it’s important to monitor available aggregate space when using thin provisioning.

**Manually create the FlexGroup with more member volumes than the default.**

If you want to keep space guarantees for volumes, then another option for high file count/small capacity environments is to create more member volumes in a FlexGroup volume.

Since inode counts are limited per FlexVol member volume based on 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.

<table>
<thead>
<tr>
<th>Total FlexGroup Size</th>
<th>Member Volume Count (Size)</th>
<th>Maximum Inode Count (entire FlexGroup)</th>
</tr>
</thead>
<tbody>
<tr>
<td>80TB (no space guarantee)</td>
<td>8 (10TB)</td>
<td>16,320,875,608</td>
</tr>
<tr>
<td>64TB (space guarantee enabled)</td>
<td>32 (2TB)</td>
<td>16,320,875,608</td>
</tr>
<tr>
<td>64TB (space guarantee enabled)</td>
<td>64 (1TB)</td>
<td>16,320,875,608</td>
</tr>
</tbody>
</table>

**Planning for Inode Counts in ONTAP**

With tools like XCP (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 in Advanced Privilege:

```
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:

```
cluster:/> df -i /vol/flexgroup/
Filesystem         iused   ifree  %iused Mounted on Vserver
/vol/flexgroup/    823 170008185   0% /flexgroup SVM
```

**Note:** Inode counts for FlexGroup volumes are available only at the FlexGroup level.
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 NetApp 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 more 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

Today, NFSv3 in ONTAP uses 32-bit file IDs by default. 32-bit file IDs are limited to 2,147,483,647 maximum unsigned integers. With the two billion inode limit in FlexVol, this value fits nicely into the architecture.

However, because NetApp 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 highly recommends enabling the NFS server option `-v3-64bit-identifiers` at the Advanced Privilege level before creating a FlexGroup volume, especially if your file system exceeds or might exceed the two billion inode threshold.

The 64-bit file identifier option is off 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 OSes can handle 64-bit file IDs without issue.

This option can currently be enabled only from Advanced Privilege 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.

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 impact SMB operations and is unnecessary with volumes only using SMB.

Using Quota Enforcement to Limit File Count

Now that ONTAP 9.5 supports quota enforcement, it’s possible to set up a quota policy that prevents a FlexGroup volume from exceeding 2 billion files if 32-bit file handles are still being used.

Since 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

<table>
<thead>
<tr>
<th>Volume</th>
<th>Tree</th>
<th>Type</th>
<th>ID</th>
<th>Used</th>
<th>Limit</th>
<th>Used</th>
<th>Limit</th>
<th>Specifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>FG4</td>
<td>twobillionfiles</td>
<td>tree</td>
<td>1</td>
<td>0B</td>
<td>-</td>
<td>1</td>
<td>2000000000</td>
<td>twobillionfiles</td>
</tr>
<tr>
<td></td>
<td></td>
<td>tree</td>
<td>*</td>
<td>0B</td>
<td>-</td>
<td>0</td>
<td>2000000000</td>
<td>*</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Volume</th>
<th>Tree</th>
<th>Type</th>
<th>ID</th>
<th>Used</th>
<th>Limit</th>
<th>Used</th>
<th>Limit</th>
<th>Specifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>FG4</td>
<td>twobillionfiles</td>
<td>tree</td>
<td>1</td>
<td>0B</td>
<td>-</td>
<td>4</td>
<td>2000000000</td>
<td>twobillionfiles</td>
</tr>
<tr>
<td></td>
<td></td>
<td>tree</td>
<td>*</td>
<td>0B</td>
<td>-</td>
<td>0</td>
<td>2000000000</td>
<td>*</td>
</tr>
</tbody>
</table>

System Manager Support for the 64-bit File ID Option

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

Figure 50) Enable or disable 64-bit file identifiers in System Manager.

![Enable or disable 64-bit file identifiers in System Manager.]

After enabling or disabling this option, you must remount all clients, because the file system IDs change, and the clients might receive stale file handle messages when attempting NFS operations.

**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 using NetApp 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
Change: 2017-01-06 15:23:58.527329000 -0600
```

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 impact SMB operations and is unnecessary with volumes only using 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, chmod, and so on.

An example of this issue can be found in bug 671319. If disabling the FSID change with NFSv3, be sure to enable the new -v3-64bit-identifiers option in ONTAP 9. However, keep in mind that this option could affect older legacy applications that require 32-bit file IDs. NetApp recommends leaving the FSID change option enabled with NetApp FlexGroup volumes to help prevent file ID collisions.

How FSIDs Operate with Snapshot Copies

When a Snapshot copy of a volume is taken, 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 spec does not require that FSIDs for a file to be identical across file versions.

Note: The -v4-fsid-change option does not apply to NetApp FlexGroup volumes, as NFSv4 is currently unsupported with FlexGroup volumes.

Directory Size Considerations

In ONTAP, there are limitations to the maximum directory size on disk. This limit is known as maxdirsize. The maxdirsize 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?

Maxdirsize can be a concern when 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 Impact FlexGroup Volumes**

Flat directory structures (many files in a single/few directories) have 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, GETATTRs, READDIR, and so on)
- CPU utilization

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

However, in flat directories, a higher ratio of remote hardlinks to local files will be seen. These remote hardlinks 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 hardlinks for the file system, you can expect the 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 ONTAP-specific commands to do this. 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
360573347 328976 drwxr-xr-x 2 root root 335544320 May 29 21:23 /flexgroup/manyfiles/folder3/topdir_0/subdir_0
347117639 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
239949155 328976 drwxr-xr-x 2 root root 335544320 May 29 14:15 /flexgroup/manyfiles/folder3/topdir_1/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
1325527652 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

XCP is mostly considered as 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 examples above can be run even faster on the storage system.

The following XCP command example allows us to run “find” only on directories with more than 2000 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 us 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 will scan 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
31.8M 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
```

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.

How Many Files Can Fit into a Single Directory with the Default maxdirsize?

This KB article (internal and partner only) explains how to determine whether a directory’s index can cause maxdirsize issues. If you don’t have access to this article, the following example shows how many files can fit into a single directory in a larger memory system by default.

CIFS/SMB files use more space per file than NFS due to the namelength size and Unicode encoding. Therefore, volumes that are accessed with SMB (either dedicated SMB or multiprotocol NAS) can hold fewer files per folder than NFS-only volumes.

Use the following formulas to determine the number of files:
- Memory in KB * 53 * 25% for SMB/multiprotocol
- Memory in KB * 128 * 25% for NFS

The maxdirsize is set to 320MB by default on larger systems, so the maximum number of files is as follows:

- 4,341,760 for SMB/multiprotocol
- 10,485,760 for NFS

**What EMS Messages Are Sent When maxdirsize Is Exceeded?**

The following 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 (example below). ActiveIQ Unified Manager can be used to monitor maxdirsize, trigger alarms, and send a notification before the 90% threshold. These EMS messages also support SNMP traps.

<table>
<thead>
<tr>
<th>Message Name</th>
<th>Severity</th>
<th>Corrective Action</th>
<th>Description</th>
<th>Supports SNMP trap</th>
</tr>
</thead>
<tbody>
<tr>
<td>wafl.dir.size.max</td>
<td>ERROR</td>
<td>Use the &quot;volume file show-inode&quot; 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 &quot;volume modify -volume vol_name -maxdir-size new_value&quot; 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.</td>
<td>This message occurs after a directory has reached its maximum directory size (maxdirsize) limit.</td>
<td>true</td>
</tr>
<tr>
<td>wafl.dir.size.max.warning</td>
<td>ERROR</td>
<td>Use the &quot;volume file show-inode&quot; 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 &quot;volume modify -volume vol_name -maxdir-size new_value&quot; 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.</td>
<td>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.</td>
<td>true</td>
</tr>
<tr>
<td>wafl.dir.size.warning</td>
<td>ERROR</td>
<td>Use the &quot;volume file show-inode&quot; 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 &quot;volume modify -volume vol_name -maxdir-size new_value&quot; 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.</td>
<td>This message occurs when a directory surpasses 90% of its current maximum directory size (maxdirsize) limit.</td>
<td>true</td>
</tr>
</tbody>
</table>
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. In ONTAP 9.2, the introduction of 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, in the case of wildcard searches and readdir operations, indexing is not of much use.

Best Practice 15: Maxdirsize Maximums

Maxdirsize values 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 NetApp 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, NetApp 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 EMS 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 EMS events, see section 13.3, "EMS 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 3 bytes. This common text includes all modern written languages, such as Chinese, Japanese, 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 3 bytes in NFS, as indicated in bug 229629. To handle character sizes that exceed 3 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 NetApp FlexGroup volumes added support for bagofbits in ONTAP 9.2.

Best Practice 17: Special Character Handling in NetApp FlexGroup Volumes

For special character handling with NetApp FlexGroup volumes, use ONTAP 9.2.

Also, ONTAP has an EMS 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 NetApp 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 3 bytes that UTF-8 encoding. ONTAP will then use 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 4 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 some legacy situations, there are NFS hosts that are configured to use non-UTF-8 encodings, and for those cases you will want to use the corresponding volume language. Use of UTF-8 has become almost universal these days, so the volume language should almost universally 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 names 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 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 using a tool like XCP.

Best Practice 18: UTF-8 or utf8mb4?

If running ONTAP 9.5 or later, it is best to use the utf8mb4 volume language to help prevent issues with file name 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 due to 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 we are able to use more hardware and WAFL affinities to spread out the delete load more efficiently. However, performance of file deletions might be slower due to remote access across the FlexGroup 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 52) Capacity imbalance after deletion of larger files.](image)

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 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 does a good job of balancing the ingest load usually, 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 XCP 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 increase the size of the existing FlexGroup member volumes to provide some relief.

**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 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 prior to 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 hardlink. Subsequent renames would create more remote hardlinks 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, attempts to resolve that many symlinks might have a negative effect on performance. The negative effect is because you are creating remote hardlinks artificially in addition to the remote hardlinks ONTAP creates.

Try to keep the number of symlinks down 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. However, because FlexGroup volumes do not currently support NFSv4.x, the mount request fails. 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
```

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 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 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, then attempts to access a FlexGroup volume from the NFSv4.x mount fail. This includes `ls -la` operations, because those require the ability to do NFSv4.x GETATTR operations.
Note in the following example that the information for the flexgroup volumes is incorrect:

```
# 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
d?????????? ?? ?? ?   ? flexgroup_local
```

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 ..
```

As a result, be sure to avoid using NFSv4.x in any path where a FlexGroup volume resides.

**NFS Mount Considerations: READDIRPLUS (READDRIP+)**

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 do not support VAAI. If `-vstorage` is enabled in the NFS options, virtualization operations such as VM migrations fail, because they attempt to use an unsupported feature. vStorage...
is disabled on NFS servers by default. If using FlexGroup volumes with virtualization workloads, make sure vstorage is disabled.

- Offload operations are not supported, so features such as rapid cloning or instant cloning do not work with datastores deployed on a FlexGroup volume.
- FlexGroup volumes do not support NFSv4.x, nor do they support continuously available shares.
- 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, except that NVE cannot be enabled on existing FlexGroup volumes. Currently, only new FlexGroup volumes can use NVE. To encrypt existing FlexGroup volumes, you must create a volume with encryption enabled and then copy the data to the volume at the file level, such as with XCP.

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)
- FAS or AFF 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 introduced NetApp Aggregate Encryption (NAE) which provides the same encryption key across multiple volumes that reside in the same aggregate. If you are using NAE on aggregates that contain FlexGroup member volumes, confirm that all aggregates spanning the FlexGroup are NAE aggregates. A mix of NAE and non-NAE in the same FlexGroup is not allowed.

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:

```
Volume Encryption Details

  Encrypted     Yes

  Encryption Status     No operation in progress
```
Rekeying a FlexGroup Volume or Encrypting Existing FlexGroup Volumes

ONTAP 9.5 has 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 Considerations

ONTAP 9.5 also includes NetApp FlexCache. This feature provides a sparse volume that can accelerate performance for NAS workloads and prevent volume hotspots 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). RAL is also what makes a FlexGroup volume a FlexGroup volume. However, source or origin volumes for FlexCache currently are only FlexVol volumes. There is no support for FlexGroup volumes as an origin for FlexCache.

ONTAP 9.6 increased the limit of maximum cache volumes per node to 100, so FlexCache has more scalability in current releases.

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 you must have the FlexGroup 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 spanning new and old aggregates.
When adding new nodes to a cluster, the same steps for adding aggregates to a cluster should be followed. Use volume moves and volume expand commands to adjust the member volumes.

Removing nodes

Removing nodes from a cluster can also be done by using nondisruptive volume moves of member volumes. For example, if you want to remove two nodes from an 8-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 6 nodes, then 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, then place six member volumes on two nodes (12) and five member volumes on four nodes (20).
NetApp FlexGroup offers benefits that Infinite Volume does not offer, in that a FlexGroup volume can be managed like a normal FlexVol volume. The following design variations are allowed with a NetApp FlexGroup volume.

NetApp FlexGroup volumes can:

- Share storage virtual machine (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)

NetApp FlexGroup volumes should not:
• Be configured to use mixed or hybrid disk or aggregate types on the same FlexGroup volume
• 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 is 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 NetApp 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 (FAS9000). The example in Table 13 is a 24-node 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 at a minimum one vsroot volume (one volume per SVM).
• The hardware has a maximum capacity (400TB aggregates prior to 9.2, 800TB per aggregate [AFF only] in 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 NetApp FlexGroup volumes on FAS9000.

<table>
<thead>
<tr>
<th>Cluster Size</th>
<th>Architectural Maximum Member Volumes per FlexGroup Volume</th>
<th>Architectural Maximum Capacity per FlexGroup Volume</th>
<th>Theoretical Maximum Inodes per FlexGroup Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>24 nodes</td>
<td>23,975 (1,000 volumes per node * 24 nodes – 25 reserved volumes)</td>
<td>176PB (based on architectural hardware limits of the FAS9000)</td>
<td>~47 trillion inodes (based on 2 billion inodes * 23,975 FlexGroup members)</td>
</tr>
</tbody>
</table>
Table 13) Architectural maximum limits for NetApp FlexGroup volumes on AFF A700.

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

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.

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

* 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 56) 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” in section 10.4.

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
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. In these cases, a small penalty is incurred (~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 NetApp 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 indirect traffic over the cluster interconnects.

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

Figure 59Figure 63 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.

Figure 59 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 59) Git clone completion times comparison.

<table>
<thead>
<tr>
<th>GIT Clone Completion Times</th>
</tr>
</thead>
<tbody>
<tr>
<td>Local FlexVol Volume Versus Local FlexGroup versus Clusterwide FlexGroup Volume</td>
</tr>
</tbody>
</table>

![Graph showing completion times comparison](image)

Figure 60 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.
In Figure 61 and Figure 62, 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 clusterwide FlexGroup volume shows the opposite trend.
Figure 62) Maximum write throughput comparison.

Figure 63 displays the total average IOPS for a local FlexVol volume versus the local and clusterwide FlexGroup configurations. The FlexGroup configurations produce twice the IOPS that the FlexVol volume does, with the local FlexGroup volume outperforming the clusterwide FlexGroup volume until the 64-thread tipping point.

Figure 63) Total average IOPS comparison.
In this test, 64 worker threads appear to be a sweet spot. Therefore, 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.

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
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 NetApp 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 67) 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 SnapMirror volumes at a more succinct level
- The ability to create a multitenant data organization without needing to create SVMs

With NetApp 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 68) 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 NetApp 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

NetApp 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.
ActiveIQ Performance Manager

ActiveIQ 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 ActiveIQ Performance Manager. Currently, only the command line and the System Manager GUI can carry out these tasks.

Figure 70 shows several FlexVol members and their corresponding performance. Each line represents a FlexVol member.

Figure 70) ActiveIQ Performance Manager FlexGroup volume view.

In Figure 71, 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 72, we can see the IOPS and MBps graphs.
ActiveIQ Unified Manager

With ActiveIQ Unified Manager, storage administrators can use a single dashboard to review the health of a NetApp ONTAP cluster. It can also integrate with ActiveIQ Performance Manager to give a single management view.
With ActiveIQ Unified Manager, you can review FlexGroup volume capacity, configurations, and storage efficiencies in a graphical format. For member volume information about performance, use ActiveIQ Performance Manager. For capacity information about member volumes, use the command line.

Figure 73) ActiveIQ Unified Manager FlexGroup capacity 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
- 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
- Maximum FlexGroup member size

Diag Privilege Level

- Detailed member volume information (capacity, used, and so on)
- FlexGroup ingests statistics

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 new NetApp FPolicy™ support available 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 ActiveIQ Unified Manager, or through the CLI in 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 due to used space eating into other volume allocations. Figure 75 illustrates how other volumes can eat into volumes with thin provisioning enabled.
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 ActiveIQ Unified Manager to monitor and alert for capacity.

Figure 76 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.

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

In ActiveIQ 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 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::> vol 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 because ONTAP is calculating against the aggregate’s available space. The FlexVol member volumes show equivalent available values depending on the aggregates where they are located.

```
cluster::> vol 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
```

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

**Best Practice 20: Using Thin Provisioning**

If you use thin provisioning with ONTAP, it is important to use tools such as ActiveIQ 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.
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 section 13.4, "Command Examples."

EMS Messages

Specific EMS messages are available so that storage administrators can alert on 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 section 13.3, "EMS Examples."

Unmodifiable values:

- Severity level
- Corrective actions
- Description
- SNMP support

Modifiable values:

- Destinations
- Allowed drops or intervals between transmissions

When an EMS 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 EMS 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%
```

EMS 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

These values can be used when testing EMS messages. When you look for which SVM is affected by the errors, use the UUID string in Advanced Privilege:

```
cluster::*> vserver show -uuid 05e7ab78-2d84-11e6-a796-00a098696ec7
Vserver Type Subtype Admin State Operational Root
---------- ------- -------- ------- ---------- ----------
SVM       data    default  running  running  vsroot aggr1_node1
```
Testing EMS Messages

To test an EMS message, use the `event generate` command (available at the Diag Privilege level). Each EMS 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 EMS message:

```plaintext
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
```

<table>
<thead>
<tr>
<th>Time</th>
<th>Node</th>
<th>Severity</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>11/28/2016 18:36:35</td>
<td>cluster-01</td>
<td>ALERT</td>
<td>monitor.volume.nearlyFull: Volume flexgroup@vserver:05e7ab78-2d84-11e6-a796-00a098696ec7 is nearly full (using or reserving 95% of space and 0% of inodes).</td>
</tr>
</tbody>
</table>

Modifying the Volume Full and Nearly Full Thresholds

With a FlexVol volume, the default values for `full` and `nearly full` 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 prior to ONTAP 9.3 because the volume autogrow functionality is 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 ActiveIQ 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, OnCommand displays active warnings and errors about capacity.
Figure 77) ActiveIQ Unified Manager capacity risks.

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

Figure 78) Capacity-related events—detailed view.

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

Figure 79) 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 80** 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 81** Volume capacity detail.
Click the Actions button to create alerts that are specific to the volume.

**Figure 82) Adding an alert.**

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

**Figure 83) Including volumes for alerts.**

Events are organized by severity and include 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 84) Volume capacity events.**
When an event is triggered, the alert mechanism in ActiveIQ Unified Manager can:

- 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).

**Editing Volume Thresholds in ActiveIQ Unified Manager**

Thresholds for Volume Nearly Full and Volume Full control when an EMS event is triggered by the cluster. This control helps storage administrators stay on top of the volume capacities to help prevent volumes from running out of space. In FlexGroup, this approach takes on an extra aspect regarding 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 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 is an Edit Thresholds option. Select that option and a window opens 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.
When you initially check the box under Capacity: Global Level, the defaults are as shown in Figure 87. These defaults are unaffiliated with the ONTAP EMS volume thresholds. Rather, they are specific to ActiveIQ Unified Manager.

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

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

You can use ActiveIQ Unified Manager alerting independently of the cluster’s EMS alerting and event destination logic, or you can use it along with them.
Inode Monitoring

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

Figure 89) Inode alerts.

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

Figure 90) 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 over provisioning 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 ~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 ~4TB available.

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

```
classroom::> 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 over provisioned 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
```

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.

```
# 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
```

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
```

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

<table>
<thead>
<tr>
<th>Filesystem</th>
<th>Size</th>
<th>Used</th>
<th>Avail</th>
<th>Use%</th>
<th>Mounted on</th>
</tr>
</thead>
<tbody>
<tr>
<td>10.193.67.220:/flexgroup</td>
<td>80T</td>
<td>70T</td>
<td>11T</td>
<td>88%</td>
<td>/flexgroup</td>
</tr>
</tbody>
</table>

## 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 91 shows just the overprovisioned FlexGroup volume to confirm the behavior.

![Figure 91) Windows capacity view.](image)

### Viewing FlexVol Member Capacity from the ONTAP Command Line

When FlexGroup volumes are created, each member is evenly divided based on 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 from the **Diag Privilege**; use `volume show -is-constituent true` from the **Admin Privilege**; or use the `volume show-space` command from the **Admin Privilege**. You can find examples in previous sections of this document, such as the section “Capacity Monitoring and Alerting with the Command Line.”

Viewing FlexVol member capacity is useful when trying to find out 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: FlexGroup Technical Overview.

### 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 over allocating 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 92) How logical space accounting works.

ONTAP 9.5 enhances this feature and adds 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 introduces support for FabricPool for FlexGroup volumes. There are no special considerations to make for FlexGroup volumes; the same FlexVol considerations apply.

Managing Quotas with NetApp FlexGroup

NetApp 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 now 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 in Admin Privilege:

```
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 in Admin Privilege. 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

Volume   Tree    Type    ID         Used  Limit    Used  Limit   Specifier
-------  --------  ------  -------  -----  -----  ------  -------  --------
flexgroup_local qtree    tree    1         1.01GB 1GB  10      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 in Admin Privilege:

```
cluster::> quota policy rule create -vserver DEMO -policy-name tree -volume flexgroup_local -type tree -target qtree
```

To enable, 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

<table>
<thead>
<tr>
<th>Volume</th>
<th>Tree</th>
<th>Type</th>
<th>ID</th>
<th>Used</th>
<th>Limit</th>
<th>Used</th>
<th>Limit</th>
<th>Specifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>flexgroup_local</td>
<td>qtree</td>
<td>tree</td>
<td>1</td>
<td>0B</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>qtree</td>
</tr>
</tbody>
</table>
```

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

<table>
<thead>
<tr>
<th>Volume</th>
<th>Tree</th>
<th>Type</th>
<th>ID</th>
<th>Used</th>
<th>Limit</th>
<th>Used</th>
<th>Limit</th>
<th>Specifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>flexgroup_local</td>
<td>qtree</td>
<td>tree</td>
<td>1</td>
<td>13.77MB</td>
<td>-</td>
<td>4</td>
<td>-</td>
<td>qtree</td>
</tr>
</tbody>
</table>
```

**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 EMS message is logged at the DEBUG severity level to notify storage administrators of the quota violation. You can configure these EMS 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

<table>
<thead>
<tr>
<th>Type</th>
<th>Target</th>
<th>Qtree</th>
<th>Masking</th>
<th>Soft User Limit</th>
<th>Soft Disk Limit</th>
<th>Soft File Limit</th>
<th>Soft Files Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>qtree</td>
<td>0</td>
<td></td>
<td></td>
<td>1GB</td>
<td>-</td>
<td>10</td>
<td>-</td>
</tr>
</tbody>
</table>
```

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
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 five files. An extra five 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
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.

```

```
------ Disk ----- Files ----- Quota
--------- ------- ------- ------ ------

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.
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 EMS 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**

Starting in ONTAP 9.4, you can also 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:

![Quota Report UI](image)

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:

![Edit Limits UI](image)

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.
The System Manager UI displays the following quota report:

<table>
<thead>
<tr>
<th>Volume</th>
<th>OwnerType</th>
<th>Users</th>
<th>Type</th>
<th>Group</th>
<th>Quota Policy</th>
<th>% Space Used</th>
<th>% File Used</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Home</td>
<td>All Users</td>
<td>User</td>
<td>User</td>
<td>student1</td>
<td>default</td>
<td>80</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>tree</td>
<td>All Users</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>tree</td>
<td>All Users</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>tree</td>
<td>All Users</td>
<td></td>
<td>15%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

An event message is also generated. The Quota Exceeded error is a DEBUG-level event, so filter by DEBUG in System Manager events:

Tree Quota Considerations

Storage virtual machines (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
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:

```bash
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.

```bash
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, quota on works, but the SVM was does not use the new tree policy.

```bash
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.

```bash
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 have to 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 clusterwide 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 section 13.4, “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 stats 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

<table>
<thead>
<tr>
<th>Counter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>cat1_tld_remote</td>
<td>2</td>
</tr>
<tr>
<td>cat2_hld_local</td>
<td>180</td>
</tr>
<tr>
<td>cat2_hld_remote</td>
<td>1292</td>
</tr>
<tr>
<td>cat3_dir_local</td>
<td>146804</td>
</tr>
<tr>
<td>cat3_dir_remote</td>
<td>283</td>
</tr>
<tr>
<td>cat4_fil_local</td>
<td>734252</td>
</tr>
<tr>
<td>cat4_fil_remote</td>
<td>1124</td>
</tr>
<tr>
<td>groupstate_analyze</td>
<td>12232</td>
</tr>
<tr>
<td>groupstate_update</td>
<td>86242</td>
</tr>
<tr>
<td>instance_name</td>
<td>0</td>
</tr>
<tr>
<td>node_name</td>
<td>cluster-01</td>
</tr>
</tbody>
</table>
```
process_name
refreshclient_create 5241
refreshclient_delete 5241
refreshserver_create 5244
refreshserver_delete 5244

The statistics capture gives a nice round up 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.

```bash
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

flexgroup show

During FlexGroup I/O, you can also view the member constituent usage and balance through the nodeshell command `flexgroup show`.

```bash
cluster::>* node run * flexgroup show flexgroup4TB_thin

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 14 describes what those values are and how to interpret them. [TR-4557: NetApp FlexGroup Technical Overview](#) covers some of these terms and concepts in more explicit detail.
Table 15) flexgroup show output column definitions.

<table>
<thead>
<tr>
<th>Column</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Idx</td>
<td>Index number of the member volume.</td>
</tr>
<tr>
<td>Member</td>
<td>Dataset ID (DSID) of the FlexGroup member.</td>
</tr>
<tr>
<td>L</td>
<td>Local or remote to the node.</td>
</tr>
<tr>
<td>Used</td>
<td>Number and overall percentage of 4K blocks used.</td>
</tr>
<tr>
<td>Urgc</td>
<td>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 based on how close to 100% used a volume’s capacity is.</td>
</tr>
<tr>
<td>Targ</td>
<td>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%.</td>
</tr>
<tr>
<td>Probabilities</td>
<td>The likelihood that a member volume is avoided for use. This number increases based on how full a member volume becomes in relation to other member volumes (tolerance).</td>
</tr>
<tr>
<td>D-Ingest and D-Alloc</td>
<td>Directory ingest and directory allocation, respectively; how many directories have been allocated to a local member volume.</td>
</tr>
<tr>
<td>F-Ingest and F-Alloc</td>
<td>File ingest and file allocation, respectively; how many files have been allocated to a local member volume.</td>
</tr>
</tbody>
</table>

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 Perstat, under `wafl_susp -w`.

**Monitoring Performance by Using Perstat**

Perfstat is a tool in the [NetApp Support site ToolChest](https://www.netapp.com) 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 a 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](https://www.netapp.com).

When a 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](https://www.latexproject.org) to view these files.

For more information, see the [KB article on how to use Perfstat](https://www.netapp.com).
Monitoring Performance (ActiveIQ Performance Manager or Unified Manager)

A more palatable and widely available tool for monitoring the performance of a FlexGroup volume is NetApp ActiveIQ 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.

ActiveIQ Performance Manager offers both real-time and historical performance information for a cluster and can be integrated with ActiveIQ Unified Manager to provide a single management point. ActiveIQ Performance Manager can give granular performance views for the entire FlexGroup volume or for individual member constituent FlexVol volumes.

Figure 95 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 ActiveIQ Performance Manager can deliver.

![Figure 95) ActiveIQ Performance Manager graphs.](image)

10.4 Failure Scenarios

The following 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 a FlexVol. Takeovers have no noticeable disruption. Nondisruptive upgrades, head swaps, rolling upgrades, and so on all perform as normal. Givebacks of stateful protocols, such as SMB, are slightly disruptive, due to 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 mirrors that of a FlexVol. 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, then ONTAP considers that Snapshot copy to be “partial” and invalidates it for 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, then the aggregate is considered offline and the member volumes that live on the offline aggregate are inaccessible. In a FlexGroup 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 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 a cluster. This feature provides flexibility when dealing with maintenance windows or when attempting to balance performance or capacity allocation in a cluster.

NetApp FlexGroup volumes also support this feature, but with even more granularity; each member volume in a FlexGroup volume can be moved 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 ActiveIQ 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 Move
Nondisruptive volume moves can come in handy in the following scenarios for NetApp 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 Move
Nondisruptive volume move for a NetApp FlexGroup member volume is available at the Admin Privilege level of the command line. There is no support currently for ONTAP System Manager to perform volume moves for FlexGroup member volumes from the GUI. ONTAP System Manager does support moving FlexVol volumes from the GUI, however.

To move a FlexGroup member volume, complete the following steps:
7. Identify the volume that needs to be moved. Use ActiveIQ Performance Manager to determine this information.

8. 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 ONTAP 8.3, a new feature, Auto Balance Aggregate, 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 NetApp FlexGroup volumes.

### 10.6 Adding Capacity to a FlexGroup Volume

A NetApp 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, NetApp 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 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.

#### 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.
**Best Practice 22: FlexGroup Capacity General Recommendations: Increasing Volume Size**

- 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 to avoid headaches later as the FlexGroup volume grows closer to full and remediation steps are needed.

**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 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.”

**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 take these individual increases into consideration 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 section 13.4, “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, left click the volume to select it. Then click Actions > Resize.

Volume Autosize (Autogrow/Autoshrink)

In ONTAP 9.3, support for volume autogrow was added for NetApp 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 introduces elastic sizing, which operates independently of volume autogrow.

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: In ONTAP 9.6, this error is mitigated usually 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, sometimes, other member volumes might have enough free space, but another member volume might 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
```
In those scenarios, a storage administrator previously would have to manually intervene and grow the member volumes or delete data to restore proper functionality. In ONTAP 9.3, volume autosize can be used to grow the member volumes automatically when the writes being attempted 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 "DEM0:fgautogrow".
The volume is set to grow to a maximum of 100g when the volume-used space is above 80%.
Volume autosize for volume 'DEM0:fgautogrow' is currently in mode grow.
```

The member volumes are all 1GB in size*, 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   available used
------- ---------------- ------ --------- -------
DEMO    fgautogrow__0001 1GB    915.6MB   57.23MB
DEMO    fgautogrow__0002 1GB    915.6MB   57.23MB
DEMO    fgautogrow__0003 1GB    915.6MB   57.23MB
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
```

*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 you land grows to the appropriate size to honor the write. In this case, member volume `fgautogrow__0003` is where the file is written.

```
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  57.26MB
DEMO    fgautogrow__0003 1.60GB  498.7MB  1.03GB
```

When this happens, an event is triggered in EMS and can be seen with `event log show`.

```
INFORMATIONAL waf1.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, send alerts through event destinations, or with ActiveIQ Unified Manager.

```
cluster::> event route show -message-name waf1.vol.autoSize.done -instance
```

<table>
<thead>
<tr>
<th>Message Name: wafl.vol.autoSize.done</th>
</tr>
</thead>
<tbody>
<tr>
<td>Severity: INFORMATIONAL</td>
</tr>
<tr>
<td>Corrective Action: (NONE)</td>
</tr>
<tr>
<td>Description: This message occurs on successful autosize of volume.</td>
</tr>
<tr>
<td>Supports SNMP trap: true</td>
</tr>
<tr>
<td>Destinations: -</td>
</tr>
<tr>
<td>Number of Drops Between Transmissions: 0</td>
</tr>
<tr>
<td>Dropping Interval (Seconds) Between Transmissions: 0</td>
</tr>
</tbody>
</table>

After a member volume has been grown through autogrow, there is an imbalance of member volume available size/allocation.

Figure 97) 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 EMS 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 and later introduces elastic sizing. Elastic sizing helps mitigate ENOSPC errors but should not be viewed as a substitution 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:

- The existing member volumes have reached their maximum capacities; or
- The physical limits of the node capacities have been reached; or
• New nodes are added to the cluster

As mentioned earlier, it is preferable to grow the member volumes and to apply volume move usually. 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 ActiveIQ Unified Manager to Automate Volume Increases

ActiveIQ Unified Manager alerting allows you to run a script when an alert is triggered. In the case of 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 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, Windows PowerShell, and .bat files.

Figure 98) Managing scripts in ActiveIQ Unified Manager.

When the script is added to ActiveIQ, modify the alert to kick off the script when it is triggered. You can manage alerts through the Administration menu that is shown in Figure 99.

Figure 99) 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 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.

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 FlexGroups.

If rebalance is necessary, consider using the XCP Migration Tool to speed up the file copy process.

10.7 Resolving Member Volume Capacity Issues

If a NetApp 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 NetApp recommends in this report 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 the best method of preventing member volumes from running out of space.

If a client receives an ENOSPC (out of space) error, you should take the following steps.

Verify Which Member Volume Is Running Out of Space

To verify which member volume is running out of space, run volume show in Diag Privilege or volume show-space in Admin Privilege 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__0009 10TB
9 entries were displayed.
```

```c
# dd if=/dev/zero of=/uneven/1TB.file count=262144000 bs=4096
# ls -lah
```

```
total 1004G
```
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.

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 favor the most empty FlexVol member so that performance and capacity could potentially suffer.

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%.

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 NetApp 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         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 offers 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.
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 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 using a combination of ActiveIQ Performance Manager and AutoSupport. The initial error seen was the following:

```plaintext
Volume flexgroup__0006@vserver:18c1fb66-57c9-1e7-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 track down 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:

```plaintext
/vol/flexgroup__0002/ 2932657360 2547943464  384713896      87%  /vol/flexgroup__0002/
/vol/flexgroup__0002/.snapshot  154350384  239276116          0     155%
/vol/flexgroup__0004/ 2932657360 2528758844  403898516      86%  /vol/flexgroup__0004/
/vol/flexgroup__0004/.snapshot  154350384  363345924          0     235%
/vol/flexgroup__0006/ 2932657360 2509467192  423190168      86%  /vol/flexgroup__0006/
/vol/flexgroup__0006/.snapshot  154350384  194732648          0     126%
/vol/flexgroup__0008/ 2932657360 2511307952  421349408      86%  /vol/flexgroup__0008/
/vol/flexgroup__0008/.snapshot  154350384  267119112          0     173%
```

The `df` section then showed this on Wednesday:

```plaintext
/vol/flexgroup__0002/ 2932657360 2345427276  587230084      80%  /vol/flexgroup__0002/
/vol/flexgroup__0002/.snapshot  154350384  216960512          0     18%
/vol/flexgroup__0004/ 2932657360 2223625248  709032112      76%  /vol/flexgroup__0004/
/vol/flexgroup__0004/.snapshot  154350384  122452102          0     0%
/vol/flexgroup__0006/ 2932657360 2199918944  617844496      76%  /vol/flexgroup__0006/
/vol/flexgroup__0006/.snapshot  154350384  66640154283744        0%  0%
/vol/flexgroup__0008/ 2932657360 2239204232  693453128      76%  /vol/flexgroup__0008/
/vol/flexgroup__0008/.snapshot  154350384  22684864131665520        15%  0%
```

When the member volumes were graphed out, the discrepancy stood out decisively.
When we look at `flexgroup show` from AutoSupport, we can see that 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.

Because we knew that the issue was likely not 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 can we narrow the time period in when the issue occurred? The answer is simple: ActiveIQ Performance Manager. ActiveIQ Performance Manager collects data over time and saves it so that we can see what workload patterns happen and at what times they occur. From ActiveIQ 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 ActiveIQ Performance Manager.

Only one member volume is writing anything here
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 a nearly 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 particular 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

NetApp 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, applying the policies is done at the FlexGroup level for simpler management of storage efficiencies. Table 16 provides guidance on which ONTAP versions support FlexGroup-level management of these features and on which ONTAP versions require more granular management of the efficiencies per member volume.

<table>
<thead>
<tr>
<th>Best Practice 25: FlexGroup Capacity General Recommendations—Applying Storage Efficiencies</th>
</tr>
</thead>
<tbody>
<tr>
<td>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.</td>
</tr>
<tr>
<td>To view storage efficiency settings, run the following commands in Diag Privilege:</td>
</tr>
<tr>
<td>volume efficiency show -volume flexgroup* -fields data-compaction,compression,inline-compression,inline-dedupe</td>
</tr>
<tr>
<td>volume show -volume flexgroup* -fields is-sis-state-enabled</td>
</tr>
</tbody>
</table>

For more information about these features, see TR-4557: FlexGroup Technical Overview and TR-4476: NetApp Data Compression, Deduplication, and Compaction.

Table 16) Storage efficiency guidance for FlexGroup in ONTAP versions.

<table>
<thead>
<tr>
<th>Feature</th>
<th>9.1RC1</th>
<th>9.1RC2</th>
<th>9.1GA and later</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thin provisioning</td>
<td>FlexGroup level</td>
<td>FlexGroup level</td>
<td>FlexGroup level</td>
</tr>
<tr>
<td>Inline deduplication</td>
<td>FlexVol member</td>
<td>FlexGroup level</td>
<td>FlexGroup level</td>
</tr>
<tr>
<td>Postprocess deduplication</td>
<td>FlexVol member</td>
<td>FlexGroup level</td>
<td>FlexGroup level</td>
</tr>
<tr>
<td>Inline data compaction</td>
<td>FlexVol member</td>
<td>FlexGroup level</td>
<td>FlexGroup level</td>
</tr>
<tr>
<td>Inline data compression</td>
<td>FlexVol member</td>
<td>FlexGroup level</td>
<td>FlexGroup level</td>
</tr>
<tr>
<td>Postprocess data compression</td>
<td>FlexVol member</td>
<td>FlexGroup level</td>
<td>FlexGroup level</td>
</tr>
<tr>
<td>Cross-volume deduplication</td>
<td>N/A</td>
<td>N/A</td>
<td>FlexGroup level</td>
</tr>
<tr>
<td>(9.2 and later)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
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 AFF or hybrid aggregate configurations.
- Cross-volume deduplication will see storage efficiencies on a per-aggregate basis.

### Applying Storage Efficiencies per FlexGroup Member Volume (ONTAP 9.1RC1)

If a FlexGroup volume does not currently have support to enable storage efficiencies at the FlexGroup level (applicable only to ONTAP 9.1RC1), use the following command to enable it on every FlexGroup member.

```bash
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.

```bash
cluster::*> volume efficiency modify -vserver SVM -volume flexgroup4* -compression true -inline-compression true -inline-dedupe true
```

8 entries were modified.

```bash
cluster::*> volume efficiency show -vserver SVM -volume flexgroup4* -fields data-compaction,compression,inline-compression,inline-dedupe
```

<table>
<thead>
<tr>
<th>Vserver</th>
<th>Volume</th>
<th>Compression</th>
<th>Inline-Compression</th>
<th>Inline-Dedupe</th>
<th>Data-Compaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM</td>
<td>flexgroup4TB__0001</td>
<td>true</td>
<td>true</td>
<td>true</td>
<td>true</td>
</tr>
<tr>
<td>SVM</td>
<td>flexgroup4TB__0002</td>
<td>true</td>
<td>true</td>
<td>true</td>
<td>true</td>
</tr>
<tr>
<td>SVM</td>
<td>flexgroup4TB__0003</td>
<td>true</td>
<td>true</td>
<td>true</td>
<td>true</td>
</tr>
<tr>
<td>SVM</td>
<td>flexgroup4TB__0004</td>
<td>true</td>
<td>true</td>
<td>true</td>
<td>true</td>
</tr>
<tr>
<td>SVM</td>
<td>flexgroup4TB__0005</td>
<td>true</td>
<td>true</td>
<td>true</td>
<td>true</td>
</tr>
<tr>
<td>SVM</td>
<td>flexgroup4TB__0006</td>
<td>true</td>
<td>true</td>
<td>true</td>
<td>true</td>
</tr>
<tr>
<td>SVM</td>
<td>flexgroup4TB__0007</td>
<td>true</td>
<td>true</td>
<td>true</td>
<td>true</td>
</tr>
<tr>
<td>SVM</td>
<td>flexgroup4TB__0008</td>
<td>true</td>
<td>true</td>
<td>true</td>
<td>true</td>
</tr>
</tbody>
</table>
```

### 10.9 Applying Storage Efficiencies to Aggregates That Own FlexGroup Volumes

ONTAP 9.2 introduced a new feature called Aggregate Inline Deduplication. This new 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 only available for AFF platforms 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. Prior to 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](https://www.netapp.com/support/trial/tr-4476-netapp-data-compression-deduplication-and-data-compaction.pdf).
11 FlexGroup Data Protection Best Practices

For the FlexGroup data protection best practices, see TR-4678: Data Protection and Backup – 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 is not considering 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.

But how do you get there?

Data migrations can take three general forms when dealing with NetApp FlexGroup:

- Migrating from non NetApp (third-party) storage to NetApp FlexGroup
- Migrating from NetApp Data ONTAP operating in 7-Mode to NetApp FlexGroup
- Migrating from FlexVol volumes, SAN LUNs, or Infinite Volume in ONTAP to NetApp FlexGroup

Data migrations to FlexGroup volumes are the best way to migrate. FlexGroup volume migrations currently cannot be performed with the following methods:

- NDMP
- FlexVol to FlexGroup volume move
- FlexVol or Infinite Volume in-place conversion
- 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 Migrating from Third-Party Storage to NetApp 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.

NFSv4 data, especially data with NFSv4 ACLs, should use a tool that has ACL preservation and NFSv4 support.

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.

12.2 Migrating from NetApp Data ONTAP Operating in 7-Mode

Migrate data from NetApp Data ONTAP operating in 7-Mode to NetApp FlexGroup in one of two ways:
• Full migration of 7-Mode systems to clustered 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 NetApp FlexGroup.

Currently, there is no migration path directly from FlexVol to NetApp FlexGroup that does not involve copy-based migrations.

12.3 Migrating from FlexVol Volumes, SAN LUNs, or Infinite Volume in ONTAP

When migrating from existing clustered 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 NetApp FlexGroup can also be used for migrating from clustered ONTAP objects. Future releases will provide more options for migrating from FlexVol and Infinite Volume to NetApp 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.

12.4 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:

• Data migration
• File or directory listings (a high-performance, flexible alternative to `ls`)
• Space reporting (a high-performance, flexible alternative to `du`)

Sometimes, XCP has 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—a performance improvement of 400 times.

XCP also gives some handy reporting graphs, as shown in Figure 103.
12.5 Using XCP to Scan Files Before Migration

When deploying a FlexGroup volume, evaluate the file system and structure to help you determine the best way to lay out member volumes, and initial sizing considerations. 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.

```bash
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:

```bash
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**

  ```
<table>
<thead>
<tr>
<th>Size</th>
<th>Depth</th>
<th>Namelen</th>
<th>Dirsize</th>
</tr>
</thead>
<tbody>
<tr>
<td>340MiB</td>
<td>9</td>
<td>86</td>
<td>500</td>
</tr>
<tr>
<td>1.61KiB</td>
<td>4</td>
<td>6</td>
<td>11</td>
</tr>
</tbody>
</table>
  ```

- **Top file extensions**

  ```
<table>
<thead>
<tr>
<th>.docx</th>
<th>.png</th>
<th>.pptx</th>
<th>.pdf</th>
<th>.css</th>
<th>other</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000038</td>
<td>260</td>
<td>175</td>
<td>128</td>
<td>91</td>
<td>33</td>
</tr>
</tbody>
</table>
  ```

- **Number of files, broken down by size ranges**

  ```
<table>
<thead>
<tr>
<th>empty</th>
<th>&lt;8KiB</th>
<th>8-64KiB</th>
<th>64KiB-1MiB</th>
<th>1-10MiB</th>
<th>10-100MiB</th>
<th>&gt;100MiB</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
  ```

For more information, see the official XCP website at [http://xcp.netapp.com](http://xcp.netapp.com).
### Space used by size range

<table>
<thead>
<tr>
<th>Size Range</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>empty</td>
<td>0</td>
</tr>
<tr>
<td>&lt;8KiB</td>
<td>0</td>
</tr>
<tr>
<td>8-64KiB</td>
<td>28.7MiB</td>
</tr>
<tr>
<td>64KiB-1MiB</td>
<td>3.94MiB</td>
</tr>
<tr>
<td>1-10MiB</td>
<td>124MiB</td>
</tr>
<tr>
<td>10-100MiB</td>
<td>695MiB</td>
</tr>
<tr>
<td>&gt;100MiB</td>
<td>272MiB</td>
</tr>
<tr>
<td>&gt;453MiB</td>
<td>453MiB</td>
</tr>
</tbody>
</table>

### Directory entries, broken down by file counts

<table>
<thead>
<tr>
<th>Count Range</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>empty</td>
<td>7</td>
</tr>
<tr>
<td>1-10</td>
<td>100118</td>
</tr>
<tr>
<td>10-100</td>
<td>30</td>
</tr>
<tr>
<td>100-1K</td>
<td>200</td>
</tr>
<tr>
<td>1K-10K</td>
<td></td>
</tr>
<tr>
<td>&gt;10k</td>
<td></td>
</tr>
</tbody>
</table>

### Directory depth ranges

<table>
<thead>
<tr>
<th>Depth Range</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-5</td>
<td>1100966</td>
</tr>
<tr>
<td>6-10</td>
<td>333</td>
</tr>
<tr>
<td>11-15</td>
<td></td>
</tr>
<tr>
<td>16-20</td>
<td></td>
</tr>
<tr>
<td>21-100</td>
<td></td>
</tr>
<tr>
<td>&gt;100</td>
<td></td>
</tr>
</tbody>
</table>

### Modified and created date ranges

#### Modified

<table>
<thead>
<tr>
<th>Date Range</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;1 year</td>
<td>579</td>
</tr>
<tr>
<td>&gt;1 month</td>
<td>1100559</td>
</tr>
<tr>
<td>1-31 days</td>
<td>11</td>
</tr>
<tr>
<td>1-24 hrs</td>
<td></td>
</tr>
<tr>
<td>&lt;1 hour</td>
<td></td>
</tr>
<tr>
<td>&lt;15 mins</td>
<td></td>
</tr>
<tr>
<td>future</td>
<td>150</td>
</tr>
</tbody>
</table>

#### Created

<table>
<thead>
<tr>
<th>Date Range</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;1 year</td>
<td>1100210</td>
</tr>
<tr>
<td>&gt;1 month</td>
<td>1089</td>
</tr>
<tr>
<td>1-31 days</td>
<td></td>
</tr>
<tr>
<td>1-24 hrs</td>
<td></td>
</tr>
<tr>
<td>&lt;1 hour</td>
<td></td>
</tr>
<tr>
<td>&lt;15 mins</td>
<td></td>
</tr>
<tr>
<td>future</td>
<td></td>
</tr>
</tbody>
</table>

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 104:

```bash
cxp scan -stats -html demo:/flexgroup_16 > /flexgroup.html
```
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 '.*?/.*?/'
```
Where to Find Additional Information

To learn more about the information described in this document, refer to the following documents and/or websites:

Technical Reports

- NetApp Thin Provisioning Deployment and Implementation Guide
- TR-3982: NetApp Clustered Data ONTAP 8.3.x and 8.2.x
- TR-4037: Introduction to NetApp Infinite Volume
- NFS Best Practice and Implementation Guide
- TR-4379: Name Services Best Practices Guide (pre-ONTAP 9.3)
- TR-4668: Name Services Best Practices Guide (ONTAP 9.3 and later)
- NetApp Data Compression, Deduplication, and Data Compaction
- NetApp Storage Solutions for Apache Spark
- NetApp FlexGroup Best Practices and Implementation Guide

NetApp FlexGroup Top Best Practices

- Electronic Design Automation Best Practices

FabricPool Best Practices

Miscellaneous Content

- Tech ONTAP Podcast—Episode 46: FlexGroups
  https://soundcloud.com/techontap_podcast/episode-46-flexgroups
- What’s New For FlexGroup Volumes in ONTAP 9.3?
- FlexGroup Volumes: An Evolution of NAS
- 7 Myths about NetApp 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=Wp6jEd4Vkgl&t=4s
EMS Examples

Example of Maxdirsize EMS 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 message 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 EMS Messages

<table>
<thead>
<tr>
<th>Message Name</th>
<th>Severity</th>
<th>Corrective Action</th>
<th>Description</th>
<th>Supports SNMP trap</th>
<th>Destinations</th>
<th>Number of Drops Between Transmissions</th>
<th>Dropping Interval (Seconds) Between Transmissions</th>
</tr>
</thead>
<tbody>
<tr>
<td>monitor.volume.full</td>
<td>DEBUG</td>
<td>(NONE)</td>
<td>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 &quot;Snapshot Reserve&quot; field from the value of the &quot;Used&quot; field of the &quot;volume show-space&quot; 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 &quot;vol.log.overalloc&quot; EMS message for more information.</td>
<td>true</td>
<td>-</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>monitor.volume.nearlyFull</td>
<td>ALERT</td>
<td>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 &quot;volume size&quot; command. To delete a volume's Snapshot(R) copies, use the &quot;volume snapshot delete&quot; command. To increase an aggregate's size, add disks by using the &quot;storage aggregate add-disks&quot; command. Aggregate Snapshot(R) copies are deleted automatically when the aggregate is full.</td>
<td>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 &quot;Snapshot Reserve&quot; field from the value of the &quot;Used&quot; field of the &quot;volume show-space&quot; command.</td>
<td>true</td>
<td>-</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>monitor.volume.ok</td>
<td>DEBUG</td>
<td>(UNKNOWN)</td>
<td>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.</td>
<td>true</td>
<td>-</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Message Name: monitor.volumes.one.ok
Severity: DEBUG
Corrective Action: (NONE)

Description: This message occurs 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
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%
```

```
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%
```
### Vserver : SVM
#### Volume : flexgroup__0003

<table>
<thead>
<tr>
<th>Feature</th>
<th>Used</th>
<th>Used%</th>
</tr>
</thead>
<tbody>
<tr>
<td>User Data</td>
<td>57.02MB</td>
<td>0%</td>
</tr>
<tr>
<td>Filesystem Metadata</td>
<td>3.66MB</td>
<td>0%</td>
</tr>
<tr>
<td>Inodes</td>
<td>84.55MB</td>
<td>0%</td>
</tr>
<tr>
<td>Snapshot Reserve</td>
<td>512GB</td>
<td>5%</td>
</tr>
<tr>
<td>Deduplication</td>
<td>12KB</td>
<td>0%</td>
</tr>
<tr>
<td>Performance Metadata</td>
<td>44KB</td>
<td>0%</td>
</tr>
<tr>
<td>Total Used</td>
<td>512.1GB</td>
<td>5%</td>
</tr>
<tr>
<td>Total Physical Used</td>
<td>145.6MB</td>
<td>0%</td>
</tr>
</tbody>
</table>

### Vserver : SVM
#### Volume : flexgroup__0004

<table>
<thead>
<tr>
<th>Feature</th>
<th>Used</th>
<th>Used%</th>
</tr>
</thead>
<tbody>
<tr>
<td>User Data</td>
<td>57.19MB</td>
<td>0%</td>
</tr>
<tr>
<td>Filesystem Metadata</td>
<td>8.93MB</td>
<td>0%</td>
</tr>
<tr>
<td>Inodes</td>
<td>82.09MB</td>
<td>0%</td>
</tr>
<tr>
<td>Snapshot Reserve</td>
<td>512GB</td>
<td>5%</td>
</tr>
<tr>
<td>Deduplication</td>
<td>12KB</td>
<td>0%</td>
</tr>
<tr>
<td>Performance Metadata</td>
<td>44KB</td>
<td>0%</td>
</tr>
<tr>
<td>Total Used</td>
<td>512.1GB</td>
<td>5%</td>
</tr>
<tr>
<td>Total Physical Used</td>
<td>148.5MB</td>
<td>0%</td>
</tr>
</tbody>
</table>

### Vserver : SVM
#### Volume : flexgroup__0005

<table>
<thead>
<tr>
<th>Feature</th>
<th>Used</th>
<th>Used%</th>
</tr>
</thead>
<tbody>
<tr>
<td>User Data</td>
<td>3.99GB</td>
<td>0%</td>
</tr>
<tr>
<td>Filesystem Metadata</td>
<td>4.88MB</td>
<td>0%</td>
</tr>
<tr>
<td>Inodes</td>
<td>83.54MB</td>
<td>0%</td>
</tr>
<tr>
<td>Snapshot Reserve</td>
<td>512GB</td>
<td>5%</td>
</tr>
<tr>
<td>Deduplication</td>
<td>12KB</td>
<td>0%</td>
</tr>
<tr>
<td>Performance Metadata</td>
<td>52KB</td>
<td>0%</td>
</tr>
<tr>
<td>Total Used</td>
<td>516.1GB</td>
<td>5%</td>
</tr>
<tr>
<td>Total Physical Used</td>
<td>4.08GB</td>
<td>0%</td>
</tr>
</tbody>
</table>

### Vserver : SVM
#### Volume : flexgroup__0006

<table>
<thead>
<tr>
<th>Feature</th>
<th>Used</th>
<th>Used%</th>
</tr>
</thead>
<tbody>
<tr>
<td>User Data</td>
<td>57.04MB</td>
<td>0%</td>
</tr>
<tr>
<td>Filesystem Metadata</td>
<td>3.50MB</td>
<td>0%</td>
</tr>
<tr>
<td>Inodes</td>
<td>87.26MB</td>
<td>0%</td>
</tr>
<tr>
<td>Snapshot Reserve</td>
<td>512GB</td>
<td>5%</td>
</tr>
<tr>
<td>Deduplication</td>
<td>12KB</td>
<td>0%</td>
</tr>
<tr>
<td>Performance Metadata</td>
<td>44KB</td>
<td>0%</td>
</tr>
<tr>
<td>Total Used</td>
<td>512.1GB</td>
<td>5%</td>
</tr>
<tr>
<td>Total Physical Used</td>
<td>148.2MB</td>
<td>0%</td>
</tr>
</tbody>
</table>

### Vserver : SVM
#### Volume : flexgroup__0007

<table>
<thead>
<tr>
<th>Feature</th>
<th>Used</th>
<th>Used%</th>
</tr>
</thead>
</table>
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                               57.02MB         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%

cluster::> vol show -is-constituent true -volume flexgroup__*

<table>
<thead>
<tr>
<th>Vserver</th>
<th>Volume</th>
<th>Aggregate</th>
<th>State</th>
<th>Type</th>
<th>Size</th>
<th>Available</th>
<th>Used</th>
<th>Used%</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM</td>
<td>flexgroup__0001</td>
<td>aggr1_node1</td>
<td>online</td>
<td>RW</td>
<td>10TB</td>
<td>5.05TB</td>
<td>49%</td>
<td></td>
</tr>
<tr>
<td>SVM</td>
<td>flexgroup__0002</td>
<td>aggr1_node2</td>
<td>online</td>
<td>RW</td>
<td>10TB</td>
<td>5.08TB</td>
<td>49%</td>
<td></td>
</tr>
<tr>
<td>SVM</td>
<td>flexgroup__0003</td>
<td>aggr1_node1</td>
<td>online</td>
<td>RW</td>
<td>10TB</td>
<td>5.05TB</td>
<td>49%</td>
<td></td>
</tr>
<tr>
<td>SVM</td>
<td>flexgroup__0004</td>
<td>aggr1_node2</td>
<td>online</td>
<td>RW</td>
<td>10TB</td>
<td>5.08TB</td>
<td>49%</td>
<td></td>
</tr>
<tr>
<td>SVM</td>
<td>flexgroup__0005</td>
<td>aggr1_node1</td>
<td>online</td>
<td>RW</td>
<td>10TB</td>
<td>5.05TB</td>
<td>49%</td>
<td></td>
</tr>
<tr>
<td>SVM</td>
<td>flexgroup__0006</td>
<td>aggr1_node2</td>
<td>online</td>
<td>RW</td>
<td>10TB</td>
<td>5.08TB</td>
<td>49%</td>
<td></td>
</tr>
<tr>
<td>SVM</td>
<td>flexgroup__0007</td>
<td>aggr1_node1</td>
<td>online</td>
<td>RW</td>
<td>10TB</td>
<td>5.05TB</td>
<td>49%</td>
<td></td>
</tr>
<tr>
<td>SVM</td>
<td>flexgroup__0008</td>
<td>aggr1_node2</td>
<td>online</td>
<td>RW</td>
<td>10TB</td>
<td>5.08TB</td>
<td>49%</td>
<td></td>
</tr>
</tbody>
</table>

8 entries were displayed.

cluster::> storage aggregate show -aggregate aggr1* -fields usedsize,size,percent-used -sort-by percent-used aggregate percent-used size usedsize

<table>
<thead>
<tr>
<th>Aggregate</th>
<th>percent-used</th>
<th>usedsize</th>
<th>size</th>
</tr>
</thead>
<tbody>
<tr>
<td>aggr1_node1</td>
<td>26%</td>
<td>7.86TB</td>
<td>2.05TB</td>
</tr>
<tr>
<td>aggr1_node2</td>
<td>26%</td>
<td>7.86TB</td>
<td>2.02TB</td>
</tr>
</tbody>
</table>

2 entries were displayed.
### Example of statistics show-periodic Command for Entire Cluster

```
cluster::*> statistics show-periodic
cpu cpu total fcache total total data data data
cpu cluster cluster: cluster.cluster: 11/30/2016 11:49:46
cpu cpu total fcache total total data data data
cpu cluster cluster: cluster.cluster: 11/30/2016 11:49:53
```

### Real-Time SVM-Level statistics show-periodic for NFSv3 Read and Write Operations Only

```
cpu cpu total fcache total total data data data
cpu cluster cluster: cluster.cluster: 11/30/2016 11:49:46
cpu cpu total fcache total total data data data
cpu cluster cluster: cluster.cluster: 11/30/2016 11:49:53
```
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
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
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.

```json
{
  "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"
      },
```
This is what the FlexGroup looks like after it's created:

```
cluster::*> vol show -vserver DEMO -volume REST*

<table>
<thead>
<tr>
<th>Vserver</th>
<th>Volume</th>
<th>Aggregate</th>
<th>State</th>
<th>Type</th>
<th>Size</th>
<th>Available</th>
<th>Used%</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEMO</td>
<td>RESTAPI_FG</td>
<td>-</td>
<td>online</td>
<td>RW</td>
<td>2TB</td>
<td>1.90TB</td>
<td>0%</td>
</tr>
<tr>
<td>DEMO</td>
<td>RESTAPI_FG__0001</td>
<td>aggr1_node1</td>
<td>online</td>
<td>RW</td>
<td>256GB</td>
<td>243.1GB</td>
<td>0%</td>
</tr>
<tr>
<td>DEMO</td>
<td>RESTAPI_FG__0002</td>
<td>aggr1_node1</td>
<td>online</td>
<td>RW</td>
<td>256GB</td>
<td>243.1GB</td>
<td>0%</td>
</tr>
<tr>
<td>DEMO</td>
<td>RESTAPI_FG__0003</td>
<td>aggr1_node1</td>
<td>online</td>
<td>RW</td>
<td>256GB</td>
<td>243.1GB</td>
<td>0%</td>
</tr>
<tr>
<td>DEMO</td>
<td>RESTAPI_FG__0004</td>
<td>aggr1_node1</td>
<td>online</td>
<td>RW</td>
<td>256GB</td>
<td>243.1GB</td>
<td>0%</td>
</tr>
<tr>
<td>DEMO</td>
<td>RESTAPI_FG__0005</td>
<td>aggr1_node1</td>
<td>online</td>
<td>RW</td>
<td>256GB</td>
<td>243.1GB</td>
<td>0%</td>
</tr>
<tr>
<td>DEMO</td>
<td>RESTAPI_FG__0006</td>
<td>aggr1_node1</td>
<td>online</td>
<td>RW</td>
<td>256GB</td>
<td>243.1GB</td>
<td>0%</td>
</tr>
<tr>
<td>DEMO</td>
<td>RESTAPI_FG__0007</td>
<td>aggr1_node1</td>
<td>online</td>
<td>RW</td>
<td>256GB</td>
<td>243.1GB</td>
<td>0%</td>
</tr>
<tr>
<td>DEMO</td>
<td>RESTAPI_FG__0008</td>
<td>aggr1_node1</td>
<td>online</td>
<td>RW</td>
<td>256GB</td>
<td>243.1GB</td>
<td>0%</td>
</tr>
</tbody>
</table>

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:

```bash
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 via the "style" option and does not specify the "constituents_per_aggregate" option.

```json
{
  "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 resultant FlexGroup:

```bash
cluster::*> vol show -vserver DEMO -volume RESTAPI_FG2*
Vserver   Volume       Aggregate    State      Type       Size  Available Used%
--------- ------------ ------------ ---------- ---- ---------- ---------- ----- 
DEMO      RESTAPI_FG2 - online     RW        2T        1.8879TB    0% 
DEMO      RESTAPI_FG2__0001 aggr1_node1 online RW        256GB    243.1GB    0% 
DEMO      RESTAPI_FG2__0002 aggr1_node2 online RW        256GB    243.1GB    0% 
DEMO      RESTAPI_FG2__0003 aggr1_node1 online RW        256GB    243.1GB    0% 
DEMO      RESTAPI_FG2__0004 aggr1_node2 online RW        256GB    243.1GB    0% 
DEMO      RESTAPI_FG2__0005 aggr1_node1 online RW        256GB    243.1GB    0% 
DEMO      RESTAPI_FG2__0006 aggr1_node2 online RW        256GB    243.1GB    0% 
DEMO      RESTAPI_FG2__0007 aggr1_node1 online RW        256GB    243.1GB    0% 
DEMO      RESTAPI_FG2__0008 aggr1_node2 online RW        256GB    243.1GB    0% 
9 entries were displayed.
```
Example of Increasing a FlexGroup Volume’s Size

```sh
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**

```bash
cluster::*> volume show -vserver SVM -volume flexgroup4*

<table>
<thead>
<tr>
<th>Vserver</th>
<th>Volume</th>
<th>Aggregate</th>
<th>State</th>
<th>Type</th>
<th>Size</th>
<th>Available</th>
<th>Used%</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM</td>
<td>flexgroup4TB -</td>
<td></td>
<td></td>
<td>RW</td>
<td>4TB</td>
<td>3.78TB</td>
<td>5%</td>
</tr>
<tr>
<td>SVM</td>
<td>flexgroup4TB__0001</td>
<td>aggr1_node1</td>
<td>online</td>
<td>RW</td>
<td>512GB</td>
<td>485.5GB</td>
<td>5%</td>
</tr>
<tr>
<td>SVM</td>
<td>flexgroup4TB__0002</td>
<td>aggr1_node2</td>
<td>online</td>
<td>RW</td>
<td>512GB</td>
<td>481.2GB</td>
<td>6%</td>
</tr>
<tr>
<td>SVM</td>
<td>flexgroup4TB__0003</td>
<td>aggr1_node1</td>
<td>online</td>
<td>RW</td>
<td>512GB</td>
<td>481.2GB</td>
<td>6%</td>
</tr>
<tr>
<td>SVM</td>
<td>flexgroup4TB__0004</td>
<td>aggr1_node2</td>
<td>online</td>
<td>RW</td>
<td>512GB</td>
<td>481.5GB</td>
<td>5%</td>
</tr>
<tr>
<td>SVM</td>
<td>flexgroup4TB__0005</td>
<td>aggr1_node1</td>
<td>online</td>
<td>RW</td>
<td>512GB</td>
<td>481.5GB</td>
<td>5%</td>
</tr>
<tr>
<td>SVM</td>
<td>flexgroup4TB__0006</td>
<td>aggr1_node2</td>
<td>online</td>
<td>RW</td>
<td>512GB</td>
<td>481.5GB</td>
<td>5%</td>
</tr>
<tr>
<td>SVM</td>
<td>flexgroup4TB__0007</td>
<td>aggr1_node1</td>
<td>online</td>
<td>RW</td>
<td>512GB</td>
<td>485.5GB</td>
<td>5%</td>
</tr>
<tr>
<td>SVM</td>
<td>flexgroup4TB__0008</td>
<td>aggr1_node2</td>
<td>online</td>
<td>RW</td>
<td>512GB</td>
<td>485.5GB</td>
<td>5%</td>
</tr>
</tbody>
</table>

cluster::*> volume expand -vserver SVM -volume flexgroup4TB -aggr-list aggr1_node1,aggr1_node2 -aggr-list-multiplier 4

cluster::*> volume show -vserver SVM -volume flexgroup4*

<table>
<thead>
<tr>
<th>Vserver</th>
<th>Volume</th>
<th>Aggregate</th>
<th>State</th>
<th>Type</th>
<th>Size</th>
<th>Available</th>
<th>Used%</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM</td>
<td>flexgroup4TB -</td>
<td></td>
<td></td>
<td>RW</td>
<td>8TB</td>
<td>7.78TB</td>
<td>1%</td>
</tr>
<tr>
<td>SVM</td>
<td>flexgroup4TB__0001</td>
<td>aggr1_node1</td>
<td>online</td>
<td>RW</td>
<td>512GB</td>
<td>485.5GB</td>
<td>1%</td>
</tr>
<tr>
<td>SVM</td>
<td>flexgroup4TB__0002</td>
<td>aggr1_node2</td>
<td>online</td>
<td>RW</td>
<td>512GB</td>
<td>481.2GB</td>
<td>1%</td>
</tr>
<tr>
<td>SVM</td>
<td>flexgroup4TB__0003</td>
<td>aggr1_node1</td>
<td>online</td>
<td>RW</td>
<td>512GB</td>
<td>481.2GB</td>
<td>1%</td>
</tr>
<tr>
<td>SVM</td>
<td>flexgroup4TB__0004</td>
<td>aggr1_node2</td>
<td>online</td>
<td>RW</td>
<td>512GB</td>
<td>481.5GB</td>
<td>1%</td>
</tr>
<tr>
<td>SVM</td>
<td>flexgroup4TB__0005</td>
<td>aggr1_node1</td>
<td>online</td>
<td>RW</td>
<td>512GB</td>
<td>485.5GB</td>
<td>1%</td>
</tr>
<tr>
<td>SVM</td>
<td>flexgroup4TB__0006</td>
<td>aggr1_node2</td>
<td>online</td>
<td>RW</td>
<td>512GB</td>
<td>485.5GB</td>
<td>1%</td>
</tr>
<tr>
<td>SVM</td>
<td>flexgroup4TB__0007</td>
<td>aggr1_node1</td>
<td>online</td>
<td>RW</td>
<td>512GB</td>
<td>485.5GB</td>
<td>1%</td>
</tr>
<tr>
<td>SVM</td>
<td>flexgroup4TB__0008</td>
<td>aggr1_node2</td>
<td>online</td>
<td>RW</td>
<td>512GB</td>
<td>485.5GB</td>
<td>1%</td>
</tr>
<tr>
<td>SVM</td>
<td>flexgroup4TB__0009</td>
<td>aggr1_node1</td>
<td>online</td>
<td>RW</td>
<td>512GB</td>
<td>485.5GB</td>
<td>1%</td>
</tr>
<tr>
<td>SVM</td>
<td>flexgroup4TB__0010</td>
<td>aggr1_node2</td>
<td>online</td>
<td>RW</td>
<td>512GB</td>
<td>481.2GB</td>
<td>1%</td>
</tr>
<tr>
<td>SVM</td>
<td>flexgroup4TB__0011</td>
<td>aggr1_node1</td>
<td>online</td>
<td>RW</td>
<td>512GB</td>
<td>481.5GB</td>
<td>1%</td>
</tr>
<tr>
<td>SVM</td>
<td>flexgroup4TB__0012</td>
<td>aggr1_node2</td>
<td>online</td>
<td>RW</td>
<td>512GB</td>
<td>485.5GB</td>
<td>1%</td>
</tr>
<tr>
<td>SVM</td>
<td>flexgroup4TB__0013</td>
<td>aggr1_node1</td>
<td>online</td>
<td>RW</td>
<td>512GB</td>
<td>485.5GB</td>
<td>1%</td>
</tr>
<tr>
<td>SVM</td>
<td>flexgroup4TB__0014</td>
<td>aggr1_node2</td>
<td>online</td>
<td>RW</td>
<td>512GB</td>
<td>485.5GB</td>
<td>1%</td>
</tr>
<tr>
<td>SVM</td>
<td>flexgroup4TB__0015</td>
<td>aggr1_node1</td>
<td>online</td>
<td>RW</td>
<td>512GB</td>
<td>485.5GB</td>
<td>1%</td>
</tr>
<tr>
<td>SVM</td>
<td>flexgroup4TB__0016</td>
<td>aggr1_node2</td>
<td>online</td>
<td>RW</td>
<td>512GB</td>
<td>485.5GB</td>
<td>1%</td>
</tr>
</tbody>
</table>
```
How to Create a SnapMirror and SnapVault Relationship (ONTAP 9.3 and earlier)

For more information about using NetApp SnapMirror and SnapVault, see TR-4678.

1. Open ONTAP System Manager. Under Volumes, determine whether the volume is unprotected. The volume to be protected is Tech_ONTAP (the NetApp Tech OnTap® podcast volume).

   ![Screenshot of ONTAP System Manager showing volumes]

   Note: Click the + sign to show more details.

2. Click Configuration in the left menu, and then select SVM Peers (for local SnapMirror) or Cluster Peers (for remote SnapMirror) if you haven’t already peered the source and destination. In this example, we’re peering SVMs to do a local SnapMirror configuration.

   ![Screenshot of ONTAP System Manager showing SVM peers]

3. Peer the SVM or cluster.

   SVM peering. Click Create and choose your SVMs. Click Initiate SVM Peering. Within a few seconds, you should see SVM Peering Successful. Click Done.
Cluster peering. Cluster peering is needed if you plan to implement intercluster SnapMirror. In ONTAP System Manager for ONTAP 9.3, this is done with Cluster Peers. Using this option also allows you to peer SVMs in the same configuration steps.

a. Select Protection > Relationships in the left menu. Then click Create.
b. Go to the CLI; System Manager doesn’t support creating SnapMirror relationships for FlexGroup volumes in ONTAP 9.3 and earlier. ONTAP 9.4 adds support for the simple creation of FlexGroup SnapMirror relationships in System Manager. Create the destination volume. It must be type data protection and have the same number of member volumes as the source FlexGroup volume. It must also be the same size as or greater than the source FlexGroup volume.

c. The Tech_ONTAP FlexGroup volume has eight-member volumes and is 10TB in size.

```
cluster::> vol show -vserver DEMO -volume Tech_ONTAP* -is-constituent true -fields size
vserver volume           size
------- ---------------- -----
DEMO    Tech_ONTAP__0001 1.25TB
DEMO    Tech_ONTAP__0002 1.25TB
DEMO    Tech_ONTAP__0003 1.25TB
DEMO    Tech_ONTAP__0004 1.25TB
DEMO    Tech_ONTAP__0005 1.25TB
DEMO    Tech_ONTAP__0006 1.25TB
DEMO    Tech_ONTAP__0007 1.25TB
DEMO    Tech_ONTAP__0008 1.25TB
8 entries were displayed.
cluster::> vol show -vserver DEMO -volume Tech_ONTAP* -fields size
vserver volume     size
------- ---------- -----
DEMO    Tech_ONTAP 10TB
```

d. The destination volume is created. It must match the source volume’s member volume count and must be as large as or larger than the source. Otherwise, a geometry error occurs.

```
Error: command failed: Geometry of the source FlexGroup does not match to that of the destination FlexGroup. Delete the relationship, fix the destination FlexGroup to match its geometry with that of the source FlexGroup and then recreate the SnapMirror relationship before running the operation again.
```

e. The FlexGroup volume spans a single aggregate and uses a multiplier of eight to create eight-member volumes per aggregate. So, in this case, it’s a FlexGroup volume with eight-member volumes.

```
cluster::*> vol create -vserver SVM1 -volume Tech_ONTAP_mirror -aggr-list aggr1_node1 -aggr-list-multiplier 8 -state online -type DP -size 10t -space-guarantee none
Warning: The FlexGroup "Tech_ONTAP_mirror" will be created with the following number of constituents of size 1.25TB: 8.
Do you want to continue? [y/n]: y
[Job 8460] Job succeeded: Successful
```

4. Decide on a SnapMirror policy. You can apply various policies to the mirror. For disaster recovery SnapMirror, use "MirrorAllSnapshots.” If not specified, this is the default. SnapVault and MirrorVault would use MirrorAndVault.

Below are the specific policies and what they do:
5. Create the SnapMirror relationship.

```
class::> snapmirror policy show -fields comment
vserver        policy    comment
------------------- --------- -----------------------------------
ontap9-tme-8040 DPDefault Default policy for DP relationship.
ontap9-tme-8040 MirrorAllSnapshots Asynchronous SnapMirror policy for mirroring all snapshots and the latest active file system.
ontap9-tme-8040 MirrorAndVault A unified Asynchronous SnapMirror and SnapVault policy for mirroring the latest active file system and daily and weekly snapshots.
ontap9-tme-8040 MirrorLatest Asynchronous SnapMirror policy for mirroring the latest active file system.
ontap9-tme-8040 Unified7year Unified SnapMirror policy with 7 year retention.
ontap9-tme-8040 XDPDefault Default policy for XDP relationship with daily and weekly rules.
6 entries were displayed.
```

```
class::> snapmirror create -source-path DEMO:Tech_ONTAP -destination-path SVM1:Tech_ONTAP_mirror -vserver SVM1 -throttle unlimited -identity-preserve false -type XDP
Operation succeeded: snapmirror create for the relationship with destination "SVM1:FGlocal_mirror".

class::> snapmirror show -destination-path SVM1:Tech_ONTAP_mirror
Source Path: DEMO:Tech_ONTAP
Destination Path: SVM1:Tech_ONTAP_mirror
Relationship Type: XDP
Relationship Group Type: flexgroup
SnapMirror Schedule: -
SnapMirror Policy Type: async-mirror
SnapMirror Policy: MirrorAllSnapshots
Tries Limit: -
Throttle (KB/sec): unlimited
Mirror State: Uninitialized
Relationship Status: Idle
File Restore File Count: -
File Restore File List: -
Transfer Snapshot: -
Snapshot Progress: -
Total Progress: -
Network Compression Ratio: -
Snapshot Checkpoint: -
Newest Snapshot: -
Newest Snapshot Timestamp: -
Exported Snapshot: -
Exported Snapshot Timestamp: -
Healthy: true
Unhealthy Reason: -
Constituent Relationship: false
Destination Volume Node: ontap9-tme-8040-01
Relationship ID: cab5fbc2-c3eb-11e7-8d0f-00a0986b1223
Current Operation ID: -
Transfer Type: -
Transfer Error: -
Current Throttle: -
Current Transfer Priority: -
Last Transfer Type: -
Last Transfer Error: -
Last Transfer Size: -
Last Transfer Network Compression Ratio: -
Last Transfer Duration: -
Last Transfer From: -
Last Transfer End Timestamp: -
Progress Last Updated: -
Relationship Capability: 8.2 and above
Lag Time: -
Identity Preserve Vserver DR: -
```
Volume MSIDs Preserved: -
Is Auto Expand Enabled: true
Number of Successful Updates: 0
Number of Failed Updates: 0
Number of Successful Resyncs: 0
Number of Failed Resyncs: 0
Number of Successful Breaks: 0
Number of Failed Breaks: 0
Total Transfer Bytes: -
Total Transfer Time in Seconds: 0

6. **Initialize the SnapMirror relationship.** This takes some time, depending on the amount of data to transfer and the speed of your connection.

```bash
cluster::> snapmirror initialize -destination-path SVM1:Tech_ONTAP_mirror
Operation is queued: snapmirror initialize of destination "SVM1:Tech_ONTAP_mirror".
```

```bash
cluster::> snapmirror show -destination-path SVM1:Tech_ONTAP_mirror -fields status,state
source-path     destination-path       state         status
--------------- ---------------------- ------------- ------------
DEMO:Tech_ONTAP SVM1:Tech_ONTAP_mirror Uninitialized Transferring
```

Although you cannot currently manage FlexGroup SnapMirror from System Manager, you can view it.

7. **To make the mirror a SnapVault relationship (keeping more Snapshot copies than just the ones on the source), modify the relationship policy to MirrorAndVault.**

```bash
cluster::> snapmirror modify -destination-path SVM1:Tech_ONTAP_mirror -policy MirrorAndVault
Operation succeeded: snapmirror modify for the relationship with destination "SVM1:Tech_ONTAP_mirror".
```

```bash
cluster::> snapmirror show -destination-path SVM1:Tech_ONTAP_mirror -fields policy
source-path     destination-path       policy
--------------- ---------------------- ------------------
DEMO:Tech_ONTAP SVM1:Tech_ONTAP_mirror MirrorAndVault
```

8. **Also consider using a Snapshot policy and label for your volume.**

   **Note:** Labels cannot be used with the async-mirror policy type.

   The policy rules support only two combinations of these labels, either just `sm_created` or both `sm_created` and `all_source_snapshots`. The label is used to define the set of Snapshot copies that you want backed up to the version-flexible SnapMirror secondary volume. Other Snapshot copies on the primary volume are ignored by the version-flexible SnapMirror relationship.

```bash
cluster::*> snapshot policy create -policy flexgroup -enabled true -schedule1 hourly -count1 255 -vserver DEMO -snapmirror-label1 all_source_snapshots
```

9. **Modify the volume to use the new policy.**

```bash
cluster::*> vol modify -vserver DEMO -volume FGmirror -snapshot-policy flexgroup
```

Warning: You are changing the Snapshot policy on volume "FGmirror" to "flexgroup". Any Snapshot copies on this volume that do not match any of the prefixes of the new Snapshot policy will not be deleted.

However, once the new Snapshot policy takes effect, depending on the new retention count, any existing Snapshot copies that continue to use the same prefixes may be deleted. See the 'volume modify' man page for more information.

Do you want to continue? [y/n]: y

[Job 8755] Job succeeded: volume modify succeeded
10. Create a schedule to apply to the SnapMirror relationship using Job Schedule Create and then apply it to your SnapMirror.

```
cluster::*> job schedule cron create -name FGSM -minute 30
cluster::*> job schedule show -name FGSM -instance
Cluster: ontap9-tme-8040
Schedule Name: FGSM
Schedule Type: cron
Description: @:30
```

```
cluster::*> snapmirror modify -destination-path SVM1:FGlocal_mirror -schedule FGSM
Operation succeeded: snapmirror modify for the relationship with destination "SVM1:FGlocal_mirror".
```

After a volume is successfully copied using SnapMirror, System Manager shows that it is protected.

For more information about SnapMirror, see the product documentation.

**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 New 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.
```
Version History

<table>
<thead>
<tr>
<th>Version</th>
<th>Date</th>
<th>Document Version History</th>
</tr>
</thead>
<tbody>
<tr>
<td>Version 1.0</td>
<td>January 2017</td>
<td>Initial release</td>
</tr>
<tr>
<td>Version 1.0.1</td>
<td>February 2017</td>
<td>Minor revisions</td>
</tr>
<tr>
<td>Version 1.1</td>
<td>May 2017</td>
<td>ONTAP 9.2RC1</td>
</tr>
<tr>
<td>Version 1.2</td>
<td>December 2017</td>
<td>ONTAP 9.3GA</td>
</tr>
<tr>
<td>Version 1.3</td>
<td>May 2018</td>
<td>ONTAP 9.4RC1</td>
</tr>
<tr>
<td>Version 1.4</td>
<td>November 2018</td>
<td>ONTAP 9.5</td>
</tr>
<tr>
<td>Version 1.5</td>
<td>June 2019</td>
<td>ONTAP 9.6</td>
</tr>
</tbody>
</table>

Acknowledgments

Special thanks to Chad Morgenstern, Richard Jernigan, Dan Tennant, Ken Cantrell, Sunitha Rao, Shriya Paramkusam, and Mrinal Devadas.
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.

**Copyright Information**

Copyright © 2019 NetApp, Inc. All Rights Reserved. Printed in the U.S. No part of this document covered by copyright may be reproduced in any form or by any means—graphic, electronic, or mechanical, including photocopying, recording, taping, or storage in an electronic retrieval system—without prior written permission of the copyright owner.

Software derived from copyrighted NetApp material is subject to the following license and disclaimer:

**THIS SOFTWARE IS PROVIDED BY NETAPP "AS IS" AND WITHOUT ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE, WHICH ARE HEREBY DISCLAIMED. IN NO EVENT SHALL NETAPP BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.**

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

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

Data contained herein pertains to a commercial item (as defined in FAR 2.101) and is proprietary to NetApp, Inc. The U.S. Government has a non-exclusive, non-transferable, non-sublicensable, worldwide, limited irrevocable license to use the Data only in connection with and in support of the U.S. Government contract under which the Data was delivered. Except as provided herein, the Data may not be used.

**Trademark Information**

NETAPP, the NETAPP logo, and the marks listed at [http://www.netapp.com/TM](http://www.netapp.com/TM) are trademarks of NetApp, Inc. Other company and product names may be trademarks of their respective owners.