Technical Report

NetApp ONTAP FlexGroup Volumes Top Best Practices
ONTAP 9.7

Justin Parisi, NetApp
January 2020 | TR-4571-a

Abstract

This technical report attempts to condense the contents of TR-4571: NetApp ONTAP FlexGroup Volumes Best Practices and Implementation Guide down into easily digestible best practices for quick consumption. For more detailed information, see TR-4571.

Information Classification

Public

■ NetApp®

TABLE OF CONTENTS

1 General Best Practices

The following best practices are condensed from TR-4571. For more details about these best practices, as well as what features are supported or unsupported with FlexGroup volumes, see TR-4571: NetApp FlexGroup Best Practices and Implementation Guide.

For technical information about how FlexGroup volumes work, see TR-4557: NetApp ONTAP FlexGroup Volumes Technical Overview.

Review Your Environment

Before deciding to deploy a FlexGroup volume, review your workload and analyze the I/O types, the average file sizes, the maximum file sizes, and so on. The XCP tool can help scan and analyze the files you have, and NetApp field representatives have methods to analyze workloads. In addition, be sure to review TR-4571 and TR-4557 to determine whether any features your environment requires are unsupported by FlexGroup volumes. To generate reports, use XCP or Data Dynamics File Insight.

Run the Latest ONTAP Version

Every release of ONTAP adds improvements to performance and stability, and includes new features for FlexGroup volumes. For example, ONTAP 9.6 adds support for elastic sizing in a FlexGroup volume, which helps protect against failed writes when a FlexGroup member volume is running out of space. Running the latest release of ONTAP (P-releases included) is the simplest way to get the most out of your FlexGroup volumes.

Homogenous Configuration

When you deploy a FlexGroup volume, the underlying hardware components should be homogenous across all member volumes to help maintain consistent and predictable performance. Node types (FAS versus NetApp NetApp AFF), disk types, spindle counts, and RAID group sizes ideally will be identical. Variations across the FlexGroup volume are supported, but not recommended.

Automated Deployment

FlexGroup volumes are designed with simplicity in mind. Although it is possible to customize FlexGroup volumes, usually it's best to let ONTAP decide how to create them. To create a FlexGroup volume, use the NetApp ONTAP System Manager GUI or the automated CLI commands. See the ONTAP 9 product documentation or TR-4571 for details.

Deploy Across Multiple Nodes and Aggregates

There are some benefits to keeping NetApp FlexVol® and FlexGroup volumes local to the nodes where network access is configured (local fast path, avoiding the cluster network, and so on). However, deploying a FlexGroup volume across multiple nodes and aggregates offers more benefits: increased number of CPU threads/volume affinities, and more efficient CPU, RAM, network, and disk use across the cluster. FlexGroup volumes provide a true scale-out file system container.

Bigger is Better

FlexGroup volumes operate best when member volumes are large enough to accommodate the average file sizes in a file system without getting too close to 70–80% of the member volume sizes. When you choose a FlexGroup volume size, clients see the combined available space of all member volumes, but ONTAP slices that total capacity into {total/number of member volumes}. For example, a 10TB FlexGroup volume with eight member FlexVol volumes has 1.25TB per member. Therefore, it's important to consider the member volume sizes in addition to the total FlexGroup size. Member volumes are held to the same

100TB limits as regular FlexVol volumes, but for best results, they should not be smaller than 100GB. If the FlexGroup size is large, files are more likely to be placed evenly.

Large File Deployments

Large files are defined in [TR-4571: NetApp FlexGroup Best Practices and Implementation Guide](#) as a percentage of the member volume size rather than a flat value. A large file in a FlexGroup volume depends on how small or large the member volume is. For example, a 1TB file might be considered large in a 16TB FlexGroup volume with eight member volumes, because it takes up 50% of the 2TB member volume's capacity immediately. However, that same 1TB file uses only 1% of the capacity of a 100TB member volume. Ideally, a file size would not use up more than 1-5% of the FlexGroup member volume size. Therefore, for environments with large files, it might make sense to manually reduce the number of member volumes in a FlexGroup volume to allow larger member volume sizes.

Networking Considerations

FlexGroup volumes can generate a large amount of throughput and IOPS in NAS environments. However, the storage can only consume the amount of data that can be sent to it. Smaller network links can't push as much data as larger ones, so the FlexGroup volume might not reach its potential with 1GB links. Therefore, NetApp recommends that you use 10GB or greater interfaces on each node in the cluster, and balance mount points across each node to even out network connectivity distribution.

Handling High-File-Count Environments

FlexGroup volumes were primarily designed for high-file-count, large-metadata workloads such as those in EDA and software development environments. The ability to present a multiple FlexVol volume deployment (and, thus, multiple volume affinities) to these workloads with minimal administrative overhead is what FlexGroup volumes do best. Therefore, use the default deployment configurations for these types of workloads.

In addition, ONTAP 9.1 introduces an NFS server option in advanced mode called `-v3-64bit-identifiers`. ONTAP 9.7 also introduces an option for 64-bit IDs for NFSv4.x operations called `-v4-64bit-identifiers`.

These options allows FlexGroup volumes to exceed the normal 32-bit unsigned integer value of 2,147,483,647, which corresponds with the 2 billion file limit in normal FlexVol volumes. When this option is enabled, existing NFS mounts must be remounted to establish the new file handle generated in 64-bit format. Enabling 64-bit file identifiers prevents file ID collisions in FlexGroup volumes and is covered in detail in [TR-4571](#).

Also, because inode counts are enforced for each member volume, NetApp recommends proactively increasing the inode maximums for the FlexGroup volume. Doing so prevents your high-file-count environment from creating an "out of inodes" scenario. See [TR-4571](#) for more details about sizing FlexGroup volumes for high file count environments.

Member Volume Counts - Mapping to Volume and Aggregate Affinities

ONTAP has a set number of volume and aggregate affinities, where each FlexVol volume maps to an affinity at boot. Affinities are essentially FlexVol to CPU mappings.

In ONTAP 9.3 and prior:

- Each volume gets one affinity.
- Each aggregate gets four affinities
- Nodes get a maximum of eight affinities

In ONTAP 9.4 and later (high-end platforms):

- Each volume gets one affinity.
- Each aggregate gets eight affinities
- Nodes get a maximum of 16 affinities.

As a result, one member volume per available affinity will get created. In ONTAP 9.4 and later, up to 16 member volumes per node are created in a FlexGroup volume.

ONTAP automates these values when the volume is created using System Manager or the `-auto-provision-as` option with the `volume create` command. Therefore, the best practice for member volume layout in most environments is to allow ONTAP to handle the creation of the FlexGroup volume.

For more information about volume affinities, see TR-4571.

Thin Provisioning

A FlexGroup volume is built with capacity and performance in mind, and NetApp recommends sizing a FlexGroup volume as large as possible to provide more consistent ingest of files. Therefore, NetApp recommends using thin provisioning for a FlexGroup volume to allow large container footprints without reserving large amounts of space in the cluster. Naturally, the same general recommendations for thin provisioning in ONTAP are applicable. However, with a FlexGroup volume, it's also important to consider what other space is being used in an aggregate when you deploy a FlexGroup volume with thin provisioning. If other FlexVol volumes are present and have space guarantees enabled, or if some aggregates are used more than other aggregates that host member volumes in the same FlexGroup, the ingest algorithms can be skewed based on total free capacity in those aggregates. Again, homogenous environments are best for FlexGroup volumes. TR-4571 goes into more detail regarding thin provisioning.

Increasing the Size of a FlexGroup Volume

Although an individual member volume is limited to 100TB, a FlexGroup volume has virtually no limitations in how large it can grow. However, keep these considerations in mind when expanding a FlexGroup volume:

- Expand the member volumes before adding new member volumes. Expanding member volumes preserves the data layout, which maintains the ingest algorithm's calculations.
- If adding new nodes to a cluster, use volume moves to balance members across the new nodes and then add new members in even multiples across all nodes rather than growing existing volumes.
- To maintain the volume affinity balance in the cluster, when possible, add new members in multiples of four or eight per node. Adding new member volumes might create ingest "hot spots," like adding new disks to RAID groups. Because there is no rebalance of data, adding multiple members can help offset this problem by providing parallel processing of new data requests.

Note: Keep in mind FlexGroup member volume count limits for features such as NetApp SnapMirror® (32 per FlexGroup, 100 per cluster) in ONTAP 9.4 and prior. ONTAP 9.5 and later removes those limits. See TR-4678: FlexGroup Volumes Data Protection Best Practices for details. Also remember that as the number of members increases, so do overhead for remote access and intervals for Snapshots copy.

Adding Nodes or Removing Nodes from a Cluster

When adding new nodes to a cluster, use volume move to migrate member volumes to the new nodes. This approach balances the load of the FlexGroup volume across nodes, provided those new nodes are identical to other nodes in the FlexGroup. For example, suppose that you have a 2-node NetApp AFF A700 all-flash storage system that has a 32 member FlexGroup (16 members per node). If you're adding 2 new A700 nodes, nondisruptively move 16 (8 from each node) of the volumes to the new HA pair to have 8 member volumes per node in the 4-node cluster. Then, if desired, use volume expand to add new members to the FlexGroup volume in the same multiples as the existing FlexGroup member volumes. In this scenario, you would add 32 new members (8 per node in the 4-node A700 cluster).

When removing nodes from a cluster, keep in mind that from ONTAP 9.6 you can shrink a FlexGroup volume. However, you cannot remove member volumes, because of the underlying hard links that might point to other files. In this case, there are two options:

- Use a combination of volume move/shrink and thin provisioning to migrate the existing member volumes off the nodes to be retired and onto nodes that will remain in the cluster. Do your best to maintain the best practices of using homogenous hardware types.
- Create a FlexGroup volume that spans only the nodes that will remain in the cluster, and migrate the existing FlexGroup data by using file-based migration tools such as XCP.

Storage Efficiencies

NetApp FlexGroup volumes support most storage efficiency features in ONTAP, including:

- Thin provisioning
- Deduplication (inline and postprocess; aggregate and volume-level)
- Inline data compaction
- Data compression (inline and postprocess)
- FabricPool (starting in ONTAP 9.5)

Inline storage efficiencies operate identically for FlexGroup and FlexVol volumes. However, postprocess storage efficiencies currently operate a little differently for a FlexGroup volume.

In a FlexVol volume, postprocess deduplication can deduplicate identical files in a single volume because the blocks reside in the same container. However, a FlexGroup volume uses multiple FlexVol volumes, and the data placement is not controlled by the client or storage administrator, but by ONTAP.

Therefore, if an identical file is written twice to a single member volume, you see postprocess deduplication work as expected. If the two identical files are written to two different members, you don't see any space savings for postprocess deduplication, unless you are using aggregate-level deduplication (inline in 9.2; postprocess starting in ONTAP 9.4 – AFF only).

Note: If space guarantee is being changed on volumes that have storage efficiencies in place, undo commands might need to be run to allow the changes to complete. One exception to managing storage efficiencies at the FlexGroup level is that volume efficiency undo commands must be run against the member volumes at advanced privilege. See TR-4571 for more information.

Migrating to FlexGroup Volumes

Prior to ONTAP 9.7, migrating from FlexVol volumes or competitor storage arrays to a FlexGroup volume is file based. You can use any file-based migration utility that operates over NAS (NFS or SMB) to migrate your data to a FlexGroup volume. NetApp often recommends the use of XCP for NFSv3 migrations because of its superior speed and robust functionality. Using SnapMirror to migrate from FlexVol to FlexGroup is not supported.

FlexVol to FlexGroup Conversion

ONTAP 9.7 introduces a new feature that allows a FlexVol volume to be converted to a single member volume FlexGroup. This is done in-place and takes less than 40 seconds to complete, regardless of capacity or file count in the FlexVol volume. Clients won't need to remount during the process. For more information, see TR-4571: NetApp FlexGroup Best Practices and Implementation Guide, or check out this Tech ONTAP Podcast episode on FlexVol to FlexGroup conversion.

For more details, see the section on migrating to FlexGroup volumes in TR-4571: NetApp FlexGroup Best Practices and Implementation Guide.

For more information, email us at flexgroups-info@netapp.com.

Version History

Version	Date	Document Version History
Version 1.0	December 2018	ONTAP 9.5
Version 1.1	June 2019	ONTAP 9.6
Version 1.2	January 2020	ONTAP 9.7

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

Trademark Information

■ NetApp®