Technical Report

NetApp ONTAP FlexGroup Volumes

Top best practices

ONTAP TME Team, NetApp
December 2023 | 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.

TABLE OF CONTENTS

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 NetApp® ONTAP® 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.

General best practice #1: 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 NetApp XCP Migration 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 to determine whether any features your environment requires are unsupported by FlexGroup volumes. To generate reports, use XCP or Data Dynamics Insight Analytix.

General best practice #2: 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 9.8 adds support for proactive resizing in a FlexGroup volume, which keeps a consistent buffer of available free space across all member volumes in a FlexGroup by growing or shrinking member volumes automatically. You can configure proactive resizing to keep the total FlexGroup volume size where it is, or you can enable volume autosize to work in conjunction with proactive resizing for an even more hands-off approach to managing capacity. Another example is ONTAP 9.12.1 and the introduction of FlexGroup rebalancing. FlexGroup volumes can have usage imbalances between the constituents. FlexGroup rebalancing can be used to correct those imbalances.

Running the latest release of ONTAP (P-releases included) is the simplest way to get the most out of your FlexGroup volumes. For more information on the newest FlexGroup features, see TR-4571.

General best practice #3: 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 (all NetApp AFF or all FAS), disk types, spindle counts, and RAID group sizes will, ideally, be identical. Variations across the FlexGroup volume are supported but not recommended.

General best practice #4: Automated deployment

FlexGroup volumes are designed with simplicity in mind. Although it is possible to customize FlexGroup volumes, usually it is best to let ONTAP decide how to create them. Using the latest ONTAP versions makes these operations even more seamless, as new features are designed and added to new releases to simplify storage management. 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.

General best practice #5: 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 when you deploy them across multiple hardware resources.

General best practice #6: 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 10TiB FlexGroup volume with eight member FlexVol volumes has 1.25TiB per member. Therefore, it is important to consider the member volume sizes in addition to the total FlexGroup size. Member volumes are held to the same 300TiB limits (larger constituent volumes introduced in ONTAP 9.12.1P2) as regular FlexVol volumes, but for best results, they should not be smaller than 100GiB. If the FlexGroup size is large, files are more likely to be placed evenly. Generally, it is better to create one large FlexGroup volume and qtrees for workload segmentation than create many smaller FlexGroup volumes.

General best practice #7: Understand 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 1TiB file might be considered large in a 16TiB FlexGroup volume with eight member volumes because it takes up 50% of the 2TiB member volume's capacity immediately. However, that same 1TiB file uses only 1% of the capacity of a 100TiB 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. ONTAP 9.8 and later reduces the need to think about member volumes' sizes with the new proactive resizing feature.

General best practice #8: 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 cannot push as much data as larger ones, so the FlexGroup volume might not reach its potential with 1GiB links. Therefore, NetApp recommends that you use 10GiB or greater interfaces on each node in the cluster, and balance mount points across each node to even out network connectivity distribution. If you decide to use Link Aggregation Control Protocol (LACP) for link aggregation, be sure to follow the guidance in TR-4571.

General best practice #9: Handling high-file-count environments

FlexGroup volumes were primarily designed for high-file-count, large-metadata workloads such as those in Electronic Design Automation (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. FlexGroup volumes have introduced new options to control what type of file IDs are handed out to clients (64-bit file IDs) for both NFSv3 and NFSv4.x.

These options enable 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, you must remount existing NFS mounts 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.

General best practice #10: Initial member volume counts

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 earlier:

- Each volume has one affinity.
- Each aggregate has four affinities
- Nodes have a maximum of eight affinities

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

- Each volume has one affinity.
- Each aggregate has eight affinities
- Nodes have a maximum of 16 affinities.

One member volume per available affinity is 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 you create a volume by using System Manager or the `-auto-provision-as` option with the `volume create` command. ONTAP System Manager 9.7 and later uses the volume affinity count to create new FlexGroup volumes. To make the process more streamlined, System Manager 9.7 also tries to create fewer member volumes when smaller FlexGroup volumes are created.

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.

General best practice #11: 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 is 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.

General best practice #12: Increasing the size or maximum number of files of a FlexGroup volume

An individual FlexVol volume is limited to 300TiB and 2 billion files, a FlexGroup volume has virtually no limitations on how large it can grow outside of the stated 60PiB/400 billion files/200 member volumes tested configuration. In some situations, you might need to increase the capacity or available files.

Keep these considerations in mind when growing a FlexGroup volume's capacity or maximum number of files:

- Grow the existing member volumes (`volume resize`) before adding new member volumes (`volume expand`), if possible. Growing member volumes preserves the data layout, which keeps more predictable data ingest.

- When adding new nodes to a cluster, use the `volume move` command to balance members across the new nodes and then either grow existing volumes or add new member volumes in even multiples across all nodes. For example, if you add two nodes to a cluster and you have 16 member volumes in your FlexGroup, use the `volume move` command to relocate eight of those member volumes (four per node) to the new nodes (giving you four member volumes on four nodes), and then add 16 new member volumes across the four nodes.
- To prevent capacity or performance bottlenecks, when adding new member volumes to a FlexGroup volume, add them in multiples of four or eight per node (depending on the existing FlexGroup member volume count). Adding new member volumes in smaller batches (such as one or two at a time) might create ingest hot spots, similar to adding new disks to RAID groups. Prior to ONTAP 9.12.1 there was no ability to rebalance data between constituents, therefore adding multiple members helped to offset this problem by providing more available paths for parallel processing of new data requests. For more information see Rebalance FlexGroup volumes.

Note: Keep in mind FlexGroup member volume count limits for features such as NetApp SnapMirror® (32 per FlexGroup volume, 100 per cluster) in ONTAP 9.4 and earlier. ONTAP 9.5 and later removes those limits. See TR-4678: FlexGroup Volumes Data Protection Best Practices for details.

General best practice #13: Adding nodes or removing nodes from a cluster

When adding new nodes to a cluster, use the `volume move` command 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 volume. For example, suppose that you have a two-node NetApp AFF A900 all-flash storage system that has a 32 member FlexGroup volume (16 members per node). If you are adding two new A900 nodes, nondisruptively move 16 (eight from each node) of the volumes to the new HA pair to have eight member volumes per node in the four-node cluster. Then, if desired, use the `volume expand` command 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 (eight per node in the four-node A900 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.

General best practice #14: Storage efficiencies

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 do not see any space savings for postprocess deduplication unless you are using aggregate-level deduplication (inline in ONTAP 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, you might need to run undo commands to allow the changes to complete. One exception to managing storage efficiencies at the FlexGroup level is that you must run `volume efficiency undo` commands against the member volumes at advanced privilege. See TR-4571 for more information.

General best practice #15: Migrating to FlexGroup volumes

In releases earlier than 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 the XCP tool for NFSv3 migrations because of its superior speed and robust functionality. Using SnapMirror to migrate from a FlexVol volume to a FlexGroup volume 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 do not 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 (initial release)
Version 1.1	June 2019	ONTAP 9.6
Version 1.2	January 2020	ONTAP 9.7
Version 1.3	January 2021	ONTAP 9.8
Version 1.4	December 2023	ONTAP 9.14.1, limits updates, FlexGroup rebalance

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

TR-4571-a-0121

NetApp