Technical Report

Storage

Cloud Manager 3.5

Aksel Davis, Ben Cammett
03/12/2020

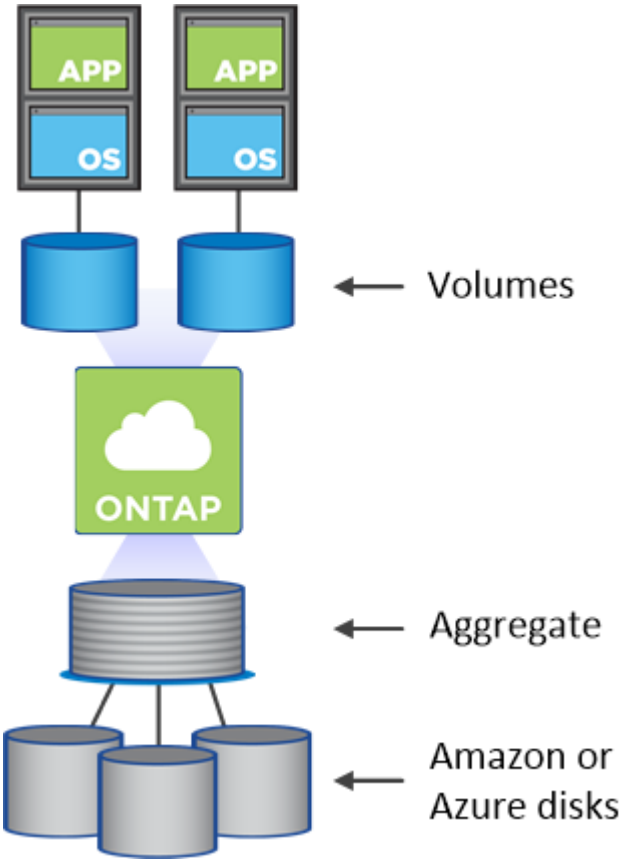
Table of Contents

Storage

Understanding how Cloud Volumes ONTAP uses cloud storage can help you understand your storage costs.

How Cloud Volumes ONTAP uses cloud storage

Cloud Volumes ONTAP uses AWS and Azure volumes as back-end storage. It sees these volumes as disks and groups them into one or more aggregates. Aggregates provide storage to one or more volumes.



In AWS, an aggregate can contain up to 6 disks, with a maximum disk size of 16 TB. The underlying EBS disk type can be either General Purpose SSD, Provisioned IOPS SSD, Throughput Optimized HDD, or Cold HDD. You can also pair an EBS disk with Amazon S3 for data tiering (not supported with Cold HDDs).

In Azure, an aggregate can contain up to 12 disks, with a maximum disk size of 4 TB. The underlying Azure managed disk type can be either Standard Storage (HDD) or Premium Storage (SSD). You can also pair a managed disk with Azure Blob storage for data tiering.

You choose the disk type when creating volumes and the default disk size when you deploy Cloud Volumes ONTAP. For more details, refer to the following:

- Choosing an AWS disk type
- Choosing an Azure disk type

- Choosing a disk size

The total amount of storage purchased from AWS or Azure is the *raw capacity*. The *usable capacity* is less because approximately 12 to 14 percent is overhead that is reserved for Cloud Volumes ONTAP use. For example, if Cloud Manager creates a 500 GB aggregate, the usable capacity is 442.94 GB.

Data tiering overview

You can reduce your storage costs by enabling automated tiering of cold data to low-cost object storage. Active data remains in high-performance SSDs or HDDs (the performance tier), while inactive data is tiered to low-cost object storage (the capacity tier). This enables you to reclaim space on your primary storage and shrink secondary storage.

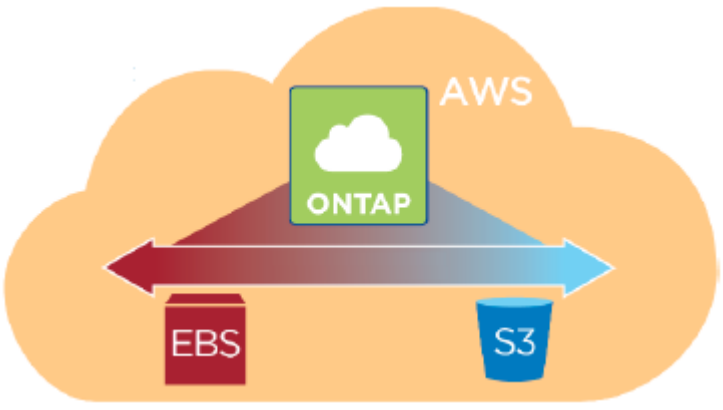
Cloud Volumes ONTAP supports data tiering in AWS and in Microsoft Azure. Data tiering is powered by FabricPool technology.

You do not need to install a feature license to enable data tiering.

- How data tiering works in AWS
- How data tiering works in Microsoft Azure
- How data tiering affects capacity limits
- Volume tiering policies
- Setting up data tiering

How data tiering works in AWS

When you enable data tiering in AWS, Cloud Volumes ONTAP uses EBS as a performance tier for hot data and AWS S3 as a capacity tier for cold data:



Performance tier in AWS

The performance tier can be General Purpose SSDs, Provisioned IOPS SSDs, or Throughput Optimized HDDs.

Capacity tier in AWS

By default, Cloud Volumes ONTAP tiers cold data to the S3 *Standard* storage class. Standard is ideal for frequently accessed data stored across multiple Availability Zones.

If you do not plan to access the cold data, you can reduce your storage costs by changing a system's tiering level to either of the following, after you deploy Cloud Volumes ONTAP:

Standard-Infrequent Access

For infrequently accessed data stored across multiple Availability Zones.

One Zone-Infrequent Access

For infrequently accessed data stored in a single Availability Zone.

The access costs are higher if you do access the data, so you must take that into consideration before you change the tiering level. For more details about S3 storage classes, refer to AWS documentation.

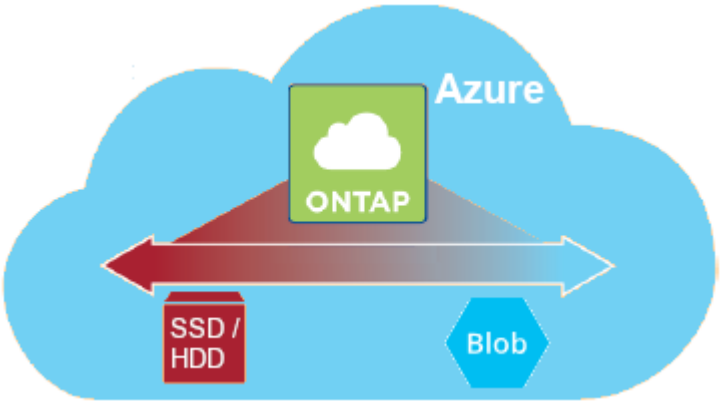
When you change the tiering level, cold data starts in the Standard storage class and moves to the storage class that you selected, if the data is not accessed after 30 days. For details about changing the tiering level, see Tiering cold data to low-cost object storage.

The tiering level is system wide—it is not per volume.

A Cloud Volumes ONTAP working environment uses an S3 bucket for all tiered data from the system. A different S3 bucket is not used for each volume. This includes an HA working environment. Cloud Manager creates an S3 bucket and names it fabric-pool-*cluster unique identifier*.

How data tiering works in Microsoft Azure

When you enable data tiering in Azure, Cloud Volumes ONTAP uses Azure managed disks as a performance tier for hot data and Azure Blob storage as a capacity tier for cold data:



Performance tier in Azure

The performance tier can be either Premium Storage (SSD) or Standard Storage (HDD).

Capacity tier in Azure

By default, Cloud Volumes ONTAP tiers cold data to the Azure *hot* storage tier, which is ideal for frequently accessed data.

If you do not plan to access the cold data, you can reduce your storage costs by changing a system's tiering level to the Azure *cool* storage tier after you deploy Cloud Volumes ONTAP. The cool tier is ideal for infrequently accessed data that will reside in the tier for at least 30 days.

The access costs are higher if you do access the data, so you must take that into consideration before you change the tiering level. For more details about Azure Blob storage tiers, refer to Azure documentation.

When you change the tiering level, cold data starts in the hot storage tier and moves to the cool storage tier, if the data is not accessed after 30 days. For details about changing the tiering level, see Tiering cold data to low-cost object storage.

The tiering level is system wide—it is not per volume.

A Cloud Volumes ONTAP working environment uses an Azure Blob container for all tiered data from the system. A different container is not used for each volume. Cloud Manager creates a new storage account with a container for each Cloud Volumes ONTAP system. The name of the storage account is random.

How data tiering affects capacity limits

If you enable data tiering, a system's capacity limit stays the same. The limit is spread across the performance tier and the capacity tier.

Volume tiering policies

To enable data tiering, you must select a volume tiering policy when you create, modify, or replicate a volume. You can select a different policy for each volume.

Some tiering policies have an associated minimum cooling period, which sets the time that user data in a volume must remain inactive for the data to be considered "cold" and moved to the capacity tier.

Cloud Volumes ONTAP supports the following tiering policies:

Snapshot Only

After an aggregate has reached 50% capacity, Cloud Volumes ONTAP tiers cold user data of Snapshot copies that are not associated with the active file system to the capacity tier. The cooling period is approximately 2 days.

If read, cold data blocks on the capacity tier become hot and are moved to the performance tier.

Auto

After an aggregate has reached 50% capacity, Cloud Volumes ONTAP tiers cold data blocks in a volume to a capacity tier. The cold data includes not just Snapshot copies but also cold user data

from the active file system. The cooling period is approximately 31 days.

This policy is supported starting with Cloud Volumes ONTAP 9.4.

If read by random reads, the cold data blocks in the capacity tier become hot and move to the performance tier. If read by sequential reads, such as those associated with index and antivirus scans, the cold data blocks stay cold and do not move to the performance tier.

Backup

When you replicate a volume for disaster recovery or long-term retention, data for the destination volume starts in the capacity tier. If you activate the destination volume, the data gradually moves to the performance tier as it is read.

None

Keeps data of a volume in the performance tier, preventing it from being moved to the capacity tier.

Setting up data tiering

For instructions and a list of supported configurations, see Tiering cold data to low-cost object storage.