Technical FAQ

NetApp Cloud Tiering Service

Frequently Asked Questions

Oded Berman, NetApp

January 2021

Abstract

This technical FAQ is meant to assist NetApp® systems engineers, professional services, support, partners, and customers by providing answers to frequently asked questions about the use and implementation of the NetApp Cloud Tiering service.

TABLE OF CONTENTS

Cloud Tiering service

What are the benefits of using the NetApp Cloud Tiering service?

Answer: Cloud Tiering addresses the challenges that come with rapid data growth, providing you with benefits such as:

- Effortless data center extension to the cloud, providing up to 50x more space
- Storage optimization, yielding an average storage savings of 70%
- Reduced total cost of ownership by 30%
- No need to refactor applications

Which NetApp ONTAP® versions does Cloud Tiering support?

Answer: Cloud Tiering supports ONTAP versions 9.2 and later.

What storage types does Cloud Tiering support?

Answer: Cloud Tiering supports the following storage types:

- **Performance Tier:** On-premises ONTAP running on NetApp AFF and SSD-backed FAS systems.
- **Cloud Tier:** Amazon S3, Azure Blob, Google Cloud Storage, and NetApp StorageGRID®. (Click the links for detailed storage classes and access tiers.)

Is Cloud Tiering only available for NetApp AFF systems?

Answer: No. Cloud Tiering is available for AFF and FAS systems with SSD aggregates. Also, Cloud Tiering supports displaying tiering information from Cloud Volumes ONTAP working environments.

What is the minimum ONTAP version that supports tiering from FAS systems with HDDs?

Answer: ONTAP 9.8.

What is the Connector or Service Connector?

Answer: To use the Cloud Tiering service, a Connector must be deployed. The Connector is a software running on a compute instance either within your cloud account or on-premises that enables Cloud Manager to securely manage cloud resources.

Can I deploy the Service Connector on-premises?

Answer: Yes. The Connector software can be downloaded and manually installed on a Linux host in your network. More information on installing the Connector can be found here.

What are the implications if the Connector fails?

Answer: In the case of a Connector failure, only the visibility into the tiered environments is affected. All the data is accessible and newly identified cold data is automatically tiered to object storage.

Can I use Cloud Tiering for both NAS and SAN data?

Answer: Yes. However, it is recommended to tier SAN data locally, to a StorageGRID system for example, because SAN protocols are more sensitive to connectivity issues than NAS.

Is an account with a cloud service provider required before using Cloud Tiering?

Answer: Yes. An account is required when defining the object storage to use. An account with a cloud storage provider is also required when setting up the Connector in the VPC/VNet.

Will Cloud Tiering retain my storage efficiency savings in the cloud tier?

Answer: Partially. Volume-level storage efficiencies, such as deduplication and compression, are preserved while aggregate-level efficiencies, such as cross-volume deduplication and compaction, are not.

Is Cloud Tiering used at the volume or aggregate level?

Answer: Cloud Tiering is enabled at the volume level by associating a tiering policy with that volume. Cold data identification is done at the block level.

How does Cloud Tiering differ from Cloud Volumes ONTAP data tiering?

Answer: The underlying tiering technology used for all ONTAP deployments is the NetApp FabricPool technology. The term **data tiering** is used within Cloud Manager to describe FabricPool on Cloud Volumes ONTAP, while **Cloud Tiering** is a solution that turns FabricPool into a managed service providing easier activation with advanced management capabilities. Currently, Cloud Tiering is used to activate and manage tiering (FabricPool) for on-premises ONTAP deployments and to display tiering-related information, in read-only format, for Cloud Volumes ONTAP. Activating data tiering with Cloud Volumes ONTAP is done directly through the working environment on the Cloud Manager **Canvas**.

What is the difference between FabricPool and Cloud Tiering?

Answer: FabricPool is ONTAP's tiering technology that can be self-managed through CLI and System Manager or managed as-a-Service through Cloud Tiering. Cloud Tiering turns FabricPool into a managed service with advanced automation processes, on both ONTAP and in the cloud, providing greater visibility and control over tiering across hybrid and multicloud deployments.

What kind of data is useful to tier to the cloud?

Answer: Essentially, any data that is considered inactive on both primary and secondary storage systems is a good target to move to the cloud. On primary systems, such data can include snapshots, historical records, and finished projects. On secondary systems, this includes all volumes that contain copies of primary data made for disaster recovery and backup purposes.

Can the data tiered to the cloud be used for disaster recovery or for backup/archive?

Answer: No. Because the volume's metadata is never tiered from the performance tier, the data stored in object storage cannot be accessed directly. However, Cloud Tiering can be used to achieve cost-effective backup and disaster recovery by enabling it on secondary systems and SnapMirror destination volumes (DP volumes), to tier off all of their data (metadata excluded), thus reducing their DC footprint and TCO.

Can I tier data from an AFF joined to a cluster that has FAS nodes with HDDs?

Answer: Yes. Cloud Tiering can be configured to tier volumes hosted on any SSD aggregate. The data tiering configuration is irrelevant to the type of controller used and whether the cluster is heterogeneous or not.

Can I tier data from FAS systems with HDDs only?

Answer: Starting with ONTAP 9.8, you can tier data from volumes hosted on HDD aggregates. Currently, configuring tiering in this case can only be done through the ONTAP CLI.

How does Cloud Tiering determine which blocks to tier to the cloud?

Answer: The tiering policy associated with the volume is the mechanism that controls which blocks are tiered and when. The policy defines the type of data blocks (snapshots, user, or both) and the cooling period. For more details, see the section "Tiering Policies."

How does Cloud Tiering affect the volume capacity?

Answer: Cloud Tiering has no effect on the volume's capacity but rather on the aggregate's performance tier usage.

What is the definition of inactive data or infrequently used data and how is that controlled?

Answer: The definition of what can also be referred to cold data is: volume blocks (metadata excluded) that have not been accessed for some amount of time. The "amount of time" is determined by a tiering policy attribute named cooling-days.

Licenses and costs

How much does using Cloud Tiering cost?

Answer: When tiering cold data to the public cloud:

- For the Pay-as-you-Go, usage-based subscription: **$0.05 per GB/Month**.
- For Annual, term-based subscription: starting from **$0.033 per GB/Month**.

  When tiering cold data to a NetApp StorageGRID (private cloud) there is **no cost**.

Can I have both a BYOL and PayGo license for the same ONTAP cluster?

Answer: Yes. Cloud Tiering allows using BYOL, PayGo, and a combination of both.

Can I use my FabricPool license (term-based or perpetual) with Cloud Tiering?

Answer: Yes. Both term-based and perpetual licenses are supported and can be activated through Cloud Tiering's licensing tab. In this case, Cloud Tiering registers the license and installs it on your ONTAP cluster.

What happens if I have reached the BYOL capacity limit?

Answer: If you reach a BYOL capacity limit, tiering of new cold data stops while all previously tiered data remains accessible. If you have a marketplace subscription to the Cloud Manager- Deploy & Manage Cloud Data Service, new cold data will continue to be tiered to the object storage and the charges would incur on a per-use basis.

Does the Cloud Tiering license include the egress charges from the cloud provider?

Answer: No, it does not.

Is rehydration of on-prem systems subject to the egress cost charged by the cloud providers?

Answer: Yes. All reads from the public cloud are subject to egress fees.

How can I estimate my cloud charges? Is there a "what if" mode for Cloud Tiering?

Answer: The best way to estimate how much a cloud provider will charge for hosting your data is to use their calculators: AWS, Azure and Google Cloud.

Are there any extra charges, by the cloud providers, for reading/retrieving data from the object storage to the on-prem storage?

Answer: Yes. Check Cloud Storage Pricing, Block Blob Pricing and Amazon S3 Pricing for additional pricing incurred with data reading/retrieving.

How can I estimate my volumes’ savings and get a cold data report before I enable Cloud Tiering?

Answer: To get an estimation, simply add your ONTAP cluster to Cloud Manager and inspect it through the Clusters Dashboard, which is located in the Tiering tab. When Inactive Data Reporting (IDR) is disabled or has not yet been activated for a long enough period of time, Cloud Tiering uses an industry-constant of 70% to calculate the estimated savings. Once IDR data is available, Cloud Tiering updates the savings to accurate figures.

How long does it take Inactive Data Reporting (IDR) to show information from the moment I start running it?

Answer: IDR starts showing information after the cooling period it uses has passed. Until ONTAP 9.7, IDR had a non-adjustable cooling period of 31 days. Starting with ONTAP 9.8, the IDR cooling-period can be configured.

Tiering policies

What are the available tiering policies?

Answer: There are four tiering policies:

- **None**: Classifies all data as always hot, preventing it from being moved to object storage.
- **Cold Snapshots** (Snapshot-only): Only cold snapshot blocks are moved to object storage.
- **Cold User Data and Snapshots** (Auto): Both cold snapshot blocks and cold user data blocks are moved to object storage.
- **All User Data** (All): Classifies all data as cold and immediately moves the entire volume to object storage.

  For more information, go to Learn About Cloud Tiering.

At which point my data is considered cold?

Answer: Because data tiering is done at the block level, a data block is considered cold after it hasn’t been accessed for a certain period of time, which is defined by the tiering policy’s **minimum-cooling-days** attribute. The applicable range for the attribute is **2-63** days, or **2-183** days starting with ONTAP 9.8.

What is the default cooling period for data before it is tiered to the cloud tier?

Answer: The default cooling period for the Cold Snapshot policy is **2 days**, while the default cooling period for Cold User Data and Snapshots is **31 days**. The cooling-days parameter is not applicable to the **All** tiering policy.

Is all the tiered data retrieved from object storage when I do a full backup?

Answer: During a full backup, all the cold data is read. The retrieval of the data depends on the tiering policy used. When using the **All** and **Cold User Data and Snapshots** policies, the cold data is not written back to the performance tier. When using the **Cold Snapshots** policy, only in case of an old snapshot being used for the backup will its cold blocks be retrieved.

Can you choose a tiering size per volume?

Answer: No. However, you can choose which volumes are eligible for tiering, the type of data to be tiered, and its cooling period. This is done by associating a tiering policy with that volume.

Is the All User Data policy the only option for data protection volumes?

Answer: No. Data protection (DP) volumes can be associated with any of the three policies available. The type of policy used on the source and destination (DP) volumes determines the write location of the data.

Does resetting the tiering policy of a volume to None rehydrate the cold data or just prevent future cold blocks from being moved to the cloud?

Answer: No rehydration takes place when a tiering policy is reset, but it will prevent new cold blocks from being moved to the cloud tier.

After tiering data to the cloud, can I change the tiering policy?

Answer: Yes. The behavior after the change depends on the new associated policy.

What should I do if I want to ensure certain data is not moved to the cloud?

Answer: Do not associate a tiering policy with the volume containing that data.

Where is the metadata of the files stored?

Answer: The metadata of a volumes is always stored locally, on the performance tier, and never tiered to the cloud.

Networking and security

What tools can I use for monitoring and reporting in order to manage cold data stored in the cloud?

Answer: Other than Cloud Tiering, Active IQ Unified Manager and Active IQ can be used for monitoring and reporting.

What are the implications if the network link to the cloud provider fails?

Answer: In case of a network failure, the local tier remains online and hot data remains accessible. However, blocks that were already moved to the cloud tier will be inaccessible and applications will receive an error message when trying to access that data. After connectivity is restored, all data will be seamlessly accessible.

Is there a network bandwidth recommendation?

Answer: The underlying FabricPool tiering technology read latency depends on the connectivity to the cloud tier. Although tiering works on any bandwidth, it is recommended to place intercluster LIFs on 10Gbps ports to provide adequate performance. There are no recommendations or bandwidth limitations for the Connector.

Is there a latency when a user attempts to access tiered data?

Answer: Yes. Cloud tiers cannot provide the same latency as the local tier since latency depends on the connectivity. To estimate the latency and throughput of an object store, Cloud Tiering provides a **Cloud Performance Test** (based on the ONTAP object store profiler) that can be used after the object store is attached and before tiering is set up.

How is my data secured?

Answer: AES-256-GCM encryption is maintained on both the performance and cloud tiers. TLS 1.2 encryption is used to encrypt data over the wire as it moves between tiers, and to encrypt communication between the Connector and both the ONTAP cluster and the object store.

Do I need an Ethernet port installed and configured on my AFF?

Answer: Yes. An intercluster LIF must be configured on an Ethernet port, on each node within an HA pair that hosts volumes with data you plan to tier to the cloud. For more information, check the Get started section in the Cloud Tiering documentation.

Object storage

What are the storage classes supported with Amazon S3?

Answer: The supported storage classes are: Standard, Standard-IA, One Zone-IA, and Intelligent-Tiering.

Why are Amazon S3 Glacier and S3 Glacier Deep Archive not supported by Cloud Tiering?

Answer: The main reason Amazon S3 Glacier and S3 Glacier Deep Archive are not supported is that Cloud Tiering was designed as a high-performance tiering solution, so data must be continuously available and quickly accessible for retrieval. With S3 Glacier and S3 Glacier Deep Archive, data retrieval can last anywhere between a few minutes to 48 hours.

Can I use other S3-compatible object storage services, such as Wasabi, with Cloud Tiering?

Answer: Yes, however, configuring them through the service UI is not supported yet.

Can I tier cold data to other on-prem storage arrays?

Answer: Yes. Currently only NetApp StorageGRID arrays are supported with more to be supported in the future.

Does Cloud Tiering support the use of lifecycle management policies offered by the different cloud object stores?

Answer: When using Amazon S3, Cloud Tiering can apply a lifecycle rule that would move data from the standard class to any of the other supported classes after 30 days.

What happens with the tiered data when you migrate a volume from one cluster to another?

Answer: When migrating a volume from one cluster to another, all the cold data is read from the cloud tier. The write location on the destination cluster depends on whether tiering was enabled and the type of tiering policy used on the source and destination volumes.

What happens with the tiered data when you move a volume from one node to another in the same cluster?

Answer: If the destination aggregate does not have an attached cloud tier, data is read from the cloud tier of the source aggregate and written entirely to the local tier of the destination aggregate. If the destination aggregate has an attached cloud tier, data is read from the cloud tier of the source aggregate and first written to the local tier of the destination aggregate, to facilitate quick cutover. Later, based on the tiering policy used, it is written to the cloud tier. Starting with ONTAP 9.6, if the destination aggregate is using the same cloud tier as the source aggregate, the cold data does not move back to the local tier.

How can I bring my tiered data back on-prem?

Answer: Write back is generally performed on reads and depends on the tiering policy type. Writing back of the entire volume can be done with a **volume move** operation, prior to ONTAP 9.8. After 9.8, the

Promote cloud retrieval policy can be used (through the ONTAP CLI) to bring all the data back to the local tier.

When replacing an existing AFF/FAS controller with a new one, would the tiered data be migrated back on-prem?

Answer: No. During the “head swap” procedure the only thing that changes is the aggregate’s ownership. In this case it will be changed to the new controller without any data movement.

Can I use the tiered data to recover a volume or a system when having a disaster recovery scenario?

Answer: No. Because the volume’s metadata is always stored on the local tier, in case there’s a disaster and the local tier is lost, the metadata is lost as well and there is no way to reference the tiered data.

Can I use the cloud provider’s console or object storage explorers to look at the data tiered to a bucket? Can I use the data stored in the object storage directly without ONTAP?

Answer: No. The objects constructed and tiered to the cloud do not contain a single file but up to 1,024 4KB blocks from multiple files. A volume’s metadata always remains on the local tier.

Can multiple buckets be attached to the same aggregate?

Answer: It is possible to attach up to two buckets per aggregate for the purpose of mirroring, where cold data is synchronously tiered to both buckets. The buckets can be from different providers and different locations.

Can different buckets be attached to different aggregates in the same cluster?

Answer: Yes. The general best practice is to attach a single bucket to multiple aggregates. However, when using the public cloud there is a maximum IOPS limitation for the object storage services, therefore multiple buckets must be considered.

Can I pick any cloud provider access tier or storage class as the destination?

Answer: Yes, from the storage classes and access tiers supported by Cloud Tiering. Multiple storage classes or access tiers are supported on AWS, Azure and Google Cloud. When creating a new bucket, the tier/class can be selected.

Version history

Version	Date	Document Version History
1.0	January 2021	Initial release.

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

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-transferrable, 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, disclosed, reproduced, modified, performed, or displayed without the prior written approval of NetApp, Inc. United States Government license rights for the Department of Defense are limited to those rights identified in DFARS clause 252.227-7015(b).

Trademark Information