Intelligent Tiering for RGW

Jiffin Tony Thottan
Ceph RGW Team
IBM

Agenda

- Intelligent Tiering
- Current Support in RGW
- Cloud Transition
- Restore Capability
- S3 Tape
- Demo
- Q and A

IBM

Intelligent Tiering

What ? Movement of Data across different storage classes

Why ? optimizes storage costs and performance

When/How? Based on Policies

Whats present in RGW

- Placement targets : control which pool are associated with a particular bucket
- StorageClasses: specify the placement of object data
- Lifecycle Policies: give users a way to set how and when objects in S3 buckets move between tiers, expire, and are deleted
- Configuration are created in zone/zonegroup level
- Lifecycle Policies defines with S3 Lifecycle rules

Cloud Transition

- Transition to cloud-endpoints (S3 compatible)
    - Storage-class of type “cloud-s3”
    - Cloudtier configuration

```
tier-type = cloud-s3
tier-config =
        {
            "access_key": <access>,
            "secret": <secret>,
            "endpoint": <endpoint>,
            "region": <region>,
            ...
         }
```

- Uni-directional
- Once transitioned, objects can be retrieved from only the cloud endpoint
- **retain_head_object** - object stub (metadata) is preserved in RGW

Restore Capability

- retain_head_object should be enabled
- Objects can be restored via
  - the **S3 restore-object API** or
  - **GET**(/read-through) on the transitioned object.
- Restore-type
  - Temporary
    - Data restored will be temporary
    - Lifecycle cloud-transition rules will be skipped
    - Post expiry, the data will be deleted and the object will be reset to stub
  - Permanent
    - Once restored, it will be treated like any other regular object
    - Subjected to Lifecycle transitions
    - Replicated across the zones

S3 Apis for client

- Using S3 restore-object in the CLI, you can restore the cloud transitioned objects.

  Syntax: **aws s3api restore-object --bucket <value> --key <value> [--version-id <value>] --restore-request (structure)**

  For example,

  - Permanent: aws s3api restore-object --bucket my-glacier-bucket --key permanent.txt [--version-id 3sL4kqtJlcpXroDTDmJ+rmSpXd3dIbrHY+MTRCxf3vjVBH40Nr8X8gdRQBpUMLUo] --restore-request {}
  - Temporary: aws s3api restore-object --bucket my-glacier-bucket --key temporary.txt [--version-id 3sL4kqtJlcpXroDTDmJ+rmSpXd3dIbrHY+MTRCxf3vjVBH40Nr8X8gdRQBpUMLUo] --restore-request Days=5
- Using S3 cli for get objects, the read through happens, the object automatically download transitioned objects

  This enabled by tier config options :

  - ***"allow_read_through" : "enable | disable"***
  - ***"read_through_restore_days" : <integer>***

Verification/Status

- The retrieval of objects is done asynchronously.
- Once the restore request is initiated -
  - The object is internally marked RestoreAlreadyInProgress
  - An asynchronous request is sent to GET object from the cloud endpoint
  - Once the data is restored, the object is marked CloudRestored
  - If restore fails, the object is marked RestoreFailed
  - This state can be checked via radosgw-admin object stat command
- x-amz-restore header also contains restoration status and can be checked via HEAD request
- Example: aws s3api --endpoint https://host02.example.com:8043 --region default head-object --key transition1 --bucket transition
- Objects are restored to STANDARD storage-class. However, for temporary objects, *x-amz-storage-class* will still return the original cloudtier storage-class

S3 Tape

- S3 Compatibility: Store and manage data using the S3 API
- Tape Storage: Data is stored on tape, providing a cost-effective and durable storage solution
- Data Integrity: Data is verified for integrity using checksums and digital signatures
- Multi-Tenant Support: Support for multiple tenants, providing a scalable and secure solution for archival storage

Future Plans

- Providing "cloud-s3-glacier" storage class for Tape
- Support restore feature in "cloud-s3-glacier"
    - PR https://github.com/ceph/ceph/pull/61558
- Provide more options in radosgw-admin cli for debugging
- Automate the restoring workflow similar to transition
- TODO : ceph tracker link

Any questions?

Thank You