

Integrating IBM Storage Ceph with PoINT Archival Gateway for Policy-Based S3 Tape Archiving(2025/5/17)

有限会社オプティカルエキスパート 角野

IBM から PAG を Ceph Object Storage に統合するための、コマンドの実行内容を含む詳細な手順書が公開されました。

2022/10 以降は、IBM Storage Ceph と呼ばれている Ceph ですが、昨年 11 月にリリースされた V8.0 から、オブジェクトストレージの Ceph RGW(Reliable Autonomous Distributed Object Storage Gateway)は、外部のクラウドストレージをバックエンドの cloud-tier placements として使用し、そこにあるオブジェクトを直接、Ceph 上に復元する S3 のライフサイクル管理上での新しい機能が追加されました。さらに、先月中旬に、Ceph 開発担当から、PAG との統合に関する記事が IBM の Web 上で公開され、さらに PAG の開発元の PoINT 社の Web でも同じ内容が公開されました。内容としては、Ceph と PAG のそれぞれの機能紹介と、PAG の Linux 版のインストールと Ceph が、PAG の Object Repository(Bucket)を、Ceph のライフサイクル管理下のクラウド S3 として設定するための、Ceph と AWS S3 で実行するコマンドの具体的な実行例が含まれており、以下にまとめました。

Benefits of Integrating PAG with IBM Storage Ceph Object Storage

PoINT Archival Gateway(PAG)を使用すると、テープストレージクラスを Ceph クラスタに S3 のライフサイクル管理を使用して統合できます。テープをアクティブなアーカイブ階層とする多階層構成を実現できます。Ceph はポリシーベースのデータアーカイブと検索機能をサポートし、PAG を S3 テープエンドポイントとして統合することで、長期保存、災害復旧、またはコストを最適化したコールドストレージを実現します。ポリシーベースの自動化を活用することで、Ceph は事前に定義されたライフサイクルルールに従ってデータが PAG に、自動的にテープに移動します。PAG は、追加のディスクストレージクラスを必要としないため、Ceph との効率的なディスクとテープのストレージ統合を実現します。

PAG and IBM Storage Ceph Integration Workflow

1. Ceph 管理者がクラウド層(テープ)ストレージクラスを設定する

Ceph 管理者は(CLI または操作 API を使用して)、オブジェクトを PoINT Archival Gateway (PAG)にルーティングするストレージ階層を設定し、S3 テープエンドポイントを提供する。

2. エンドユーザーが Ceph バケットにオブジェクトを作成する

エンドユーザー(またはアプリケーション)は、標準の S3 API 経由でオブジェクト(JPG ファイルなど)を Ceph Object が管理する「ユーザーデータバケット」にアップロードする。新しく作成されたオブジェクトは、最初は構成済みの「ホット」(または標準)階層に存在する。

3. ライフサイクルポリシーによる移行

ライフサイクルポリシーはユーザーバケットに定義される。例えば、365 日後に、そのしきい値より古いオブジェクトは PAG を通じて低コストの長期ストレージティア(テープ)に移行するよう指定する。

4. テープへの自動アーカイブ

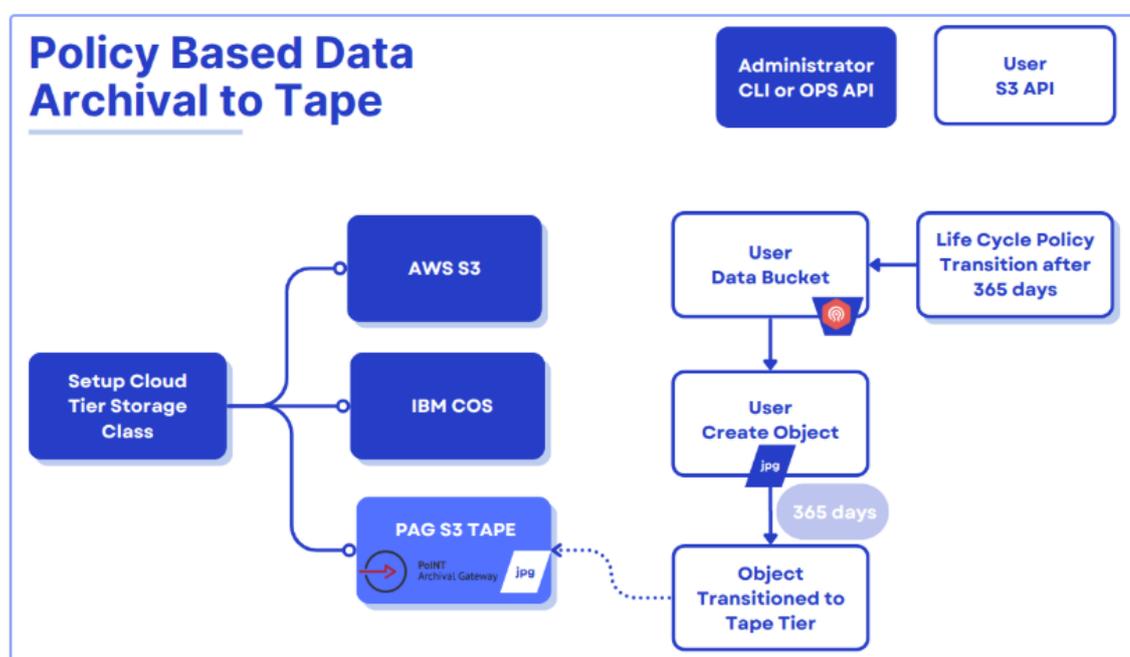
オブジェクトの年齢がポリシールールを満たすと(365 日後)、Ceph は自動的に PAG を通じてテープ階層にオブジェクトを移行する。

5. 長期保存と検索

アーカイブ後、オブジェクトは PAG 経由でテープメディアに保存される。オブジェクトは引き続き Ceph のメタデータで参照され、同じ S3 Ceph エンドポイント、PAG テープティアを使用して後で取得できる。

Ceph 管理者は (CLI または運用 API を使用して)、オブジェクトを PoINT Archival Gateway (PAG) にルーティングするストレージ階層を構成し、S3 Tape エンドポイントを提供出来ます。

以下の図は、左側が、Ceph オブジェクトストレージの外部クラウド S3 として、設定されている AWS S3/IBM COS と PAG を示しています。右側が、Ceph Object Storage で、S3 のライフサイクルポリシーで、365 日経過すると、PAG が管理しているテープに、オブジェクトが自動的に移行される様子を示しています。



IBM Storage Ceph 8.0 はポリシーベースのデータ検索を導入し、その機能が進化しました。この機能強化により、ユーザーは PAG のような S3 Tape エンドポイントからアーカイブされたオブジェクトをオンプレミスの Ceph 環境に直接取り出すことが出来るようになりました。データは、一時的か永久的に Ceph のバケットにオブジェクトとして戻すことが可能です。

Temporary restores: 復元されたデータは、ライフサイクルのクラウド移行ルールを回避し、指定された時間後に自動的に削除され、オブジェクトは以前のスタブ状態に戻る。

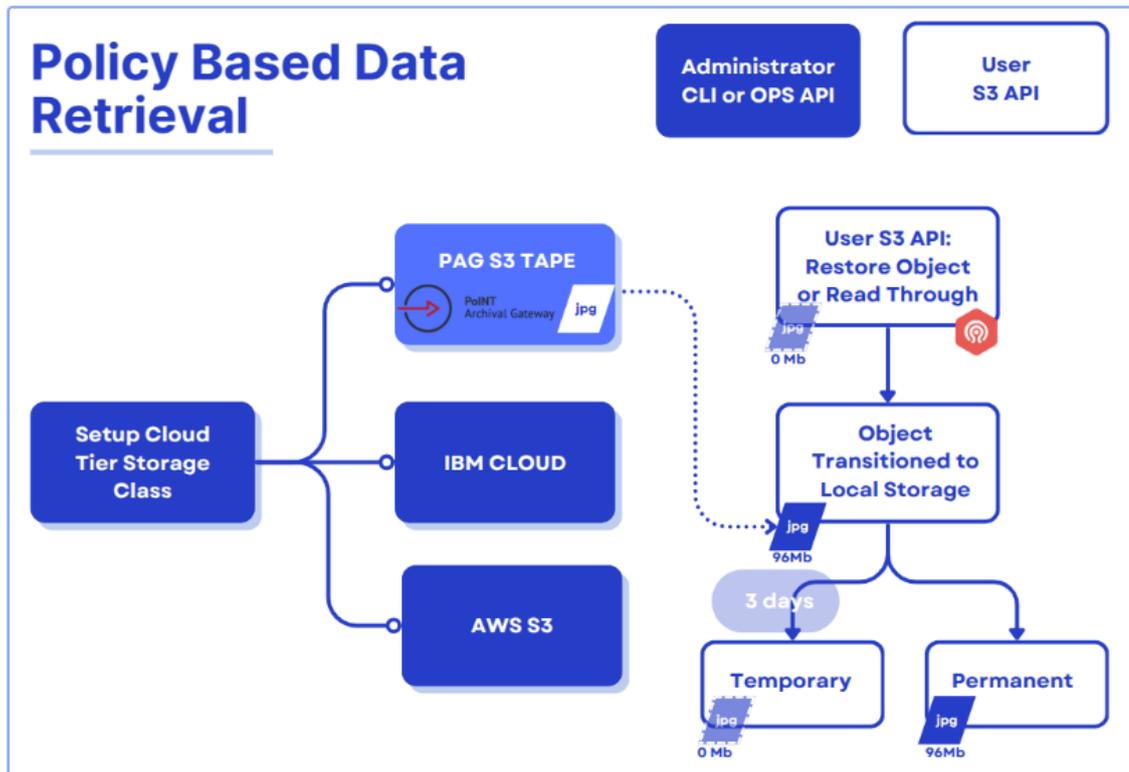
Permanent restores: これにより、オブジェクトは Ceph クラスタに完全に再統合され、通常のオブジェクトと同様に扱われ (そして通常のオブジェクトになり)、標準的なライフサイクルポリシーとレプリケーションプロセスに従います。

このオブジェクトの復元には、以下の二つの機能を選択して使用出来ます。

S3 RestoreObject API: S3RestoreObject API リクエストを使用して、リモート S3 エンドポイントからオブジェクトを取得出来ます。

Read-through Object Retrieval: この機能により、移行されたオブジェクトに対する標準的な S3 GET リクエストで、Ceph クラスタに透過的にリストア出来ます。

以下の図では、左側が外部クラウド S3 で、右側が、Ceph Object Storage でのオブジェクトの復元の様子で、復元されたオブジェクトを一時的に保存するか、永久的に保存するかを選択出来ることを示しています。



ここからは、実際のコマンド実行例(Hands-On)です。IBMのWebページでは、PAGのLinux版の簡単なインストール手順と設定が紹介されていますがここでは割愛します。。PAGの設定内容としては、Partition名が、Ceph_s3_partitionで、Object Repository名(Bucket)が、cephs3tapeに設定されています。

Integrating PAG as a Storage Class within Ceph

Ceph RGW(Reliable Autonomous Distributed Object Storage Gateway)は、外部のクラウドストレージをバックエンドの *cloud-tier placements* として使用出来ます。新しく、point-tape を storage class として追加します。これは、Ceph の 2024/11 にリリースされた V8.0 からの Transitioning Data to Amazon S3 Cloud Service 機能を使用します。以下の最初のコマンドで、point-tape の Storage Class をクラウド S3 として指定します。次に、PAG の S3 サーバーと target_path で Object Repository(Bucket)や Access Key 等を追加します。PAG の Object Repository が、target-path=cephs3tape です。

```
# radosgw-admin zonegroup placement add --rgw-zonegroup=default ¥
  --placement-id=default-placement --storage-class=point-tape ¥
  --tier-type=cloud-s3

# radosgw-admin zonegroup placement modify --rgw-zonegroup default ¥
```

```
--placement-id=default-placement --storage-class point-tape ¥
--tier config=endpoint=http://linux1.cephlabs.com:4080, ¥
access_key=9FD33A27642C45480260, ¥
secret="YvFLFqQD+fZF+2gwVD4hbbgYzNoo4QeUiprh0Tv", ¥
target_path=cephs3tape, ¥
multipart_sync_threshold=44432, ¥
multipart_min_part_size=44432, ¥
retain_head_object=true,region=default,allow_read_through=true
```

placement のリストを表示して、point-tape が表示されることを確認します。

```
# radosgw-admin zonegroup placement list
```

Note: もしも、Ceph がマルチサイトとして稼働していない場合には、変更を有効にするために、RGW を再起動します。

```
# ceph orch restart rgw.default
```

Bucket Creation & Lifecycle Policy

Bucket を、dataset 名で aws のコマンドで作成します。

```
# aws --profile tiering --endpoint https://s3.cephlabs.com ¥
s3 mb s3://dataset --region default
```

ライフサイクルポリシーを以下のように作成し、point-tape-lc.json として保存し、ライフサイクル設定を aws のコマンドを使用して追加します。

```
{
  "Rules": [
    {
      "ID": "Testing LC. move to tape after 1 day",
      "Prefix": "",
      "Status": "Enabled",
      "Transitions": [
        { "Days": 1, "StorageClass": "point-tape" }
      ]
    }
  ]
}
```

```
# aws --profile tiering s3api put-bucket-lifecycle-configuration ¥
--lifecycle-configuration file://point-tape-lc.json --bucket dataset
# aws --profile tiering s3api get-bucket-lifecycle-configuration --bucket dataset
{
  "Rules": [
    {
      "ID": "move to tape after 1 day",
```

```
    "Prefix": "",
    "Status": "Enabled",
    "Transitions": [
      { "Days": 1, "StorageClass": "point-tape" }
    ]
  }
]
```

Testing the Integrated Setup

サンプルファイルを upload して、リストを表示し、内容を確認します。

```
# aws --profile tiering s3 cp 10mb_file s3://dataset/
upload: ./10mb_file to s3://dataset/10mb_file

# aws --profile tiering s3api list-objects-v2 --bucket dataset
{
  "Contents": [
    {
      "Key": "10mb_file",
      "LastModified": "2025-03-24T15:40:55.879Z",
      "ETag": "¥75821af1e9df6bbc5e8816f5b2065899-2¥",
      "Size": 10000000,
      "StorageClass": "STANDARD"
    }
  ]
}
```

ライフサイクルデーモンが起動すると、以下のように StorageClass が変更され、オブジェクトがテープに移動されます。

```
# radosgw-admin lc list
# aws --profile tiering s3api list-objects-v2 --bucket dataset
{
  "Contents": [
    {
      "Key": "10mb_file",
      "LastModified": "2025-03-24T15:43:02.891Z",
      "ETag": "¥75821af1e9df6bbc5e8816f5b2065899-2¥",
      "Size": 0,
      "StorageClass": "point-tape"
    }
  ]
}
```

バックエンドの PAG のオブジェクトを確認します。:

```
# aws --profile points3 --endpoint http://linux1.cephlabs.com:4080 ¥
```

```
s3api head-object --bucket ceph3tape --key dataset/10mb_file
```

Object Retrieval Workflow

Temporary restore (3 days): 一時的にオブジェクトをリストアします。

```
# aws --profile tiering s3api restore-object ¥
  --bucket dataset --key hosts --restore-request Days=3
# aws --profile tiering s3 ls s3://dataset
2025-03-24 11:43:02  10000000 10mb_file
# aws --profile tiering s3api head-object ¥
  --bucket dataset --key 10mb_file
{
  "AcceptRanges": "bytes",
  "Restore": "ongoing-request=¥"false¥", expiry-date=¥"Thu, 27 Mar 2025 15:45:25 GMT¥",
  "LastModified": "2025-03-24T15:43:02Z",
  "ContentLength": 10000000,

  "StorageClass": "point-tape"
}
```

Permanent restore: オブジェクトを永久的に復元します。

```
# aws --profile tiering s3 cp 20mb_file s3://dataset/
upload: ./20mb_file to s3://dataset/20mb_file
# aws --profile tiering s3api head-object ¥
  --bucket dataset --key 20mb_file | grep StorageClass
"StorageClass": "point-tape"
# aws --profile tiering s3api restore-object ¥
  --bucket dataset --key 20mb_file --restore-request {}
# aws --profile tiering s3api head-object --bucket dataset --key 20mb_file
{
  "AcceptRanges": "bytes",
  "LastModified": "2025-03-24T15:55:10Z",
  "ContentLength": 20000000,
  "StorageClass": "STANDARD"
}
```