

Using gsutil to Perform Operations on Buckets and Objects

Overview

In this lab you will use `gsutil` to create a bucket and perform operations on objects. `gsutil` is a Python application that lets you access Cloud Storage from the command line. The `gsutil` tool has commands such as `mb` and `cp` to perform operations. Each command has a set of options that are used to customize settings further.

What you'll learn to do

- Create a bucket
- Copy files from a local folder to a bucket
- Synchronize the contents of the local folder with the contents of the bucket
- Change access control permissions on objects
- Delete a bucket.

In Cloud Shell session execute the following command to download sample data for this lab from a git repository:

```
git clone https://github.com/GoogleCloudPlatform/training-data-analyst
```

Change to the **blogs** directory:

```
cd training-data-analyst/blogs
```

Working with buckets and objects

First, set some environment variables: `PROJECT_ID=$(gcloud config get-value project)`

```
BUCKET=${PROJECT_ID}-bucket
```

Create a bucket

Create a bucket and multi-regional storage class: `gsutil mb -c multi_regional`

```
gs://${BUCKET}
```

Upload objects to your bucket

Run the following to copy the `endpointslambda` object to your bucket:

```
gsutil -m cp -r endpointslambda gs://${BUCKET}
```

List objects

To list objects in your bucket, execute the following command:

```
gsutil ls gs://${BUCKET}/*
```

Sync changes with bucket

Use the following commands to rename and delete some files:

```
mv endpointslambda/Apache2_0License.txt endpointslambda/old.txt  
rm endpointslambda/aeflex-endpoints/app.yaml
```

Now synchronize the local changes with the bucket:

```
gsutil -m rsync -d -r endpointslambda gs://${BUCKET}/endpointslambda
```

In this command, the `-d` option deletes files from the target if they're missing in the source (in this case, it deletes **app.yaml** from the bucket). The `-r` option runs the command recursively on directories.

To verify that the bucket is now in sync with your local changes, list the files in the bucket again: `gsutil ls gs://${BUCKET}/*`

Make objects public

To allow public access to all files under the `endpointslambda` folder in your bucket, execute the following command:

```
gsutil -m acl set -R -a public-read gs://{BUCKET}
```

To confirm files are viewable by the public, open the following link in a new incognito or private browser window, replacing <your-bucket-name> with the full name of your bucket, not the environment variable:

```
http://storage.googleapis.com/<your-bucket-name>/endpointslambda/old.txt
```

This URL uses the Cloud Storage API link to view the object without authentication. For more information, see [Accessing Public Data](#).

Copy with different storage class

Next, copy a file with [Nearline storage class](#) instead of the bucket's default Multi-regional storage class:

```
gsutil cp -s nearline ghn/ghn_on_bq.ipynb gs://{BUCKET}
```

Check storage classes

Run the following to check the storage classes and view other detailed information about the objects in your bucket: `gsutil ls -Lr gs://{BUCKET} | more`

Press the space key to continue viewing the rest of the command's output. The output shows that the **ghn_on_bq.ipynb** object has NEARLINE storage class while the other objects have MULTI_REGIONAL storage class.

Delete your bucket

Before deleting a bucket, you must first delete all objects in the bucket. To delete all objects, execute the following command: `gsutil rm -rf gs://{BUCKET}/*`

Now delete the bucket: `gsutil rb gs://{BUCKET}`