GCS to GCS Storage Transfer Strategy using STS & Event Triggers

# **1. Project Setup**

* Two GCP projects: GDW (Project 1) and APMF (Project 2).
* Both under the same GCP organization.
* Goal: Cross-project GCS to GCS transfers using Storage Transfer Service (STS).

# **2. Transfer Strategy Overview**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Direction | Trigger Project | STS Job Project | Type | Status |
| GDW ➝ APMF (Push) | GDW | GDW | Event-driven | ✅ Working |
| APMF ➝ GDW (Pull) | APMF | GDW | Event-driven | ✅ Working |
| GDW ➝ APMF (Batch) | GDW | GDW | One-time batch | ✅ Done |
| APMF ➝ GDW (Batch) | APMF | GDW | One-time batch | ✅ Done |

# **3. Event Flow Architecture**

* GCS Bucket triggers a Pub/Sub topic on Object Finalize.
* Cloud Function is triggered by the topic.
* Cloud Function extracts bucket + file info and triggers Composer DAG.
* Composer DAG runs pre-created STS job.

# **4. Cloud Function**

* Language: Python 3.9
* Uses functions\_framework and triggered via Pub/Sub.
* Entrypoint set to: main (not main.py).
* Triggers Composer DAG using REST API with auth token.
* Passes file info (`bucket`, `name`) as config to DAG.

# **5. Composer DAGs**

* Uses `CloudDataTransferServiceRunJobOperator`.
* Pre-created STS jobs are triggered via `job\_name`.
* DAGs are parameterized using config passed by Cloud Function.

# **6. IAM Roles & Terraform**

* No Admin roles used; tightly scoped service accounts.
* Terraform used to assign roles via modules/iam/locals.tf.
* Roles like composer.worker, storage.objectViewer, cloudfunctions.invoker used.
* Cross-project access enabled by IAM for service account triggering Composer in another project.

# **7. Cloud Function Deployment via Terraform**

* Code packaged using `archive\_file` and deployed to GCS.
* Terraform module: wf-cloudfunctions-factory used.
* Each Cloud Function is zipped from `src2/`, `src3/` etc.
* Runtime: python39, entry\_point: main

# **8. Debugging and Fixes**

* 401 Unauthorized resolved by granting composer.worker role to CF SA in target project.
* Function execution and DAG triggers verified via logs.
* Pub/Sub was confirmed to be delivering messages correctly.

# **9. File & Module Structure**

|  |  |
| --- | --- |
| Component | Location |
| DAGs | corp-apmf/composer/, corp-gdw/composer/ |
| Cloud Function | modules/cloudfunction/src3/main.py |
| IAM Roles | modules/iam/locals.tf |
| Pub/Sub | modules/pubsub/ |
| GCS Trigger | modules/gcs/ |