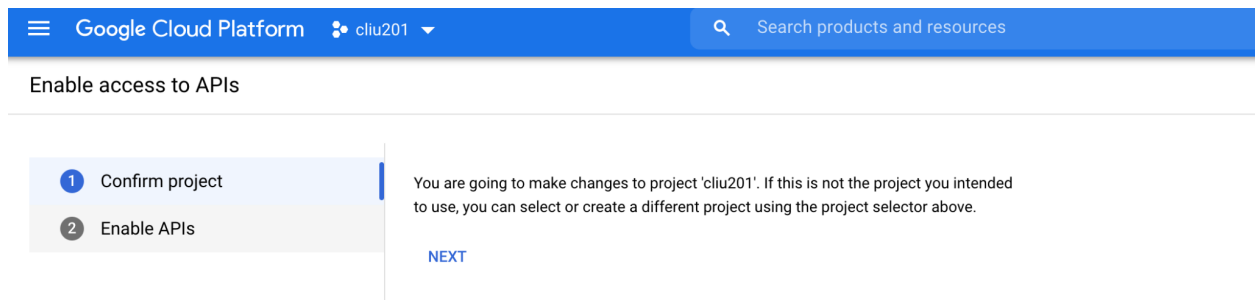


1.Full batch import bq into spanner

1.1.enable api

https://console.cloud.google.com/flows/enableapi?apiid=spanner.googleapis.com,storage_component,compute,dataflow&_ga=2.147151671.1836176879.1627838639-2032371454.1625546262



1.

1.create a schema for your database

Create a schema for your database

The Cloud Console provides two ways to create, alter, and delete tables and indexes in a database:

- By using the default Database editor to specify each part of tables, columns, and indexes.
- By switching the editor to **Edit as text** mode and entering SQL statements in Cloud Spanner Data Definition Language (DDL) syntax.

This quickstart uses DDL.

1. In the left pane of the Cloud Console, click the **example-db** database.
2. Click **Create table**.
3. Click the **Edit as text** toggle.
4. In the **DDL statement** field, enter:

```
CREATE TABLE MobileLog (  
  id STRING(1024) NOT NULL,  
  price STRING(1024),  
  product STRING(1024)  
) PRIMARY KEY (id);
```

1.2.export bq tables into avro file on GCS

Explorer + ADD DATA

mobilelog

SCHEMA DETAILS PREVIEW

Row Id price product

1	1	\$1.99	shirt
2	1	\$1.99	shirt
3	1	\$1.99	shirt
4	1	\$1.99	shirt
5	202006091557	\$1.99	shirt
6	202006091600	\$1.99	shirt
7	202006091600	\$1.99	shirt
8	202006101600	\$1.99	shirt
9	202006101600	\$1.99	shirt
10	202006101600	\$1.99	shirt
11	202006101600	\$1.99	shirt
12	202006101600	\$1.99	shirt
13	202006111600	\$1.99	shirt
14	202006141124	\$1.99	shirt
15	3	\$3.99	shirt
16	3	\$3.99	shirt
17	3	\$3.99	shirt

Explore with Sheets

Explore with Data Studio

Export to GCS

Scan with DLP

Specify the folder and the

1.3. Creating a spanner-export.json file

Notes: it's relative path to spanner-export.json, not absolute path

```
{
  "tables": [
    {
      "name": "MobileLog",
      "dataFiles": [
        "mobilelogtospanner"
      ]
    }
  ]
}
```

1.4. Create bq to spanner dataflow job

Running the Cloud Storage Avro to Cloud Spanner template

[Console](#) [gcloud](#) [API](#)

- Go to the Dataflow **Create job from template** page.

[Go to Create job from template](#)
- In the **Job name** field, enter a unique job name.

For the job to show up in the Spanner [Instances](#) page of the Cloud Console, the job name must match the following format:

```
cloud-spanner-import-SPANNER_INSTANCE_ID - SPANNER_DATABASE_NAME
```

Replace the following:

 - SPANNER_INSTANCE_ID: your Spanner instance's ID
 - SPANNER_DATABASE_NAME: your Spanner database's name
- Optional: For **Regional endpoint**, select a value from the drop-down menu. The default regional endpoint is `us-central1`.

For a list of regions where you can run a Dataflow job, see [Dataflow locations](#).
- From the **Dataflow template** drop-down menu, select the **Avro Files on Cloud Storage to Cloud Spanner** template.
- In the provided parameter fields, enter your parameter values.
- Click **Run job**.

Google Cloud Platform

cliu201

Search products and resources

Dataflow

Jobs

Snapshots

Notebooks

Pipelines

SQL Workspace

Jobs

+

CREATE JOB FROM TEMPLATE

+

CREATE JOB FROM SQL


Dataflow provides unified streaming and batch data processing that's serverless, fast, and cost-effective. [Learn more](#)

Running

Filter

Filter jobs

Name	Type	End time	Elapsed time	Start time	Status	SDK version	ID	Region
------	------	----------	--------------	------------	--------	-------------	----	--------



Job name *

bq-to-spanner-2

Must be unique among running jobs

Regional endpoint *

us-central1



Choose a Dataflow regional endpoint to deploy worker instances and store job metadata. You can optionally deploy worker instances to any available Google Cloud region or zone by using the worker region or worker zone parameters. Job metadata is always stored in the Dataflow regional endpoint. [Learn more](#)

Dataflow template *

Avro Files on Cloud Storage to Cloud Spanner



A pipeline to import a Cloud Spanner database from a set of Avro files in GCS.

[OPEN TUTORIAL](#)

Required parameters

Cloud Spanner instance id *

cliu201-spanner

The instance id of the Cloud Spanner database that you want to import to.

Cloud Spanner database id *

example-db

The database id of the Cloud Spanner database that you want to import into (must already exist).

Cloud storage input directory *

gs://cliu201-bucket/spanner/

The GCS path where the Avro files should be imported from.

Temporary location *

gs://cliu201-bucket/temp/

Path and filename prefix for writing temporary files. Ex: gs://your-bucket/temp

2.Incremental import bq into spanner

2.1 bq export by “filter”

```
EXPORT DATA OPTIONS(  
  uri='gs://cliu201-bucket/spanner/0803*',  
  overwrite=true,  
  format='AVRO') AS  
SELECT * FROM `cliu201.ds201.mobilelog` where id = '0803'
```

← Bucket details

cliu201-bucket

OBJECTS

CONFIGURATION

PERMISSIONS

RETENTION

L

Buckets > cliu201-bucket > spanner 

UPLOAD FILES




UPLOAD FOLDER

CREATE FOLDER

MANAGE HOLDS

Filter by name prefix only ▼

 Filter Filter objects and folders

<input type="checkbox"/>	Name	Size	Type
<input type="checkbox"/>	 0803000000000000	255 B	application/octet-
<input type="checkbox"/>	 mobilelogtospinner	2.3 KB	application/octet-
<input type="checkbox"/>	 spanner-export.json	124 B	application/json

2.2.Creating a spanner-export.json file

Notes: it's relative path to spanner-export.json, not absolute path

```
{
  "tables": [
    {
      "name": "MobileLog",
      "dataFiles": [
        "0803000000000000"
      ]
    }
  ]
}
```

2.3.Create bq to spanner dataflow job

Reference:

[Importing data from non-Cloud Spanner databases](#)