



Tot plugin GCP BigQuery

Control de versiones	1
Modelo de integración	1
Extracción de metadatos	2
Muestreo de datos	3
Gobierno activo	3
Credenciales requeridas	3
Creación de la cuenta de servicio	3
Extracción de metadatos	4
Muestreo de datos	4
Gobierno activo	5
Limitaciones BigQuery	6
Ejemplo de configuración	6

Control de versiones

Versión	Fecha de modificación	Responsable	Aprobador	Resumen de cambios
1.0	28/10/2022	Anjana Producto	Anjana Producto	Creación del documento
2.0	29/03/2023	Anjana Producto	Anjana Producto	Limitaciones

Modelo de integración

Extracción de metadatos

Se usará una conexión BigQuery con la credencial configurada para la extracción de metadatos así como listado por recurso.

Extrae los siguientes atributos que deben llamarse igual en la tabla `attribute_definition`, campo `name` para que aparezcan en la plantilla.

- **schema** con el valor de schema en la base de datos
- **physicalName** y **name** con el mismo valor, el nombre de la tabla
- **path** con la concatenación de los valores de dataset bigQuery y tabla bigQuery
- **infrastructure** con el valor seleccionado
- **technology** con el valor seleccionado
- **zone** con el valor seleccionado

También nos enviará los siguientes atributos relativos a los campos del recurso pedido:

- **name** con el valor del campo correspondiente
- **physicalName** con el valor del campo correspondiente
- **fieldDataType** con el tipo de dato definido para el campo correspondiente
- **position** posición que ocupa el campo correspondiente
- **nullable** indicando si el campo correspondiente es nullable
- **description** con el valor correspondiente para el campo

Muestreo de datos

Se usará una conexión BigQuery con la credencial configurada para ejecutar un select de los campos inventariados en Anjana Data con límite de registros.

Gobierno activo

La gestión de acceso requiere el plugin “Tot plugin GCP IAM” para que genere los roles (funciones) custom que representan a los DSA.

El presente plugin asociará dichos roles custom con usuarios y condiciones de acceso a nivel tabla siguiendo la recomendación del fabricante.




<https://cloud.google.com/bigquery/docs/table-access-controls#api>

Credenciales requeridas

Las credenciales requeridas se deberán configurar en el fichero yaml que utilizaremos para arrancar el servicio en la propiedad "totplugin.connection.credentialsContent"

Creación de la cuenta de servicio

Para GCP será necesario crear una cuenta de servicio en IAM para cada plugin de forma individual y tras eso asignarle los permisos necesarios para la ejecución de las tareas específicas de cada plugin.

	<code>gcp-bigquery@anjana-data-qa.iam.gserviceaccount.com</code>	<code>gcp-bigquery</code>
	<code>gcp-storage@anjana-data-qa.iam.gserviceaccount.com</code>	<code>gcp-storage</code>
	<code>gpc-iam@anjana-data-qa.iam.gserviceaccount.com</code>	<code>gpc-iam</code>

Para personalizar los permisos de forma más acorde será necesaria la creación de roles personalizados en los cuáles se engloban los permisos que luego son asociados a las cuentas de servicio.

IAM y administración < Crear rol

Las funciones personalizadas permiten agrupar permisos y asignarlos a las principales de tu organización o proyecto. Puedes seleccionar permisos de forma manual o importarlos desde otra función. [Más información](#)

Título *
tot-plugin-gcp-bigquery 23 de 100 caracteres

Descripción
Fecha de creación: 2022-10-21 29 de 256 caracteres

ID *
totpluggingcpbigquery

Etapa de lanzamiento de la función
Alfa

[+ AGREGAR PERMISOS](#)

No hay permisos asignados

Filtro Ingresar el nombre o el valor de la propiedad

<input type="checkbox"/>	Permisos ↑	Estado
No hay filas para mostrar		

i Es posible que algunos permisos estén asociados con terceros y que estos los revisen. Estos permisos contienen el servicio y nombre de dominio del tercero en el prefijo del permiso.

[CREAR](#) [CANCELAR](#)

Extracción de metadatos

Los permisos utilizados son los siguientes:

- `bigquery.datasets.get`
- `bigquery.tables.get`
- `bigquery.tables.list`

Muestreo de datos

Los permisos utilizados son los siguientes:

- `bigquery.datasets.get`
- `bigquery.tables.get`
- `bigquery.tables.getData`
- `bigquery.tables.list`
- `bigquery.jobs.create`

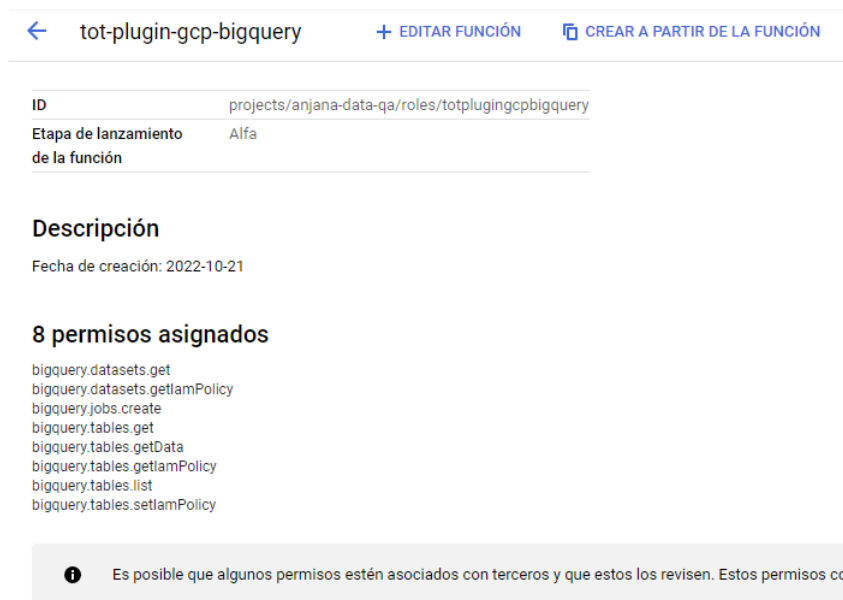
Gobierno activo

La gestión de acceso requiere el plugin “Tot plugin GCP IAM” para que genere los roles(funciones) custom que representan a los DSA.

Los permisos utilizados son los siguientes:

- `bigquery.datasets.get`
- `bigquery.tables.get`
- `bigquery.tables.getIamPolicy`
- `bigquery.tables.setIamPolicy`

En resumen los permisos utilizados para el rol personalizado serán los siguientes:



← tot-plugin-gcp-bigquery + EDITAR FUNCIÓN CREAR A PARTIR DE LA FUNCIÓN

ID: projects/anjana-data-qa/roles/totplugin-gcp-bigquery

Etapa de lanzamiento de la función: Alfa

Descripción

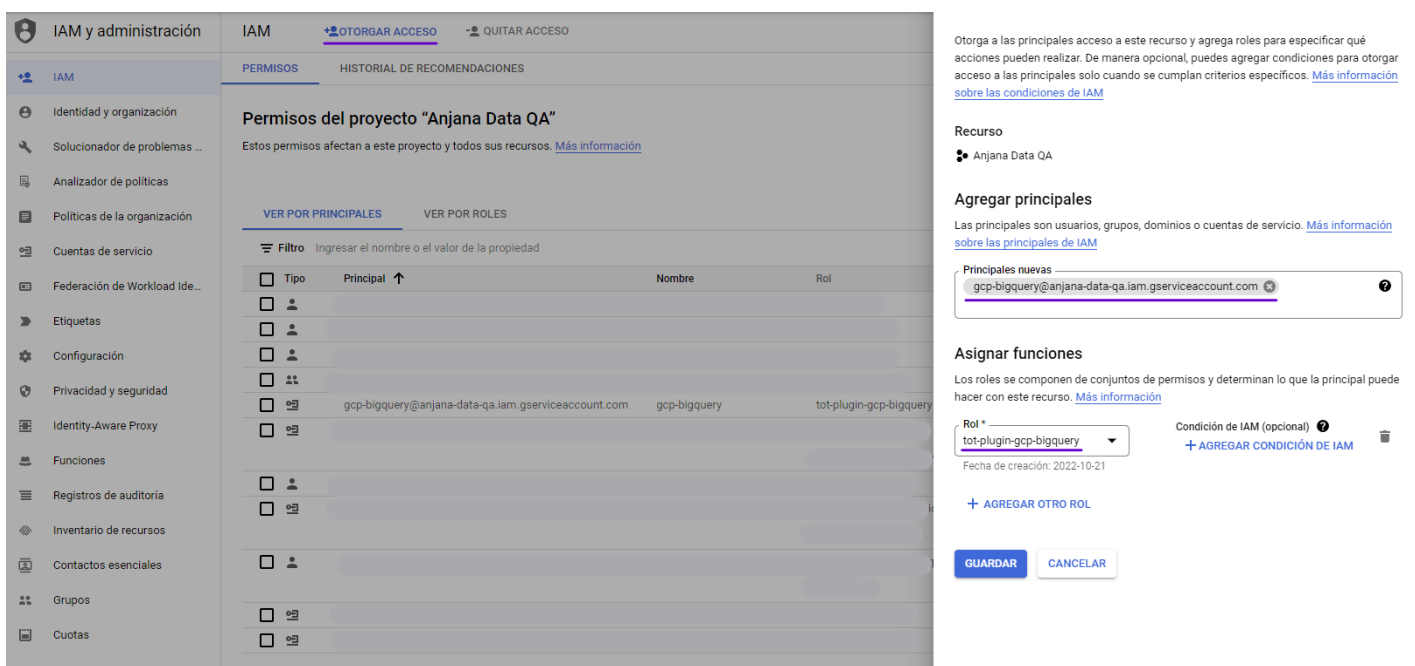
Fecha de creación: 2022-10-21

8 permisos asignados

- bigquery.datasets.get
- bigquery.datasets.getIamPolicy
- bigquery.jobs.create
- bigquery.tables.get
- bigquery.tables.getData
- bigquery.tables.getIamPolicy
- bigquery.tables.list
- bigquery.tables.setIamPolicy

! Es posible que algunos permisos estén asociados con terceros y que estos los revisen. Estos permisos con

Para asignar los permisos a la cuenta de servicio de bigquery tendremos que:



IAM y administración IAM OTORGAR ACCESO QUITAR ACCESO

PERMISOS HISTORIAL DE RECOMENDACIONES

Permisos del proyecto “Anjana Data QA”

Estos permisos afectan a este proyecto y todos sus recursos. [Más información](#)

VER POR PRINCIPALES VER POR ROLES

Filtro Ingresar el nombre o el valor de la propiedad

Tipo	Principal ↑	Nombre	Rol
<input type="checkbox"/>	<input type="checkbox"/>		
<input type="checkbox"/>	<input type="checkbox"/>		
<input type="checkbox"/>	<input type="checkbox"/>		
<input type="checkbox"/>	<input type="checkbox"/>		
<input type="checkbox"/>	<input checked="" type="checkbox"/>	gcp-bigquery@anjana-data-qa.iam.gserviceaccount.com	gcp-bigquery
<input type="checkbox"/>	<input type="checkbox"/>		tot-plugin-gcp-bigquery
<input type="checkbox"/>	<input type="checkbox"/>		
<input type="checkbox"/>	<input type="checkbox"/>		
<input type="checkbox"/>	<input type="checkbox"/>		
<input type="checkbox"/>	<input type="checkbox"/>		

Otorga a las principales acceso a este recurso y agrega roles para especificar qué acciones pueden realizar. De manera opcional, puedes agregar condiciones para otorgar acceso a las principales solo cuando se cumplan criterios específicos. [Más información sobre las condiciones de IAM](#)

Recurso
• Anjana Data QA

Agregar principales

Las principales son usuarios, grupos, dominios o cuentas de servicio. [Más información sobre las principales de IAM](#)

Principales nuevas
gcp-bigquery@anjana-data-qa.iam.gserviceaccount.com

Asignar funciones

Los roles se componen de conjuntos de permisos y determinan lo que la principal puede hacer con este recurso. [Más información](#)

Rol *
tot-plugin-gcp-bigquery

Condición de IAM (opcional) + AGREGAR CONDICIÓN DE IAM

Fecha de creación: 2022-10-21

+ AGREGAR OTRO ROL

GUARDAR CANCELAR

Limitaciones BigQuery

El número máximo de bindings para usuarios en una tabla es 1500, lo que quiere decir es que como máximo anjana podrá tener 1500 usuarios entre propietarios y adheridos en los DSAs que contengan un tabla en particular.

Ejemplo de configuración

Se han de revisar las configuraciones comunes en el doc de configuraciones “Anjana Data - Microservices configuration”

Configuraciones específicas:

- **connection:**
 - **credentialsContent:** Credenciales de acceso a GCP.
 - **sample-rows:** Tamaño del muestreo de datos.
 - **path-separator:** Separador que GCP usa en los roles

```
server:
  port: 15002

totplugin:
  location: http://totplugingcpbigqueryserver:15002/plugin/bigquery/api/v1
  server:
    url: http://totserver:15000/tot/
  aris:
    - ari: "anja:totplugin:extract:/Google/gcpBigQuery/devQA/"
    - ari: "anja:totplugin:sample:/Google/gcpBigQuery/devQA/"
    - ari: "anja:totplugin:im:/Google/gcpBigQuery/devQA/"
      imAri: "anja:totplugin:im:/Google/gcpIam/devQA/"
  connection:
    credentialsContent: |
      {
        "type": "service_account",
        "project_id": "anjana-data-qa",
        "private_key_id": "*****",
        "private_key": "-----BEGIN PRIVATE KEY-----\n\n-----END PRIVATE KEY-----\n",
        "client_email": "gcp-bigquery@*****.com",
        "client_id": "*****",
        "auth_uri": "https://accounts.google.com/o/oauth2/auth",
        "token_uri": "https://oauth2.googleapis.com/token",
        "auth_provider_x509_cert_url": "https://www.googleapis.com/oauth2/v1/certs",
        "client_x509_cert_url":
"https://www.googleapis.com/robot/v1/metadata/x509/gcp-bigquery%40*****.com"
      }
    sample-rows: 15
    path-separator: "/"
```