

Creación de Conexión entre Hive y un Index de ElasticSearch

Francisco Clemente Fernández

PARTE 1 - Configuración ES-Hadoop

1- Creamos un cluster llamado CLuster-Hadoop-project:

Clústeres							
+ CREATE CLUSTER ACTUALIZAR INICIAR DETENER BORRAR REGIONES + 5 ALERTAS RECC							
Filtro Busca el clúster por propiedades y presiona Intro							
El servidor no pudo completar tu solicitud.							
<input type="checkbox"/>	Nombre ↑	Estado	Región	Zona	Total de nodos trabajadores	¿Tiene VMs flexibles?	Eliminación por
<input type="checkbox"/>	cluster-hadoop-project	En ejecución	europa-west10	europa-west10-a	2	No	Desactivado

2- Creamos un Bucket llamado **Bucket-paraelastic**:

Buckets						
+ CREAR ACTUALIZAR IR A RUTA APRENDIZAJE						
Review the soft delete settings on your buckets. Billing for soft deleted objects began on September 1st. MÁS INFORMACIÓN ADMINISTRA POLÍTICAS DE BORRAR DE FORMA NO DEFINITIVA						
Filtro Filtrar depósitos						
<input type="checkbox"/>	Nombre ↑	Fecha de creación	Tipo de ubicación	Ubicación	Clase de almacenamiento predeterminada	Última modificación
<input type="checkbox"/>	bucket-paraelastic	3 nov 2024 20:44:08	Multi-region	us	Standard	3 nov 2024 20:44:08
<input type="checkbox"/>	dataproc-staging-europe-west10-731439...	3 nov 2024 20:38:24	Region	europa-west10	Standard	3 nov 2024 20:38:24
<input type="checkbox"/>	dataproc-temp-europe-west10-7314398...	3 nov 2024 20:38:25	Region	europa-west10	Standard	3 nov 2024 20:38:25

3- Extraemos el contenido del zip que hemos descargado previamente, y **subimos** ambos **jar (commons y elasticsearch-hadoop)** a nuestro **bucket** en google storage dentro del directorio **Jar** :

Detalles del bucket

IR A RUTA

ACTUALIZAR

APRENDIZAJE

bucket-paraelastic

Ubicación

us (varias regiones en Estados Unidos)

Clase de almacenamiento

Standard

Acceso público

No público

Protección

Borrar de forma no definitiva

OBJETOS

CONFIGURACIÓN

PERMISOS

PROTECCIÓN

CICLO DE VIDA

OBSERVABILIDAD

INFORMES DE INVENTARIO

OPERACIONES

Navegador de carpetas

bucket-paraelastic

jar/

Depósitos > bucket-paraelastic > jar

CREAR CARPETA

SUBIR



TRANSFERIR LOS DATOS

OTROS SERVICIOS

Filtrar solo por prefijo de nombre

Filtro Filtrar objetos y carpetas

Mostrar Solo objetos activos

<input type="checkbox"/>	Nombre	Tamaño	Tipo	Fecha de creación	Clase de almacenamiento	Última modificación
<input type="checkbox"/>	 commons-httpclient-3.1.jar	297.9 KB	application/octet-stream	3 nov 2024 20:44:46	Standard	3 nov 2024 20:44:46
<input type="checkbox"/>	 elasticsearch-hadoop-8.14.1.jar	2.1 MB	application/octet-stream	3 nov 2024 20:44:47	Standard	3 nov 2024 20:44:47

4- Descargamos los **jar** desde el **bucket** al **filesystem** del cluster.

```
individual files in /usr/share/doc/*/copyright.

Ubuntu comes with ABSOLUTELY NO WARRANTY, to the extent permitted by
applicable law.

fdezeta@cluster-hadoop-project-m:~$ gs://bucket-paraelastic/jar/commons-httpclient-3.1.jar
-bash: gs://bucket-paraelastic/jar/commons-httpclient-3.1.jar: No such file or directory
fdezeta@cluster-hadoop-project-m:~$ gduutil cp gs://bucket-paraelastic/jar/commons-httpclient-3.1.jar

Command 'gduutil' not found, did you mean:

  command 'gsutil' from snap google-cloud-cli (499.0.0)
  command 'gsutil' from snap google-cloud-sdk (499.0.0)
  command 'gsutil' from deb gsutil (3.1-1)

See 'snap info <snapname>' for additional versions.

fdezeta@cluster-hadoop-project-m:~$ gsutil cp gs://bucket-paraelastic/jar/commons-httpclient-3.1.jar
CommandException: Wrong number of arguments for "cp" command.
fdezeta@cluster-hadoop-project-m:~$ gsutil cp gs://bucket-paraelastic/jar/commons-httpclient-3.1.jar .
Copying gs://bucket-paraelastic/jar/commons-httpclient-3.1.jar...
/ [1 files][297.8 KiB/297.8 KiB]
Operation completed over 1 objects/297.8 KiB.
fdezeta@cluster-hadoop-project-m:~$ gsutil cp gs://bucket-paraelastic/jar/elasticsearch-hadoop-8.14.1.jar .
Copying gs://bucket-paraelastic/jar/elasticsearch-hadoop-8.14.1.jar...
- [1 files][ 2.1 MiB/ 2.1 MiB]
Operation completed over 1 objects/2.1 MiB.
```

PARTE 2 - Configuración server Elasticsearch

5- Creamos una nueva **VM** para elastic: