

CODEFEST AD ASTRA

**PRESENTADO A:
FUERZA AEREA DE COLOMBIA**

**PRESENTADO POR:
VICTOR HUGO CARDENAS
JUAN BARRERAS
NATALIA HERNANDEZ
IVAN MUÑOZ
JUAN OSPITIA**

UNICOMFACAUCA

2022

Contenido

DESCRIPCION DEL SOFTWARE USADO	3
GitHub	3
Azure	3
Python	3
QGis.....	3

DESCRIPCION DEL SOFTWARE USADO

GitHub

Dado que se pide una licencia “permisiva” se debe tener en cuenta que se otorga permiso público de uso, modificación y distribución, sin ninguna condición de bajada. Las licencias más populares de código abierto son MIT, Apache 2.0, ISC y BSD.

Nosotros elegimos Apache 2.0 debido a que usamos Python como lenguaje referente en el repositorio. Además, se añadió el README FILE, y se creó el repositorio colaborativo.

Azure

Se creo el clúster y la maquina virtual con las especificaciones dadas por el proyecto, se enlazó el repositorio de imágenes requeridas por los organizadores del proyecto al almacenamiento de la nube mediante dirección SAS y su respectivo token ,posteriormente se copiaron al contenedor creado por el equipo, después se configuró el clúster mediante Python DataBrick y la máquina virtual, de los cuales esta ultima se conectó al contenedor creado anteriormente mediante Microsoft Azure Storage Explorer y se instalaron los programas de SIG requeridos para visualizar las imágenes.

Python

Se creo un script de Python el cual usa las librerías Rasterio para la manipulación de imágenes satelitales, Numpy permite leer las imágenes en formato de array y Base64 el cual será usado para la codificación de las imágenes.

QGis

Es un software GIS de código abierto el cual es compatible con Python y será el software de referencia para probar y manipular las imágenes modificadas por el software y le permitirá al operador probar las imágenes y manipularlas de acuerdo a su criterio.

CONFIGURACION DE AZURE

Data Factory

De acuerdo a la dirección SAS y el Token suministrados por los organizadores del concurso procedemos a realizar la ingesta en la página principal Data Factory.



Se configura el origen como un Azure Blob con la SAS y el token proporcionado por los Organizadores y se establece como copia binaria.

Almacén de datos de origen

Especifique el almacén de datos de origen para la tarea de copia. Puede utilizar una conexión de almacén de datos existente o especificar un nuevo almacén de datos.

Tipo de origen: Almacenamiento de blobs de Azure

Conexión *: Codefest [Editar] [Nueva conexión]

Archivo o carpeta *: imgs/ [Examinar]

Opciones:

- ☒ Copia binaria ⓘ

Tipo de compresión: Ninguno

Recursivamente ⓘ: ☒

Eliminar archivos tras la finalización ⓘ

< Anterior [Siguiente >] Cancelar

Se crea un contenedor en Microsoft Storage y se configura la salida como Azure Blob y se envía al contenedor creado.

Almacén de datos de destino

Especifique el almacén de datos de destino para la tarea de copia. Puede utilizar una conexión de almacén de datos existente o especificar un nuevo almacén de datos.

Tipo de destino

Almacenamiento de blobs de Azure

Conexión *

AzureBlobStorage1

Editar Nueva conexión

Ruta de acceso de la carpeta *

Si la identidad que usa para acceder al almacén de datos solo tiene permiso para el subdirectorio en lugar de tenerlo para toda la cuenta, especifique la ruta de acceso a examinar.

contenedor

Examinar

Nombre de archivo

Los nombres de archivo los define el origen.

Tipo de compresión

Ninguno

Comportamiento de copia ①

Ninguno

Número máximo de conexiones simultáneas ①

< Anterior

Siguiente >

Cancelar

Se revisa, prueba y finaliza.

Almacenamiento de blobs de Azure
Región: East US 2

Correcto
Región de Azure IR: East US 2

Almacenamiento de blobs de Azure
Región: East US 2

Datos leídos: ①61,435 GB

Datos escritos: ①61,435 GB

Archivos leídos: ①33

Archivos escritos: ①33

Conexiones máximas: ①160

Conexiones máximas: ①245

Duración de la copia00:02:37

Rendimiento: ①400,698 MB/s

✓ Almacenamiento de blobs de Azure → Almacenamiento de blobs de Azure

Hora de inicioSep 24, 2022, 6:22:22 am

DIU utilizados ①32

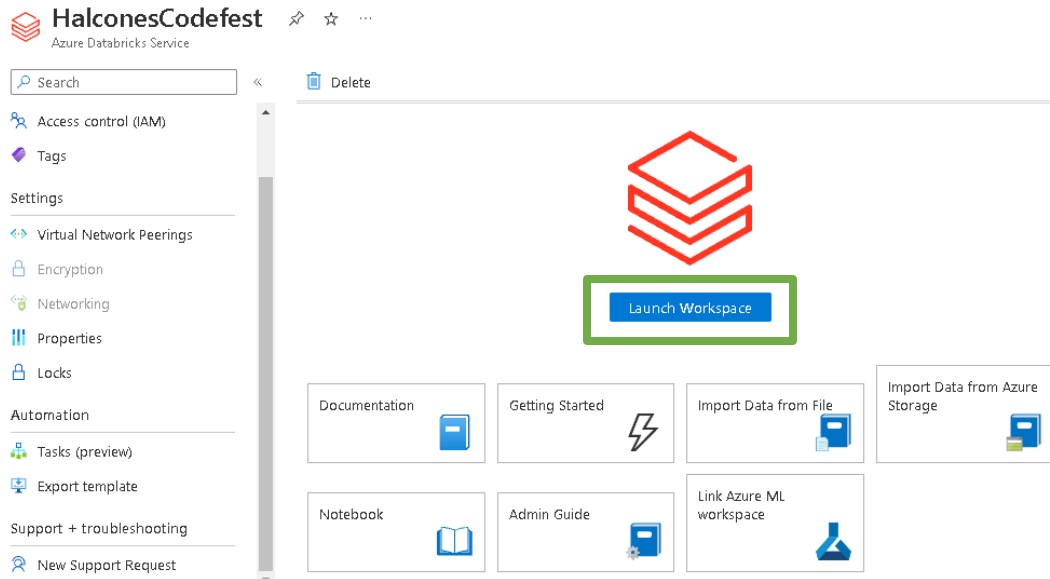
Copias en paralelo utilizadas ①16

✓ Duración00:02:37

Detalles	Duración del trabajo	Duración total
✓ Cola ①		00:00:03
✓ Transferir ①	<div>Origen de la lista ①00:00:01</div> <div>Leyendo desde el origen ①00:02:18</div> <div>Escribiendo en el receptor ①00:02:28</div>	00:02:32

DataBricks

Se configura el Databricks y le damos en Launch Workspace.



Creamos el cluster


CD. CAMILO's Cluster

Performance

Databricks runtime version 

Runtime: 11.2 ML (Scala 2.12, Spark 3.3.0) | v

☐ Use Photon Acceleration 

Worker type 



Standard_E8d_v4 64 GB Memory, 8 Cores | v

Min workers

2


Max workers


03

 ☐ Spot instances 

Driver type

Same as worker 64 GB Memory, 8 Cores | v

☒ Enable autoscaling 

☐ Terminate after minutes of inactivity 

Tags

Add tags

Create Cluster

Cancel