



Diplomado en Ciencia de Datos UNAM

Módulo 13 Datos Masivos

Septiembre de 2023

Sergio Ibarra Ramírez

Pantalla de Cluster en Ejecucion

Estado de la prueba gratuita: Te quedan \$5,044.78 de crédito y 90 días. Con una cuenta completa, obtendrás acceso ilimitado a todas las funciones de Google Cloud Platform. **DESCARTAR** **ACTIVAR**

Google Cloud My First Project

Dataproc Detalles del cl... **ENVIAR TRABAJO** **ACTUALIZAR** **INICIAR**

Trabajos alojados en clústeres

- Clústeres
- Trabajos
- Flujos de trabajo
- Políticas de ajuste de escala...

Sin servidores

- Lotes
- Interactiva

Servicios de Metastore

- Metastore
- Federación

Servicios públicos

- Intercambio de componentes
- Workbench

Consider using Auto Zone rather than selecting a zone manually. See <https://cloud.google.com/dataproc/docs/concepts/configuring-clusters/auto-zone>

Nombre cluster-sir-dcd05

UUID del clúster 782eaaa4-0df7-4fa9-be40-060c18582916

Tipo Clúster de Dataproc

Estado ✓ En ejecución

SUPERVISIÓN TRABAJOS INSTANCIAS DE VM CONFIGURACIÓN

GUARDAR COMO PANEL RESTABLECER EL ZOOM 1 hora

Memoria YARN

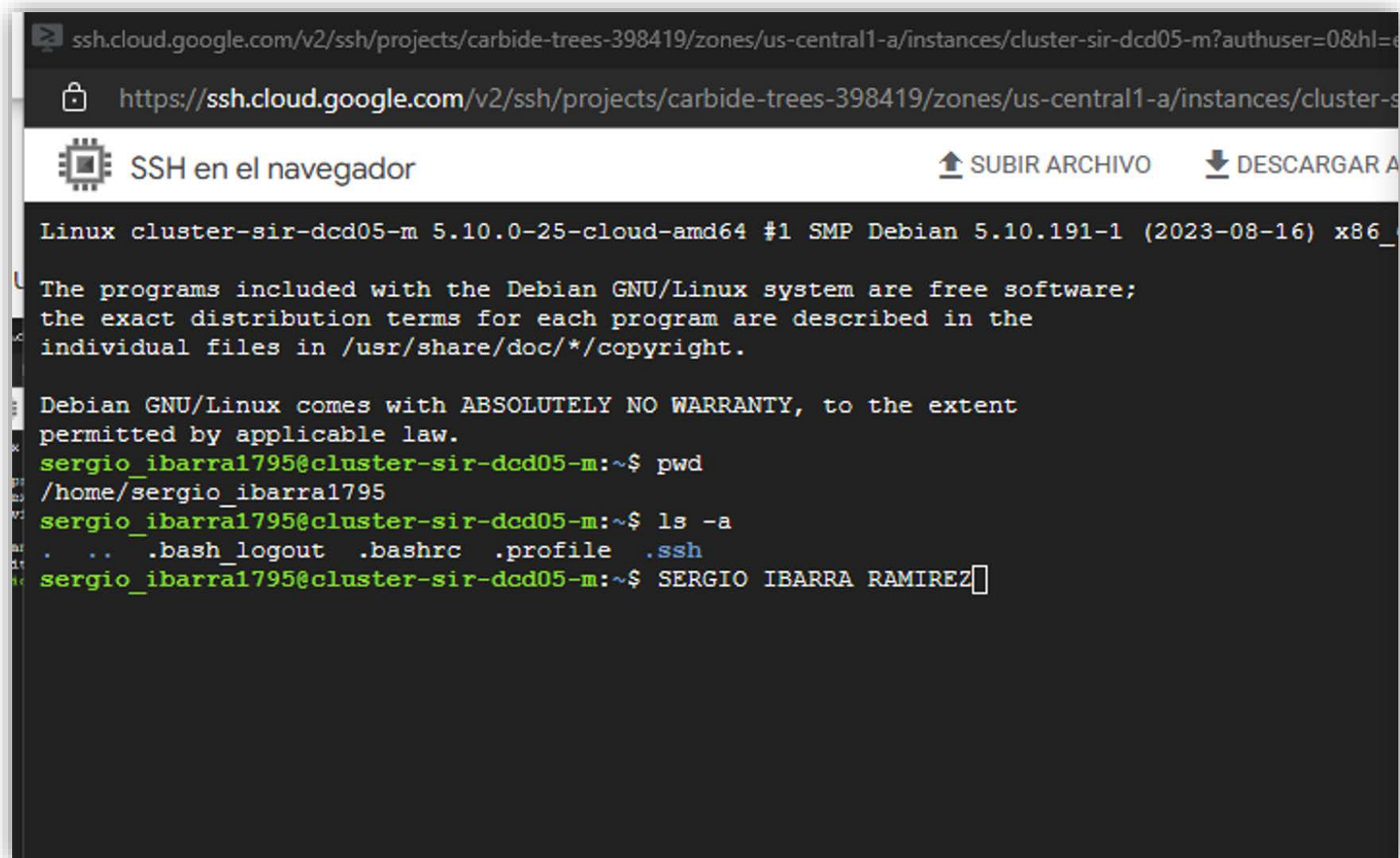
15GiB

10GiB

5GiB

allocated

Pantalla de SSH del Master



The screenshot shows a web browser window with the address bar displaying the URL: `ssh.cloud.google.com/v2/ssh/projects/carbide-trees-398419/zones/us-central1-a/instances/cluster-sir-dcd05-m?authuser=0&hl=es`. The browser's address bar also shows the full URL: `https://ssh.cloud.google.com/v2/ssh/projects/carbide-trees-398419/zones/us-central1-a/instances/cluster-sir-dcd05-m?authuser=0&hl=es`. The browser's title bar reads "SSH en el navegador". In the top right corner of the browser window, there are two buttons: "SUBIR ARCHIVO" (Upload File) and "DESCARGAR ARCHIVO" (Download File). The main content area of the browser is a terminal window. The terminal output is as follows:

```
Linux cluster-sir-dcd05-m 5.10.0-25-cloud-amd64 #1 SMP Debian 5.10.191-1 (2023-08-16) x86_64

The programs included with the Debian GNU/Linux system are free software;
the exact distribution terms for each program are described in the
individual files in /usr/share/doc/*/copyright.

Debian GNU/Linux comes with ABSOLUTELY NO WARRANTY, to the extent
permitted by applicable law.
sergio_ibarra1795@cluster-sir-dcd05-m:~$ pwd
/home/sergio_ibarra1795
sergio_ibarra1795@cluster-sir-dcd05-m:~$ ls -la
.  ..  .bash_logout  .bashrc  .profile  .ssh
sergio_ibarra1795@cluster-sir-dcd05-m:~$ SERGIO IBARRA RAMIREZ
```

MapReduce Job History

carbide-trees-398419 > cluster-sir-1105 [Sign out](#)

Logged in as: dr.who

MapReduce Job job_1694277344266_0001

Application

- Job**
 - [Overview](#)
 - [Counters](#)
 - [Configuration](#)
 - [Map tasks](#)
 - [Reduce tasks](#)
- Tools**

Job Overview

Job Name: word count
User Name: sergio_ibarra1795
Queue: default
State: SUCCEEDED
Uberized: false
Submitted: Sat Sep 09 18:08:48 UTC 2023
Started: Sat Sep 09 18:09:03 UTC 2023
Finished: Sat Sep 09 18:09:28 UTC 2023
Elapsed: 25sec
Diagnostics:
Average Map Time: 9sec
Average Shuffle Time: 7sec
Average Merge Time: 0sec
Average Reduce Time: 0sec

ApplicationMaster

Attempt Number	Start Time	Node	Logs
1	Sat Sep 09 18:08:56 UTC 2023	cluster-sir-dcd05-w-0.us-central1-a.c.carbide-trees-398419.internal:8042	/gateway/default/jobhistory/logs

Task Type	Total	Complete	
Map	1	1	
Reduce	3	3	
Attempt Type	Failed	Killed	Successful
Maps	0	0	1
Reduces	0	0	3

carbide-trees-398419 > cluster-sir-1105 [Sign out](#)

Hadoop

[Overview](#) [Datanodes](#) [Datanode Volume Failures](#) [Snapshot](#) [Startup Progress](#) [Utilities](#)

Browse Directory

[Go!](#) [New](#) [Refresh](#) [Download](#)

Show entries Search:

<input type="checkbox"/>	Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name	
<input type="checkbox"/>	drwxr-xr-x	sergio_ibarra1795	hadoop	0 B	Sep 09 12:39	0	0 B	sirilo	Delete
<input type="checkbox"/>	drwxr-xr-x	sergio_ibarra1795	hadoop	0 B	Sep 09 12:09	0	0 B	wordcount	Delete

Showing 1 to 2 of 2 entries [Previous](#) [1](#) [Next](#)

Hadoop, 2022.

Jupyter Ecobici

