

Creación de cluster

My cluster Updated less than a minute ago [Terminate](#) [Clone in AWS CLI](#) [Clone](#)

▼ Summary

Cluster info

Cluster ID
j-2D8DPY7VBW9FA

Cluster ARN
arn:aws:elasticmapreduce:us-east-1:923015016075:cluster/j-2D8DPY7VBW9FA

Cluster configuration
Instance groups

Capacity
1 Primary 1 Core 1 Task

Applications

Amazon EMR version
emr-7.11.0

Installed applications
Flink 1.20.0, HCatalog 3.1.3, Hadoop 3.4.1, Hive 3.1.3, Hue 4.11.0, JupyterHub 1.5.0, Livy 0.8.0, Spark 3.5.6, Tez 0.10.2, Zeppelin 0.11.1

Cluster management

Log destination in Amazon S3
emrsamuel

Persistent application UIs
[Spark History Server](#)
[YARN Timeline Server](#)
[Tez UI](#)

Primary node public DNS
ec2-98-92-71-66.compute-1.amazonaws.com
[Connect to the Primary node using SSH](#)
[Connect to the Primary node using SSM](#)

Status and time

Status
Waiting

Creation time
November 09, 2025, 22:03 (UTC-05:00)

Elapsed time
1 hour, 9 minutes

Properties | Bootstrap actions | Instances (Hardware) | Steps | Applications | Configurations | Monitoring | Events | Tags (0)

Operating system [info](#)

Amazon Linux release
2023.9.20251014.0

Cluster logs [info](#)

Archive log files to Amazon S3
Turned on

Amazon S3 location
s3://emrsamuel/ [link](#)

Encryption for logs
Turned off

Cluster termination and node replacement [info](#)

[Edit](#)

Termination option
Automatically terminate cluster after idle time

Idle time
1 hour

Termination protection
Off

Unhealthy node replacement
On

Acceso ssh

```
#####
\##|
\#/
V~' '-> https://aws.amazon.com/linux/amazon-linux-2023

Last login: Mon Nov 17 02:00:08 2025

EEEEEEEEEEEEEEEEEEEE MMMMMMMM MMMMMMMM RRRRRRRRRRRRRR
E::::::::::::::::::E M::::::::M M::::::::M R::::::::::::R
EE::::::::::::::::::E M::::::::M M::::::::M R::::RRRRR::::R
E::::E EEEEE M::::::::M M::::::::M RR::::R R::::R
E::::E M::::::::M M::M::M M::M::M R::R R::R
E::::EEEEEEEE M::M M::M M::M R::RRRRR::::R
E::::::::::::::::::E M::M M::M M::M R::::::::::::RR
E::::EEEEEEEE M::M M::M M::M R::RRRRR::::R
E::::E M::M M::M M::M R::R R::R
E::::E EEEEE M::M MMM M::M R::R R::R
EE::::::::::::::::::E M::M M::M R::R R::R
E::::::::::::::::::E M::M M::M RR::R R::R
EEEEEEEEEEEEEEEEEEEE MMMMMMMM MMMMMMMM RRRRRR RRRRRR

[hadoop@ip-172-31-71-159 ~]$ hdfs dfs -ls /user
Found 9 items
drwxrwxrwx - hadoop hdfsadmingroup 0 2025-11-17 01:48 /user/hadoop
drwxr-xr-x - mapred mapred 0 2025-11-17 01:48 /user/history
drwxrwxrwx - hdfs hdfsadmingroup 0 2025-11-17 01:48 /user/hive
drwxrwxrwx - hue hue 0 2025-11-17 01:48 /user/hue
drwxrwxrwx - livy livy 0 2025-11-17 01:48 /user/livy
drwxrwxrwx - oozie oozie 0 2025-11-17 01:48 /user/oozie
drwxrwxrwx - root hdfsadmingroup 0 2025-11-17 01:48 /user/root
drwxrwxrwx - spark spark 0 2025-11-17 01:48 /user/spark
drwxrwxrwx - zeppelin hdfsadmingroup 0 2025-11-17 01:48 /user/zeppelin
[hadoop@ip-172-31-71-159 ~]$
```

Security groups del cluster

EC2 > Grupos de seguridad > sg-094a6e3d4ffc780d8 - ElasticMapReduce-master > Editar reglas de entrada

Editar reglas de entrada Información

Las reglas de entrada controlan el tráfico entrante que puede llegar a la instancia.

ID de la regla del grupo de seguridad	Tipo <small>Información</small>	Protocolo <small>Información</small>	Intervalo de puertos <small>Información</small>	Origen <small>Información</small>	Descripción: opcional <small>Información</small>	<small>Eliminar</small>
sg-0d56431f19ca7120b	TCP personalizado	TCP	9870	Persona... Q 0.0.0.0/0		<small>Eliminar</small>
sg-0efdb63a66b23966e	TCP personalizado	TCP	8890	Persona... Q 0.0.0.0/0		<small>Eliminar</small>
sg-052288ef6c2b2e17	TCP personalizado	TCP	8443	Persona... Q pl-f8bd5e91		<small>Eliminar</small>
sg-011a04e586d9e8960	TCP personalizado	TCP	8888	Persona... Q 0.0.0.0/0		<small>Eliminar</small>
sg-01d7a6ebb354c7b95	TCP personalizado	TCP	9443	Persona... Q 0.0.0.0/0		<small>Eliminar</small>
sg-03f32252e1aa8df9c	Todos los UDP	UDP	0 - 65535	Persona... Q sg-094a6e3d4ffc780d8		<small>Eliminar</small>
sg-0cbc9ee8a7bde42dc	Todos los ICMP IPv4	ICMP	Todo	Persona... Q sg-094a6e3d4ffc780d8		<small>Eliminar</small>
sg-0201718b901fc665e	SSH	TCP	22	Persona... Q 0.0.0.0/0		<small>Eliminar</small>
sg-0897276182a3fc0b9	TCP personalizado	TCP	14000	Persona... Q 0.0.0.0/0		<small>Eliminar</small>

Actividades del video

Acceso a JupyterHub

jupyterhub Untitled1 Last Checkpoint: hace unos segundos (autosaved) Logout Control Panel

File Edit View Insert Cell Kernel Widgets Help Trusted | PySpark


```
In [1]: spark
Starting Spark application
  ID      YARN Application ID  Kind  State  Spark UI  Driver log  User  Current session?
  0  application_1763344109923_0001  pyspark  idle           None           ✓

SparkSession available as 'spark'.
<pyspark.sql.session.SparkSession object at 0xfffff92ac3730>

In [2]: sc
<SparkContext master=yarn appName=livy-session-0>

In [ ]:
```

Acceso a Athena

 **Zeppelin**

Notebook ▾Job

prueba ▶ 🔍 📄 📌 📁 📤 📥 🔄 🔒 ⚙️ Head 🔍 🗑️

```
%spark.pyspark
spark

<pyspark.sql.session.SparkSession object at 0xffff73f50040>









Took 0 sec. Last updated by anonymous at November 16 2025, 9:21:27 PM.
```

```
%spark.pyspark
sc

<SparkContext master=yarn appName=Zeppelin>

Took 0 sec. Last updated by anonymous at November 16 2025, 9:21:35 PM.
```

```
%sql
show databases
```

 settings ▾

namespace
default