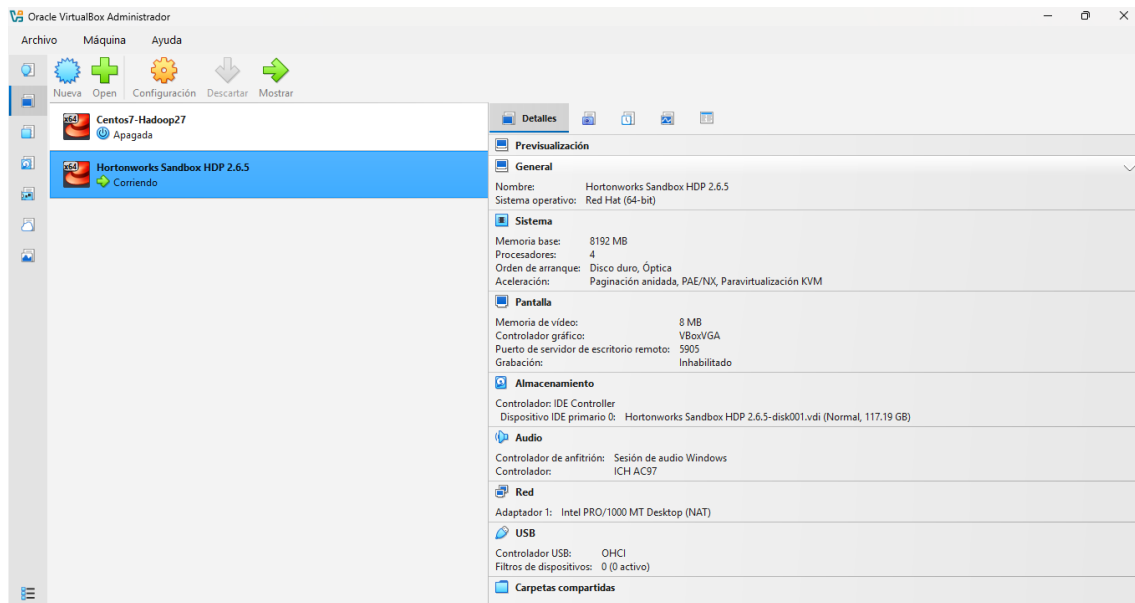


## LABORATORIO 1

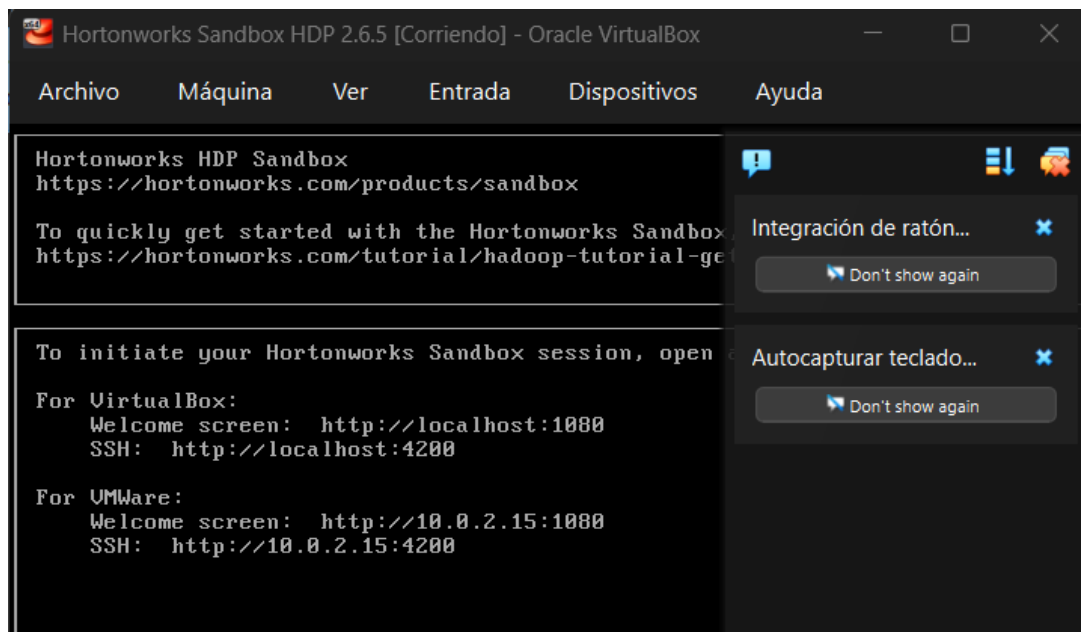
### GRUPO 9:

- Cabana Cazani, Gabriel
- Larico Cruz, Diego

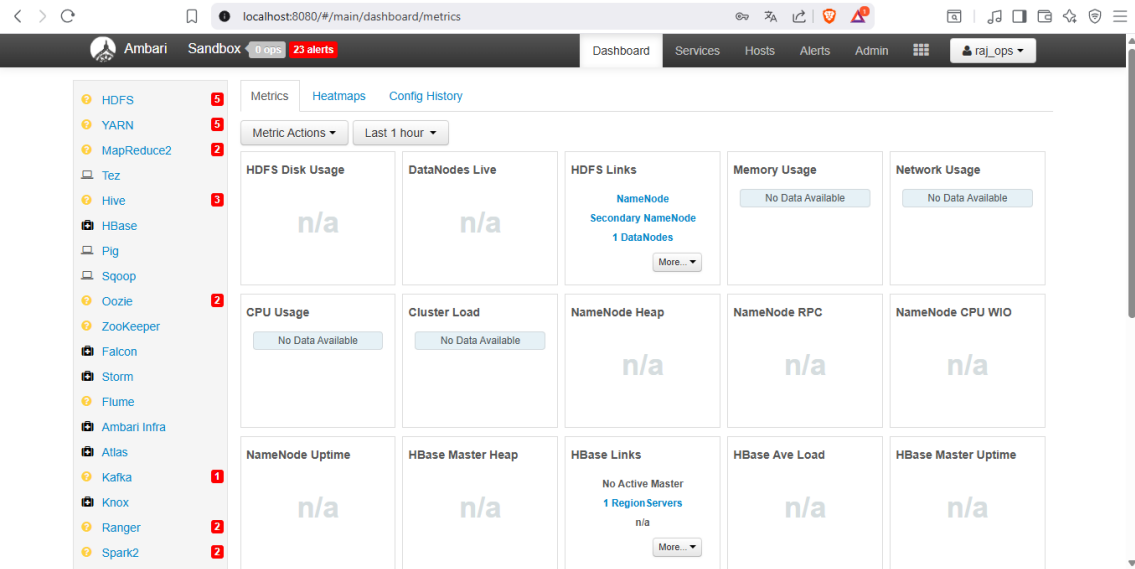
### Arranque de Hortonworks HDP



### Arranque desde VBox



Arranque de Ambari



1 Background Operation Running

Operations	Start Time	Duration	Show: All (10)
Start All Services	Today 15:26	244.05 secs	31%
Start All Services	Today 14:23	1.03 hours	100%
Stop required services	Mon Jun 18 2018 11:17	22.30 secs	100%
Stop required services	Mon Jun 18 2018 11:17	8.83 secs	100%
Stop required services	Mon Jun 18 2018 11:17	7.35 secs	100%
Stop required services	Mon Jun 18 2018 11:17	3.30 secs	100%
Stop required services	Mon Jun 18 2018 11:16	33.38 secs	100%
Stop required services	Mon Jun 18 2018 11:16	3.85 secs	100%
Stop required services	Mon Jun 18 2018 11:16	12.76 secs	100%

OK

### Version de Spark

```
< > ↺ localHost:4200
```

```
sandbox-hdp login: root
root@sandbox-hdp.hortonworks.com's password:
You are required to change your password immediately (root enforced)
Last login: Mon Sep 22 20:36:36 2025 from 172.18.0.2
Changing password for root.
(current) UNIX password:
New password:
Retype new password:
[root@sandbox-hdp ~]# spark version
-bash: spark: command not found
[root@sandbox-hdp ~]# hdfs dfs -ls
[root@sandbox-hdp ~]# spark -shell -version
-bash: spark: command not found
[root@sandbox-hdp ~]# spark-shell -version
SPARK_MAJOR_VERSION is set to 2, using Spark2
Setting default log level to "WARN".
To adjust logging level use sc.setLogLevel(newLevel). For SparkR, use setLogLevel(newLevel).
Spark context Web UI available at http://sandbox-hdp.hortonworks.com:4040
Spark context available as 'sc' (master = local[*], app id = local-1758573730345).
Spark session available as 'spark'.
Welcome to

 _/_/   _/_/   _/_/   _/_/   _/_/
/_/_/   _/_/   _/_/   _/_/   _/_/
/_/_/   _/_/   _/_/   _/_/   _/_/

version 2.3.0.2.6.5.0-292

Using Scala version 2.11.8 (OpenJDK 64-Bit Server VM, Java 1.8.0_171)
Type in expressions to have them evaluated.
Type :help for more information.

scala>
```

Subimos archivo flights.csv

localhost:8080/#/main/view/FILES/auto\_files\_instance

Ambari Sandbox 0 ops 3 alerts

Dashboard Services Hosts Alerts Admin

raji\_ops

Home Files Recent / > user > admin

Total: 1 files or folders

+ Select All + New Folder Upload

Search in current directory...

Name >	Size >	Last Modified >	Owner >	Group >	Permission
flights.csv	40.9 MB	2025-09-22 16:05	raji_ops	hdfs	-rw-r--r--

Ambari

Sandbox

0 ops

3 alerts

Dashboard

Services

Hosts

Alerts

Admin

raj\_ops

Hive

Query

Saved Queries

History

UDFs

Upload Table

Database Explorer

default

Search tables...

Search columns in result tables...

Databases

Search Results

default

flights

Query Editor

Worksheet

1 CREATE EXTERNAL TABLE flights (  
2 year STRING,  
3 month STRING,  
4 day STRING,  
5 dep\_time STRING,  
6 dep\_delay STRING,  
7 arr\_time STRING,  
8 arr\_delay STRING,  
9 carrier STRING,  
10 tailnum STRING,  
11 flight STRING,  
12 origin STRING,  
13 dest STRING,  
14 air\_time STRING,  
15 distance STRING,  
16 hour STRING,  
17 minute STRING  
18 )  
19 ROW FORMAT DELIMITED  
20 FIELDS TERMINATED BY ','

Execute

Explain

Upload

Save as...

New Worksheet

SQL

TEZ

Ambari

Sandbox

0 ops

3 alerts

Dashboard

Services

Hosts

Alerts

Admin

raj\_ops

Hive

Query

Saved Queries

History

UDFs

Upload Table

Upload from Local

File type

CSV

Database

foodmart

Stored as

ORC

Upload from HDFS

Select from local

Seleccionar archivo

Ningún archivo seleccionado

Table name

flights

Contains endlines?

Upload Table

column1

column2

column3

column4

colour

STRING

STRING

STRING

STRING

STRIN

id

year

month

day

dep\_tin

0

2013

1

1

517.0

1

2013

1

1

533.0

2

2013

1

1

542.0

3

2013

1

1

544.0