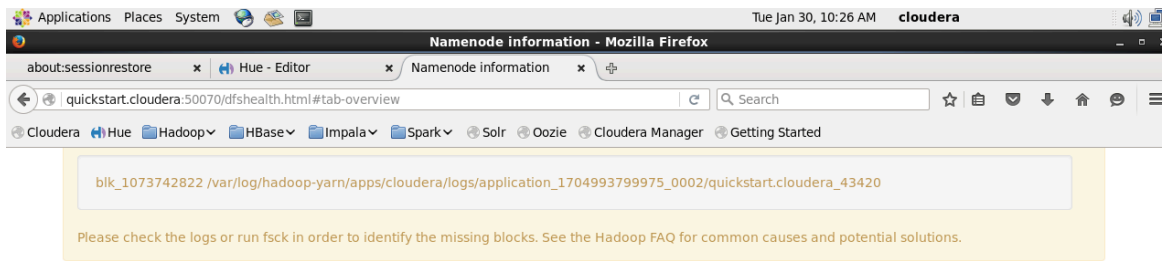


# Configuración y monitorización de sistemas Hadoop

# ACTIVIDAD 1

## WEB UI

Pantalla de información general:



### Overview 'quickstart.cloudera:8020' (active)

|                |  |
|----------------|--|
| Started:       | Wed Jan 17 12:12:14 -0800 2024                             |
| Version:       | 2.6.0-cdh5.13.0, r42e8860b182e55321bd5f5605264da4adc8882be |
| Compiled:      | Wed Oct 04 11:08:00 -0700 2017 by jenkins from Unknown     |
| Cluster ID:    | CID-a24185f9-a545-40fe-9553-84c3fdca489f                   |
| Block Pool ID: | BP-1067413441-127.0.0.1-1508775264580                      |

### Summary

Security is off.

Safemode is off.

1,130 files and directories, 965 blocks = 2,095 total filesystem object(s).

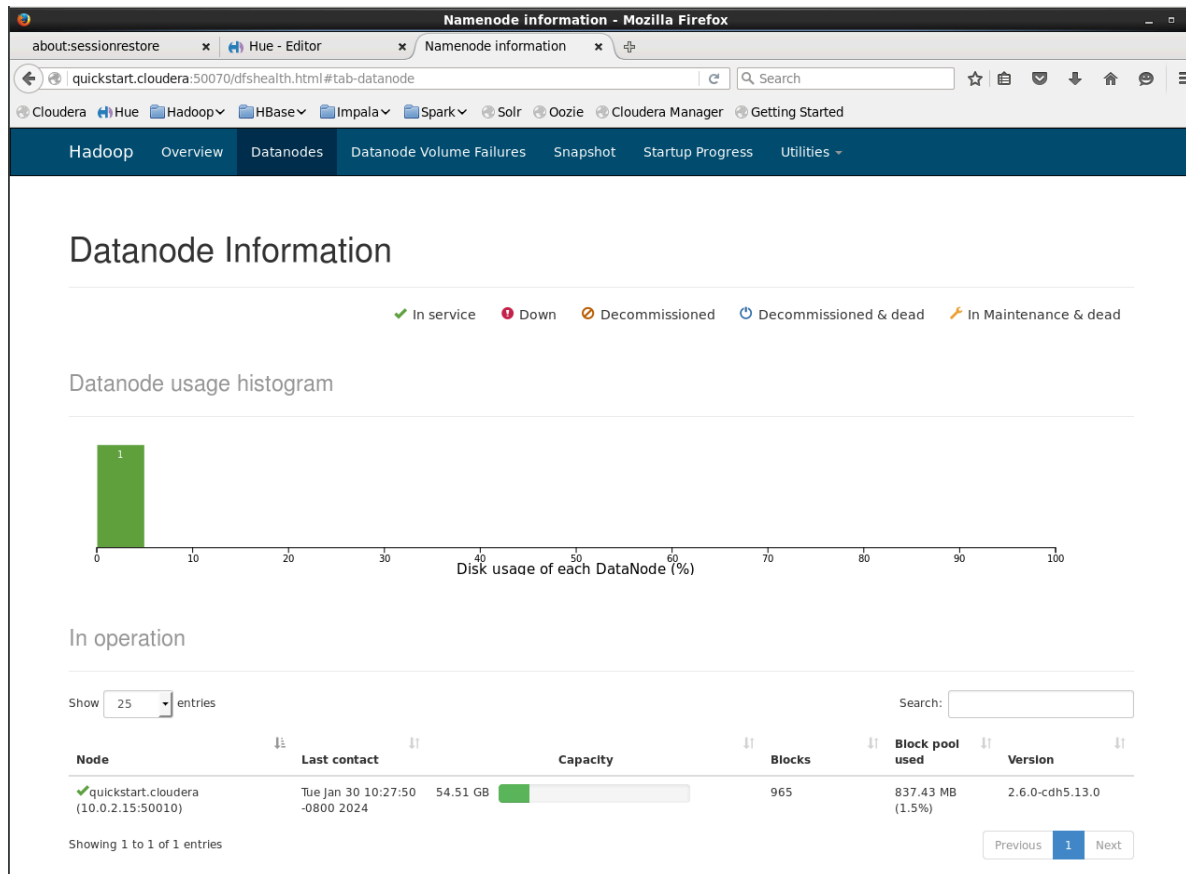
Heap Memory used 33.31 MB of 61.17 MB Heap Memory. Max Heap Memory is 966.69 MB.

Non Heap Memory used 48.74 MB of 48.81 MB Committed Non Heap Memory. Max Non Heap Memory is 130 MB.

|                      |                  |
|----------------------|------------------|
| Configured Capacity: | 54.51 GB         |
| DFS Used:            | 837.43 MB (1.5%) |
| Non DFS Used:        | 8.13 GB          |

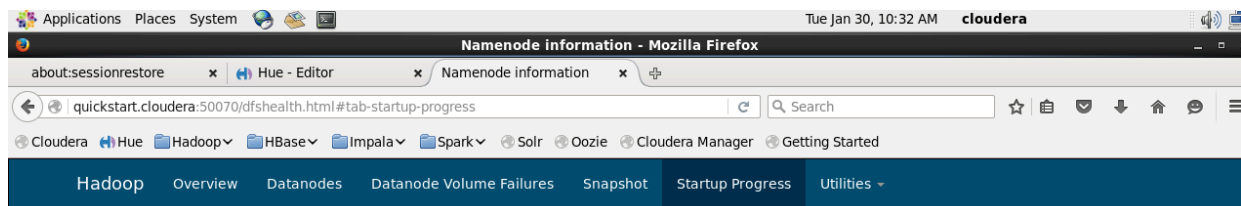
## DATANODE

Pinchando sobre alguno de los Datanodes que aparecen en el listado, se puede ver el detalle sobre los bloques que contiene, la capacidad, la ruta de HDFS en el disco local, etc.



## STARTUP PROGRESS

Muestra cómo ha sido el proceso de arranque de HDFS y el estado en el que se encuentra ahora:



### Startup Progress

Elapsed Time: 1 sec, Percent Complete: 100%

| Phase  | Completion | Elapsed Time |
|--|------------|--------------|
| <b>Loading fsimage</b> /var/lib/hadoop-hdfs/cache/hdfs/dfs/name/current/fsimage_0000000000000015290 86.75 KB | 100%       | 0 sec        |
| inodes (1077/1077)   | 100%       |              |
| delegation tokens (0/0)  | 100%       |              |
| cache pools (0/0)  | 100%       |              |
| <b>Loading edits</b>   | 100%       | 0 sec        |
| /var/lib/hadoop-hdfs/cache/hdfs/dfs/name/current/edits_0000000000000015291-0000000000000015372 1 MB (82/82)  | 100%       |              |
| <b>Saving checkpoint</b>   | 100%       | 0 sec        |
| inodes /var/lib/hadoop-hdfs/cache/hdfs/dfs/name/current/fsimage.ckpt_0000000000000015372 (0/0)               | 100%       |              |
| delegation tokens /var/lib/hadoop-hdfs/cache/hdfs/dfs/name/current/fsimage.ckpt_0000000000000015372 (0/0)    | 100%       |              |
| cache pools /var/lib/hadoop-hdfs/cache/hdfs/dfs/name/current/fsimage.ckpt_0000000000000015372 (0/0)          | 100%       |              |
| <b>Safe mode</b>   | 100%       | 0 sec        |
| awaiting reported blocks (946/947)   | 100%       |              |

Hadoop, 2017.

Legacy  
UI

UTILIDADES

VISUALIZAR EL CONTENIDO DE HDFS

Browsing HDFS - Mozilla Firefox

about:sessionrestore x Hue - Editor x Browsing HDFS x

quickstart.cloudera:50070/explorer.html#/

Search

Cloudera Hue Hadoop HBase Impala Spark Solr Oozie Cloudera Manager Getting Started

Hadoop Overview Datanodes Snapshot Startup Progress Utilities

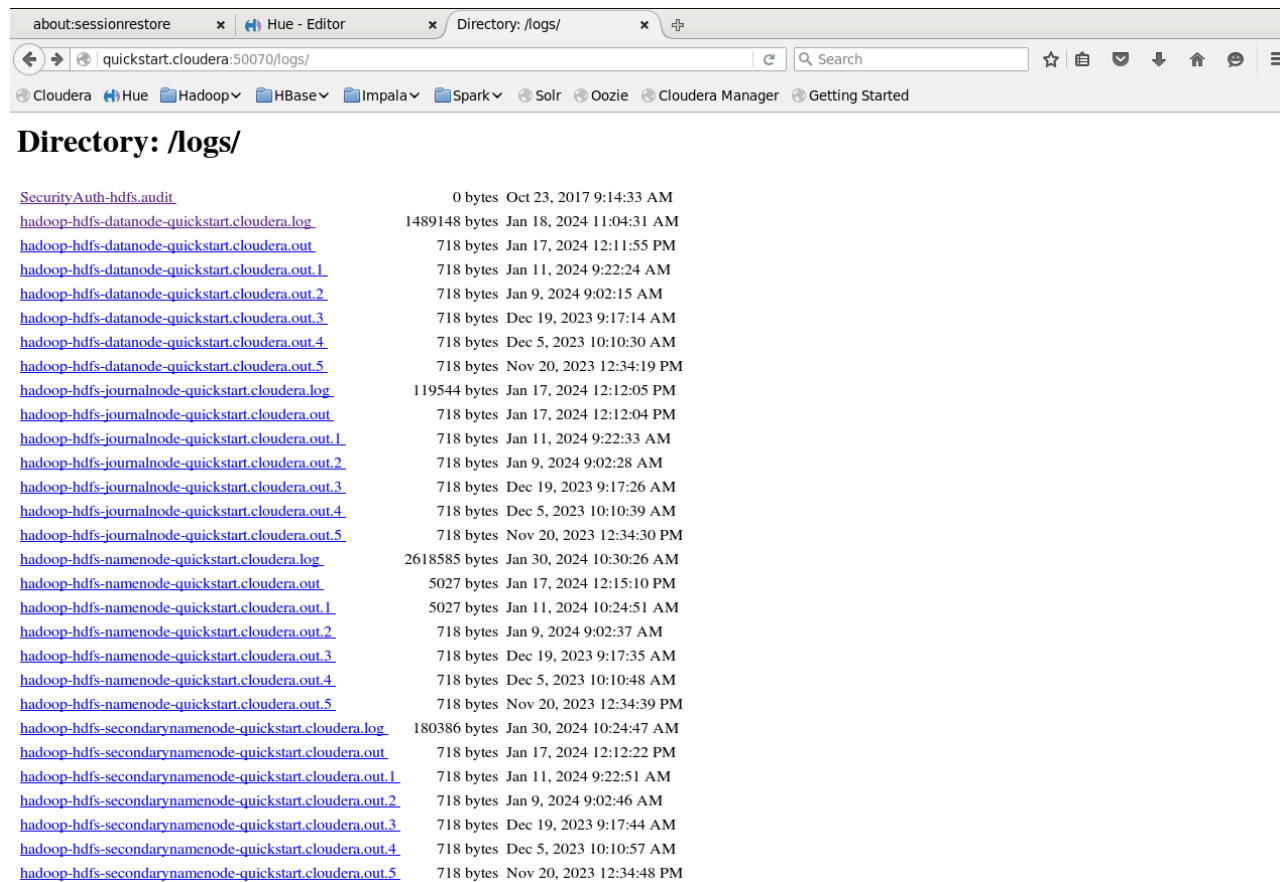
### Browse Directory

/ Go!

| Permission | Owner    | Group      | Size    | Last Modified                  | Replication | Block Size | Name                       |
|------------|----------|------------|---------|--------------------------------|-------------|------------|----------------------------|
| drwxrwxrwx | hdfs     | supergroup | 0 B     | Mon Oct 23 09:15:43 -0700 2017 | 0           | 0 B        | <a href="#">benchmarks</a> |
| -rw-r--r-- | cloudera | supergroup | 0 B     | Thu Nov 23 09:03:01 -0800 2023 | 1           | 128 MB     | <a href="#">data</a>       |
| drwxr-xr-x | cloudera | supergroup | 0 B     | Tue Jan 09 10:53:18 -0800 2024 | 0           | 0 B        | <a href="#">flume</a>      |
| drwxr-xr-x | hbase    | supergroup | 0 B     | Wed Jan 17 12:13:07 -0800 2024 | 0           | 0 B        | <a href="#">hbase</a>      |
| -rw-r--r-- | cloudera | supergroup | 1.61 MB | Thu Jan 11 10:32:58 -0800 2024 | 1           | 128 MB     | <a href="#">pig</a>        |
| drwxr-xr-x | solr     | solr       | 0 B     | Mon Oct 23 09:18:01 -0700 2017 | 0           | 0 B        | <a href="#">solr</a>       |
| drwxrwxrwt | hdfs     | supergroup | 0 B     | Tue Nov 21 09:00:36 -0800 2023 | 0           | 0 B        | <a href="#">tmp</a>        |
| drwxr-xr-x | hdfs     | supergroup | 0 B     | Mon Oct 23 09:17:33 -0700 2017 | 0           | 0 B        | <a href="#">user</a>       |
| drwxr-xr-x | hdfs     | supergroup | 0 B     | Mon Oct 23 09:17:24 -0700 2017 | 0           | 0 B        | <a href="#">var</a>        |

Hadoop, 2017.

## LOGS



The screenshot shows the Hue web interface with the directory `/logs/` selected. The browser tabs include `about:sessionrestore`, `Hue - Editor`, and `Directory: /logs/`. The address bar shows `quickstart.cloudera:50070/logs/`. The navigation bar includes links for Cloudera, Hue, Hadoop, HBase, Impala, Spark, Solr, Oozie, Cloudera Manager, and Getting Started.

**Directory: /logs/**

| File Name   | Size          | Timestamp                |
|---|---------------|--------------------------|
| <a href="#">SecurityAuth-hdfs.audit</a>                                 | 0 bytes       | Oct 23, 2017 9:14:33 AM  |
| <a href="#">hadoop-hdfs-datanode-quickstart.cloudera.log</a>            | 1489148 bytes | Jan 18, 2024 11:04:31 AM |
| <a href="#">hadoop-hdfs-datanode-quickstart.cloudera.out</a>            | 718 bytes     | Jan 17, 2024 12:11:55 PM |
| <a href="#">hadoop-hdfs-datanode-quickstart.cloudera.out.1</a>          | 718 bytes     | Jan 11, 2024 9:22:24 AM  |
| <a href="#">hadoop-hdfs-datanode-quickstart.cloudera.out.2</a>          | 718 bytes     | Jan 9, 2024 9:02:15 AM   |
| <a href="#">hadoop-hdfs-datanode-quickstart.cloudera.out.3</a>          | 718 bytes     | Dec 19, 2023 9:17:14 AM  |
| <a href="#">hadoop-hdfs-datanode-quickstart.cloudera.out.4</a>          | 718 bytes     | Dec 5, 2023 10:10:30 AM  |
| <a href="#">hadoop-hdfs-datanode-quickstart.cloudera.out.5</a>          | 718 bytes     | Nov 20, 2023 12:34:19 PM |
| <a href="#">hadoop-hdfs-journalnode-quickstart.cloudera.log</a>         | 119544 bytes  | Jan 17, 2024 12:12:05 PM |
| <a href="#">hadoop-hdfs-journalnode-quickstart.cloudera.out</a>         | 718 bytes     | Jan 17, 2024 12:12:04 PM |
| <a href="#">hadoop-hdfs-journalnode-quickstart.cloudera.out.1</a>       | 718 bytes     | Jan 11, 2024 9:22:33 AM  |
| <a href="#">hadoop-hdfs-journalnode-quickstart.cloudera.out.2</a>       | 718 bytes     | Jan 9, 2024 9:02:28 AM   |
| <a href="#">hadoop-hdfs-journalnode-quickstart.cloudera.out.3</a>       | 718 bytes     | Dec 19, 2023 9:17:26 AM  |
| <a href="#">hadoop-hdfs-journalnode-quickstart.cloudera.out.4</a>       | 718 bytes     | Dec 5, 2023 10:10:39 AM  |
| <a href="#">hadoop-hdfs-journalnode-quickstart.cloudera.out.5</a>       | 718 bytes     | Nov 20, 2023 12:34:30 PM |
| <a href="#">hadoop-hdfs-namenode-quickstart.cloudera.log</a>            | 2618585 bytes | Jan 30, 2024 10:30:26 AM |
| <a href="#">hadoop-hdfs-namenode-quickstart.cloudera.out</a>            | 5027 bytes    | Jan 17, 2024 12:15:10 PM |
| <a href="#">hadoop-hdfs-namenode-quickstart.cloudera.out.1</a>          | 5027 bytes    | Jan 11, 2024 10:24:51 AM |
| <a href="#">hadoop-hdfs-namenode-quickstart.cloudera.out.2</a>          | 718 bytes     | Jan 9, 2024 9:02:37 AM   |
| <a href="#">hadoop-hdfs-namenode-quickstart.cloudera.out.3</a>          | 718 bytes     | Dec 19, 2023 9:17:35 AM  |
| <a href="#">hadoop-hdfs-namenode-quickstart.cloudera.out.4</a>          | 718 bytes     | Dec 5, 2023 10:10:48 AM  |
| <a href="#">hadoop-hdfs-namenode-quickstart.cloudera.out.5</a>          | 718 bytes     | Nov 20, 2023 12:34:39 PM |
| <a href="#">hadoop-hdfs-secondarynamenode-quickstart.cloudera.log</a>   | 180386 bytes  | Jan 30, 2024 10:24:47 AM |
| <a href="#">hadoop-hdfs-secondarynamenode-quickstart.cloudera.out</a>   | 718 bytes     | Jan 17, 2024 12:12:22 PM |
| <a href="#">hadoop-hdfs-secondarynamenode-quickstart.cloudera.out.1</a> | 718 bytes     | Jan 11, 2024 9:22:51 AM  |
| <a href="#">hadoop-hdfs-secondarynamenode-quickstart.cloudera.out.2</a> | 718 bytes     | Jan 9, 2024 9:02:46 AM   |
| <a href="#">hadoop-hdfs-secondarynamenode-quickstart.cloudera.out.3</a> | 718 bytes     | Dec 19, 2023 9:17:44 AM  |
| <a href="#">hadoop-hdfs-secondarynamenode-quickstart.cloudera.out.4</a> | 718 bytes     | Dec 5, 2023 10:10:57 AM  |
| <a href="#">hadoop-hdfs-secondarynamenode-quickstart.cloudera.out.5</a> | 718 bytes     | Nov 20, 2023 12:34:48 PM |

No he podido acceder al archivo de configuración del HDFS, ni al volcado de hilos en ejecución.

## ACTIVIDAD 2

### YARN RESOURCES MANAGER

Al acceder a la pantalla principal, se muestra un resumen del sistema:

**Cluster Metrics**

| Apps Submitted | Apps Pending | Apps Running | Apps Completed | Containers Running | Memory Used | Memory Total |
|----------------|--------------|--------------|----------------|--------------------|-------------|--------------|
| 3              | 0            | 0            | 3              | 0                  | 0 B         | 8 GB         |

**Cluster Nodes Metrics**

| Active Nodes | Decommissioning Nodes | Decommissioned Nodes | Lost Nodes |
|--------------|-----------------------|----------------------|------------|
| 1            | 0                     | 0                    | 0          |

**User Metrics for dr.who**

| Apps Submitted | Apps Pending | Apps Running | Apps Completed | Containers Running | Containers Pending | Containers Reserved | Memory Used |
|----------------|--------------|--------------|----------------|--------------------|--------------------|---------------------|-------------|
| 0              | 0            | 0            | 0              | 0                  | 0                  | 0                   | 0 B         |

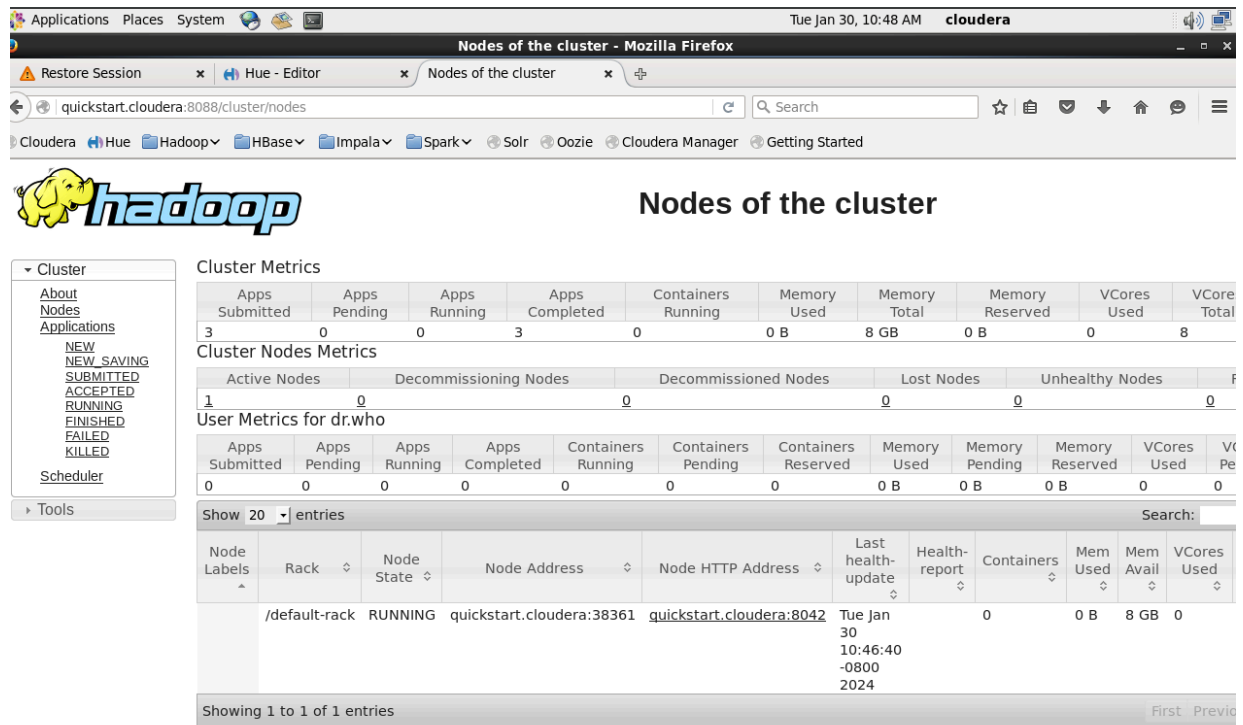
Showing 20 entries

| ID                             | User     | Name  | Application Type | Queue         | StartTime                      | FinishTime                     | State    | FinalStatus |
|--------------------------------|----------|---|------------------|---------------|--------------------------------|--------------------------------|----------|-------------|
| application_1705522371168_0003 | cloudera | -- Ejecutar una consulta en Impala SELE...75(Stage-1) | MAPREDUCE        | root.cloudera | Wed Jan 17 12:33:03 -0800 2024 | Wed Jan 17 12:33:27 -0800 2024 | FINISHED | SUCCEEDED   |
| application_1705522371168_0002 | cloudera | -- Ejecutar una consulta en Impala SELE...75(Stage-1) | MAPREDUCE        | root.cloudera | Wed Jan 17 12:29:54 -0800 2024 | Wed Jan 17 12:30:16 -0800 2024 | FINISHED | SUCCEEDED   |
| application_1705522371168_0001 | cloudera | -- Insertar algunos registros en la ta...78)(Stage-1) | MAPREDUCE        | root.cloudera | Wed Jan 17 12:20:30 -0800 2024 | Wed Jan 17 12:20:47 -0800 2024 | FINISHED | SUCCEEDED   |

Showing 1 to 3 of 3 entries

## NODES

En el menú de la izquierda se tienen diferentes opciones, por ejemplo, al pinchar en la opción Nodes se puede ver la información de los nodos worker, pudiendo comprobar cuántos hay, en qué estado se encuentran, la dirección interna del nodo dentro del clúster así como los recursos utilizados y disponibles.



The screenshot shows the Cloudera Hue interface in a Mozilla Firefox browser window. The page title is "Nodes of the cluster". The left sidebar contains a menu with options: Cluster, About, Nodes, Applications, NEW, NEW SAVING, SUBMITTED, ACCEPTED, RUNNING, FINISHED, FAILED, KILLED, Scheduler, and Tools. The main content area displays several tables:

**Cluster Metrics**

| Apps Submitted | Apps Pending | Apps Running | Apps Completed | Containers Running | Memory Used | Memory Total | Memory Reserved | VCores Used | VCores Total |
|----------------|--------------|--------------|----------------|--------------------|-------------|--------------|-----------------|-------------|--------------|
| 3              | 0            | 0            | 3              | 0                  | 0 B         | 8 GB         | 0 B             | 0           | 8            |

**Cluster Nodes Metrics**

| Active Nodes | Decommissioning Nodes | Decommissioned Nodes | Lost Nodes | Unhealthy Nodes |
|--------------|-----------------------|----------------------|------------|-----------------|
| 1            | 0                     | 0                    | 0          | 0               |

**User Metrics for dr.who**

| Apps Submitted | Apps Pending | Apps Running | Apps Completed | Containers Running | Containers Pending | Containers Reserved | Memory Used | Memory Pending | Memory Reserved | VCores Used | VCores Pending |
|----------------|--------------|--------------|----------------|--------------------|--------------------|---------------------|-------------|----------------|-----------------|-------------|----------------|
| 0              | 0            | 0            | 0              | 0                  | 0                  | 0                   | 0 B         | 0 B            | 0 B             | 0           | 0              |

**Node List**

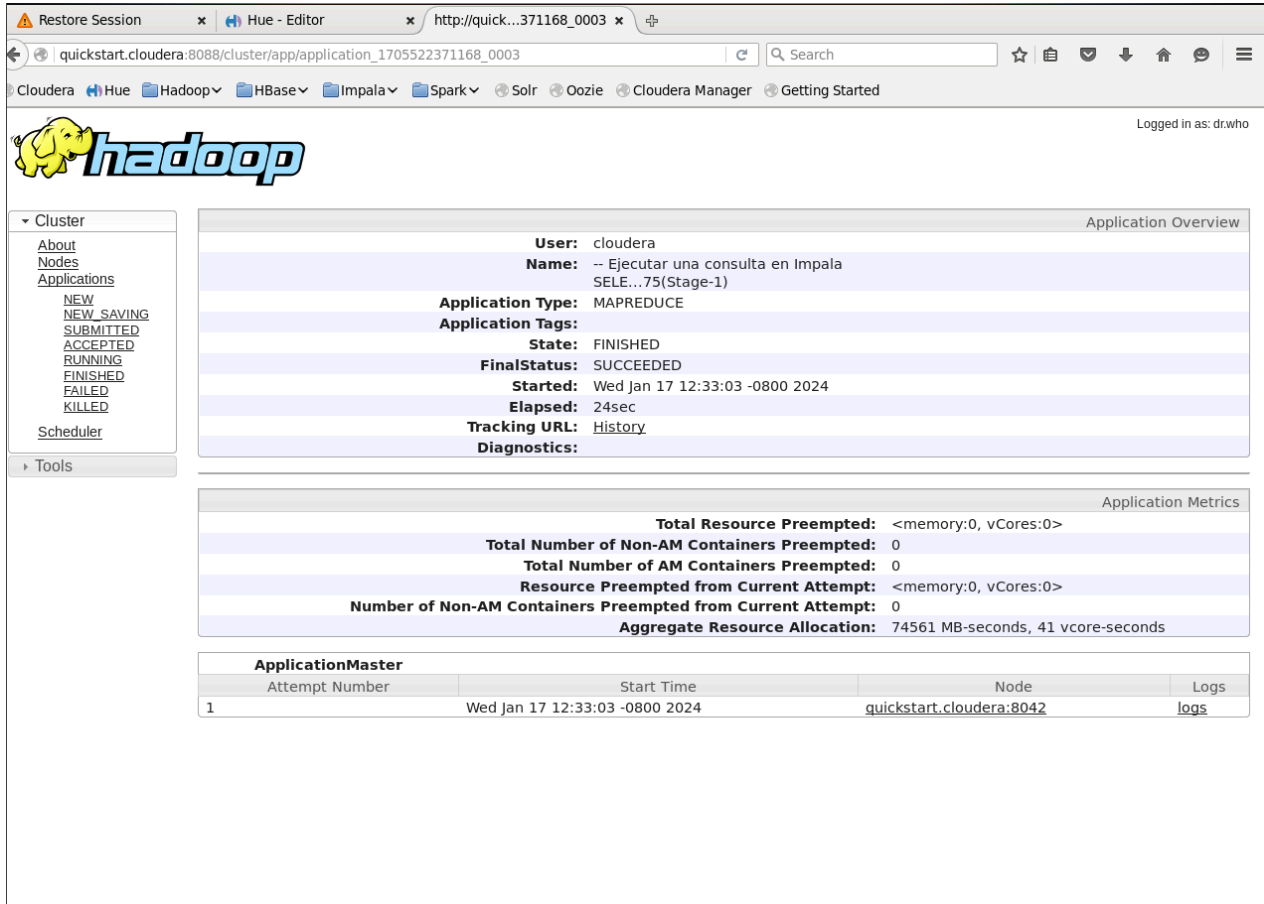
| Node Labels   | Rack    | Node State                | Node Address             | Node HTTP Address              | Last health-update | Health-report | Containers | Mem Used | Mem Avail | VCores Used |
|---------------|---------|---------------------------|--------------------------|--------------------------------|--------------------|---------------|------------|----------|-----------|-------------|
| /default-rack | RUNNING | quickstart.cloudera:38361 | quickstart.cloudera:8042 | Tue Jan 30 10:46:40 -0800 2024 | 0                  | 0 B           | 8 GB       | 0        |           |             |

Showing 1 to 1 of 1 entries



## APPLICATION

Haciendo clic sobre una aplicación, se puede consultar su detalle de ejecución. Por ejemplo, entrando en el detalle de la segunda aplicación ejecutada, se obtiene la siguiente información:



The screenshot shows the Hue web interface with the Hadoop logo and navigation menu on the left. The main content area displays the details of a specific application. The 'Application Overview' section provides key information about the job, including its user, name, type, state, and completion time. The 'Application Metrics' section shows resource usage statistics. The 'ApplicationMaster' table lists the attempt details.

**Application Overview**

|                          |  |
|--------------------------|--|
| <b>User:</b>             | cloudera   |
| <b>Name:</b>             | -- Ejecutar una consulta en Impala<br>SELE...75(Stage-1) |
| <b>Application Type:</b> | MAPREDUCE  |
| <b>Application Tags:</b> |  |
| <b>State:</b>            | FINISHED   |
| <b>FinalStatus:</b>      | SUCCEEDED  |
| <b>Started:</b>          | Wed Jan 17 12:33:03 -0800 2024                           |
| <b>Elapsed:</b>          | 24sec  |
| <b>Tracking URL:</b>     | <a href="#">History</a>                                  |
| <b>Diagnostics:</b>      |  |

**Application Metrics**

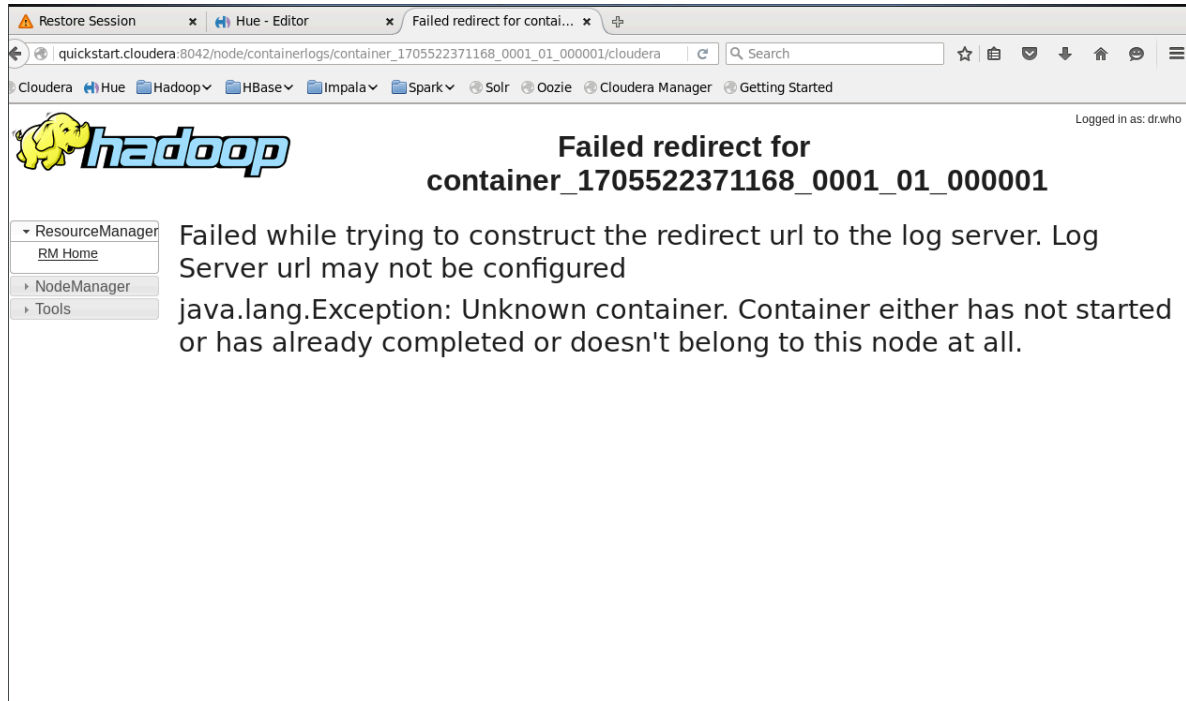
|  |                                    |
|--|------------------------------------|
| <b>Total Resource Preempted:</b>                                   | <memory:0, vCores:0>               |
| <b>Total Number of Non-AM Containers Preempted:</b>                | 0                                  |
| <b>Total Number of AM Containers Preempted:</b>                    | 0                                  |
| <b>Resource Preempted from Current Attempt:</b>                    | <memory:0, vCores:0>               |
| <b>Number of Non-AM Containers Preempted from Current Attempt:</b> | 0                                  |
| <b>Aggregate Resource Allocation:</b>                              | 74561 MB-seconds, 41 vcore-seconds |

**ApplicationMaster**

| Attempt Number | Start Time                     | Node                                     | Logs                 |
|----------------|--------------------------------|--|----------------------|
| 1              | Wed Jan 17 12:33:03 -0800 2024 | <a href="#">quickstart.cloudera:8042</a> | <a href="#">logs</a> |

## LOG

No puedo acceder al log:



The screenshot shows the Hue web interface in a browser. The address bar displays the URL: `quickstart.cloudera:8042/node/containerlogs/container_1705522371168_0001_01_000001/cloudera`. The page header includes the Cloudera logo and navigation links for Hue, Hadoop, HBase, Impala, Spark, Solr, Oozie, Cloudera Manager, and Getting Started. The main content area displays a red error message: "Failed redirect for container\_1705522371168\_0001\_01\_000001". Below the error message, a sidebar on the left shows a menu with "ResourceManager" (expanded), "RM Home", "NodeManager", and "Tools". The main text area contains the following error details:

**Failed redirect for  
container\_1705522371168\_0001\_01\_000001**

Failed while trying to construct the redirect url to the log server. Log Server url may not be configured

java.lang.Exception: Unknown container. Container either has not started or has already completed or doesn't belong to this node at all.

## SCHEDULER



NEW,NEW\_SAVING,SUBMITTED,ACCEPTED,RUNNING Applications

Cluster
About Nodes
Applications
NEW
NEW\_SAVING
SUBMITTED
ACCEPTED
RUNNING
FINISHED
FAILED
KILLED
Scheduler
Tools

### Cluster Metrics

| Apps Submitted | Apps Pending | Apps Running | Apps Completed | Containers Running | Memory Used | Memory Total | Memory Reserved | VCores Used |
|----------------|--------------|--------------|----------------|--------------------|-------------|--------------|-----------------|-------------|
| 3              | 0            | 0            | 3              | 0                  | 0 B         | 8 GB         | 0 B             | 0           |

### Cluster Nodes Metrics

| Active Nodes | Decommissioning Nodes | Decommissioned Nodes | Lost Nodes | Unhealthy Nodes |
|--------------|-----------------------|----------------------|------------|-----------------|
| 1            | 0                     | 0                    | 0          | 0               |

### User Metrics for dr.who

| Apps Submitted | Apps Pending | Apps Running | Apps Completed | Containers Running | Containers Pending | Containers Reserved | Memory Used | Memory Pending | Memory Reserved | VCores Used |
|----------------|--------------|--------------|----------------|--------------------|--------------------|---------------------|-------------|----------------|-----------------|-------------|
| 0              | 0            | 0            | 0              | 0                  | 0                  | 0                   | 0 B         | 0 B            | 0 B             | 0           |

### Application Queues

Legend: Steady Fair Share Instantaneous Fair Share Used Used (over fair share) Max Capacity

- root
  - root.cloudera
  - root.default

Show 20 entries

| ID                         | User | Name | Application Type | Queue | Fair Share | StartTime | FinishTime | State | FinalStatus | Running Containers | Allocated CPU VCoers | Allocated Memory MB | Reserved CPU VCoers | Reserved Memory MB |
|----------------------------|------|------|------------------|-------|------------|-----------|------------|-------|-------------|--------------------|----------------------|---------------------|---------------------|--------------------|
| No data available in table |      |      |                  |       |            |           |            |       |             |                    |                      |                     |                     |                    |

Showing 0 to 0 of 0 entries

### Application Queues

Legend: Steady Fair Share Instantaneous Fair Share Used Used (over fair share) Max Capacity

- root
  - root.cloudera

Used Resources: <memory:0, vCores:0>  
Demand Resources: <memory:0, vCores:0>  
AM Used Resources: <memory:0, vCores:0>  
AM Max Resources: <memory:4096, vCores:4>  
Num Active Applications: 0  
Num Pending Applications: 0  
Min Resources: <memory:0, vCores:0>  
Max Resources: <memory:8192, vCores:8>  
Reserved Resources: <memory:0, vCores:0>  
Steady Fair Share: <memory:4096, vCores:0>  
Instantaneous Fair Share: <memory:0, vCores:0>  
Preemptable: true
  - root.default

Used Resources: <memory:0, vCores:0>  
Demand Resources: <memory:0, vCores:0>  
AM Used Resources: <memory:0, vCores:0>  
AM Max Resources: <memory:0, vCores:0>  
Num Active Applications: 0  
Num Pending Applications: 0  
Min Resources: <memory:0, vCores:0>  
Max Resources: <memory:8192, vCores:8>  
Reserved Resources: <memory:0, vCores:0>  
Steady Fair Share: <memory:4096, vCores:0>  
Instantaneous Fair Share: <memory:0, vCores:0>  
Preemptable: true

No me aparece la opción 'joblauncher'.