

Java-Success.com

Prepare to fast-track, choose & go places with 800+ Java & Big Data Q&As with lots of code & diagrams.

[Home](#) [Why? ▾](#) [300+ Java FAQs ▾](#) [300+ Big Data FAQs ▾](#) [Courses ▾](#)

[Membership ▾](#) [Your Career ▾](#)

[Home](#) > [bigdata-success.com](#) > [Tutorials - Big Data](#) > [TUT - Cloudera](#) > Before running
a Spark job on a YARN cluster in Cloudera and about the Spark history server

Before running a Spark job on a YARN cluster in Cloudera and about the Spark history server

 Posted on [November 18, 2017](#)

Problem: When you run a Spark job via “spark-submit” command on a “YARN” cluster as shown below in a terminal,

```
1  
2 -bash-4.1$ spark-submit --class com.mytutorial.Spa  
3
```

300+ Java Interview FAQs

300+ Java FAQs



16+ Java Key Areas Q&As



150+ Java Architect FAQs



80+ Java Code Quality Q&As



150+ Java Coding Q&As



300+ Big Data Interview FAQs

300+ Big Data FAQs



Tutorials - Big Data



TUT -  Starting Big Data

TUT - Starting Spark & Scala

It creates a folder and files in HDFS under

```
1
2 /user/spark/applicationHistory/
3
```

The files will be named with “application ids” like “application_1510971789108_0015”. This will cause a permission issue, and the history server running at “<http://quickstart.cloudera:18088/>” will not be able to read the files in the folder “/user/spark/applicationHistory/”.

Setting the HDFS permissions to enable viewing via Spark History server

Step 1: Before you can run the Spark history server, you must create the “/user/spark/applicationHistory/” directory in HDFS and set ownership and permissions as follows:

```
1
2 $ sudo -u hdfs hadoop fs -mkdir /user/spark
3 $ sudo -u hdfs hadoop fs -mkdir /user/spark/applicationHistory/
4 $ sudo -u hdfs hadoop fs -chown -R spark:spark /user/spark/applicationHistory/
5 $ sudo -u hdfs hadoop fs -chmod 1777 /user/spark/applicationHistory/
6
```

Now, if you run the Spark job on a YARN cluster in Cloudera

```
1
2 -bash-4.1$ spark-submit --class com.mytutorial.SparkHistoryServer
3
```

You can see the following created in “/user/spark/applicationHistory/” with the

TUT - Starting with Python

TUT - Kafka

TUT - Pig

TUT - Apache Storm

TUT - Spark Scala on Zeppelin

TUT - Cloudera

TUT - Cloudera on Docker

TUT - File Formats

TUT - Spark on Docker

TUT - Flume

TUT - Hadoop (HDFS)

TUT - HBase (NoSQL)

TUT - Hive (SQL)

TUT - Hadoop & Spark

TUT - MapReduce

TUT - Spark and Scala

TUT - Spark & Java

TUT - PySpark on Databricks

TUT - Zookeeper

800+ Java Interview Q&As

300+ Core Java Q&As



300+ Enterprise Java Q&As



150+ Java Frameworks Q&As



120+ Companion Tech Q&As



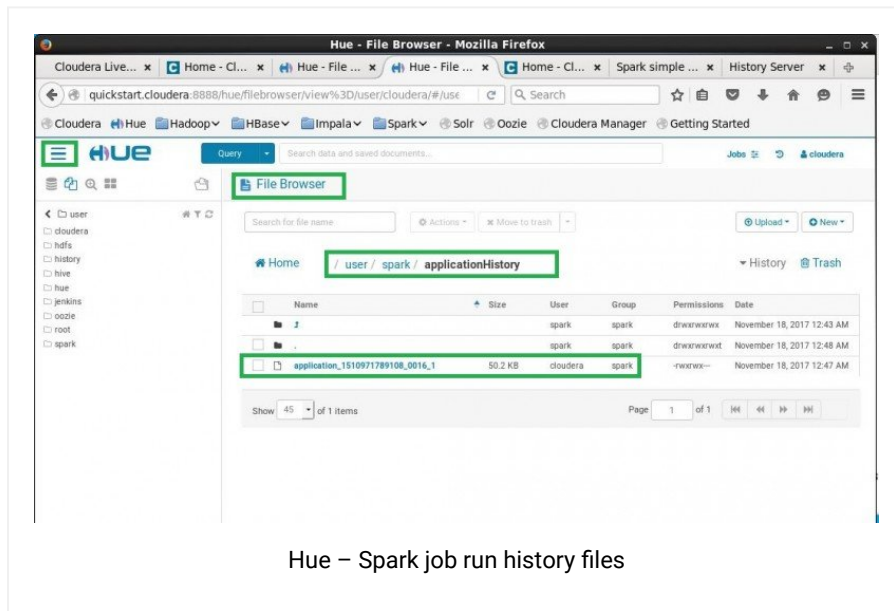
Tutorials - Enterprise Java



application id "application_1510971789108_0016"

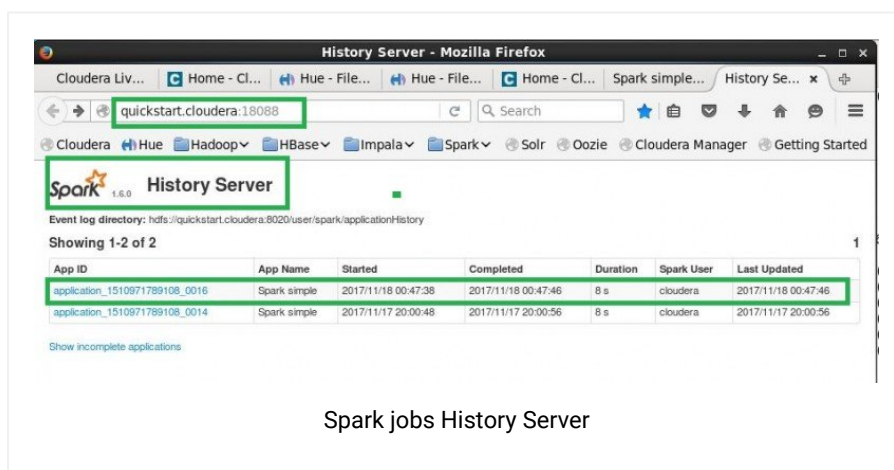
Spark run creates a file in HDFS

The file is created with the Spark run application id as depicted below.



Spark history server shows the job history

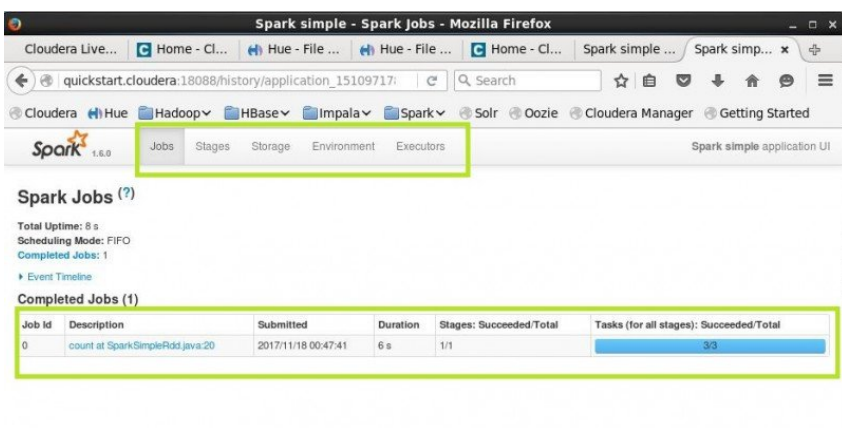
The history server can be accessed via the URL <http://quickstart.cloudera:18088/>. You can also get to this via the "Cloudera Manager" Web UI by selecting "Spark" service -> "History Server Web UI".



Drill into each job via Spark UI

to look at the logs, memory usage, stage failures, etc.

Spark UI is a great tool get more insights to performance tune your Spark jobs.



Spark UI to analyze a job

◀ Understanding Cloudera Hadoop users

09: Running a Spark job on YARN cluster in Cloudera ▶

Disclaimer

The contents in this Java-Success are copyrighted and from EmpoweringTech pty Ltd. The EmpoweringTech pty Ltd has the right to correct or enhance the current content without any prior notice. These are general advice only, and one needs to take his/her own circumstances into consideration. The EmpoweringTech pty Ltd will not be held liable for any damages caused or alleged to be caused either directly or indirectly by these materials and resources. Any trademarked names or labels used in this blog remain the property of their respective trademark owners. Links to external sites do not imply endorsement of the linked-to sites. [Privacy Policy](#)