



# Big Data Hadoop and Spark Developer

[Lesson-End Project Solution](#)



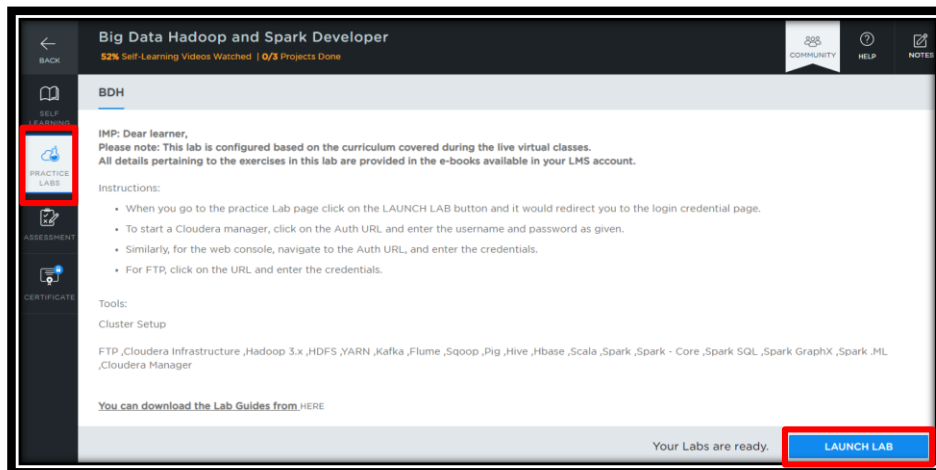
## Working with Yarn

### Steps to Perform:

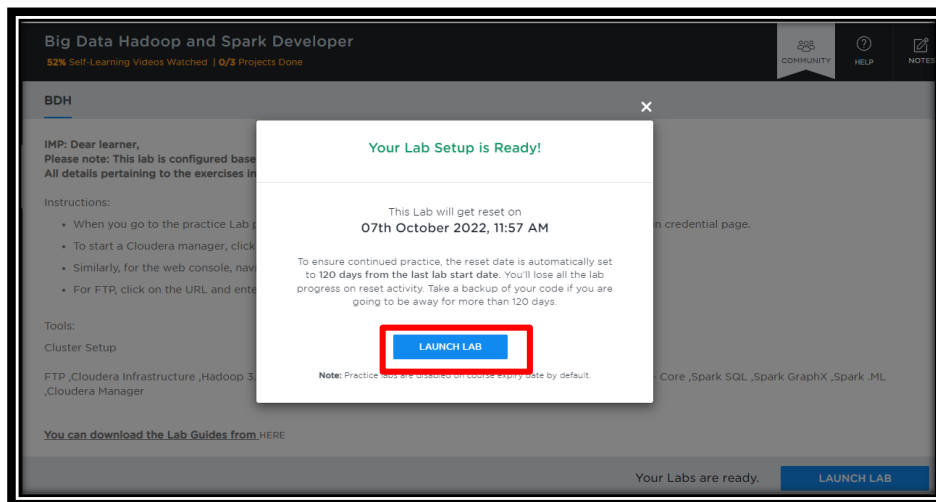
#### Step 1: Log in to the “Web console”

1.1 Open the course “**Big Data Hadoop and Spark Developer**”

1.2 On the left side, click on the “**PRACTICE LABS**” tab and click on the “**LAUNCH LAB**” button



1.3 Again, click on the “**LAUNCH LAB**” button



#### 1.4 Login to the “Web Console” lab

##### Step 2: List the yarn applications

The application list can help us fetch the following details:

- 2.1 Application Id for each application
- 2.2 User who has submitted the job
- 2.3 Queue the job has been submitted too

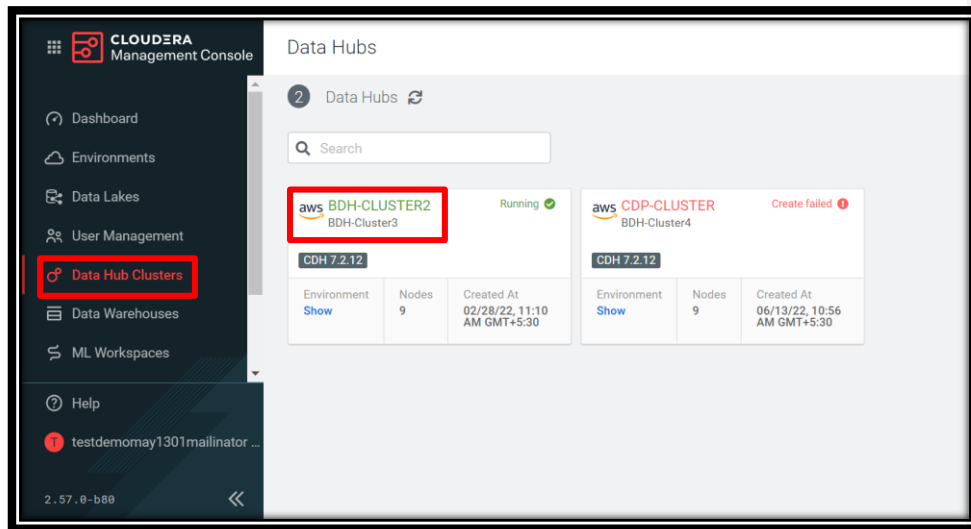
##### Command:

yarn application -list

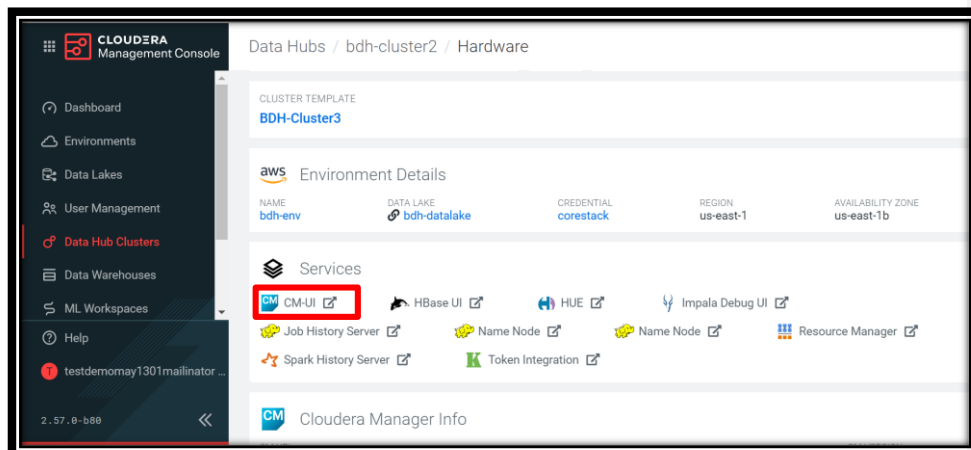
```
[testdemomay1301mailinator@bdh-cluster2-edgenode10 ~]$ yarn application -list
WARNING: YARN_OPTS has been replaced by HADOOP_OPTS. Using value of YARN_OPTS.
Total number of applications (application-types: [], states: [SUBMITTED, ACCEPTED, RUNNING] and tags: []):0
  Final-State   Progress   Application-Id   Application-Name   Application-Type   User   Queue   State
[testdemomay1301mailinator@bdh-cluster2-edgenode10 ~]$
```

##### Step 3: Navigate to Yarn using the Cloudera manager

3.1 Open the Cloudera Manager Console and click on “Data Hub Cluster”





3.2 Click on “BDH-CLUSTER2” and then click on “CM-UI”




3.3 Click on “bdh-cluster2”

Home [Switch to Classic View](#)


Status All Health Issues Configuration  All Recent Commands

Last Updated: Jun 13, 8:37:50 AM UTC 

### Compute Clusters

Status ↑	Name	Linked Data Context	Issues	Hosts	Services	Version
✓	<b>bdh-cluster2</b>	bdh-datalake (bdh-datala...		9 Hosts	15 Services	7.2.12 (Parcels)

### Other

Status	Name	Issues
✓	MGMT	

1 - 1 of 1


3.4 Scroll down and click on “yarn”

**CLOUDERA Manager**

Search

- Clusters
- Hosts
- Diagnostics
- Charts
- Running Commands
- Support
- testdemomay1301mailinator

7.5.2

- ✓ bdh-datalake
- ✓ hbase 
- ✓ hdfs
- ✓ hue
- ✓ impala
- ✓ oozie
- ✓ spark\_on\_yarn
- sqoop
- tez
- ✓ **yarn**

bytes

08:15 08:30

Total Bytes Re... 77.9K/s Total Bytes Tr... 77.5K/s

HDFS IO

bytes / seco...

08:15 08:30

hdfs ns1, Total By... 1b/s hdfs ns1, Total ... 1.9b/s

hdfs, Total Bytes ... 1b/s hdfs, Total Byte... 1.9b/s

Completed Impala Queries

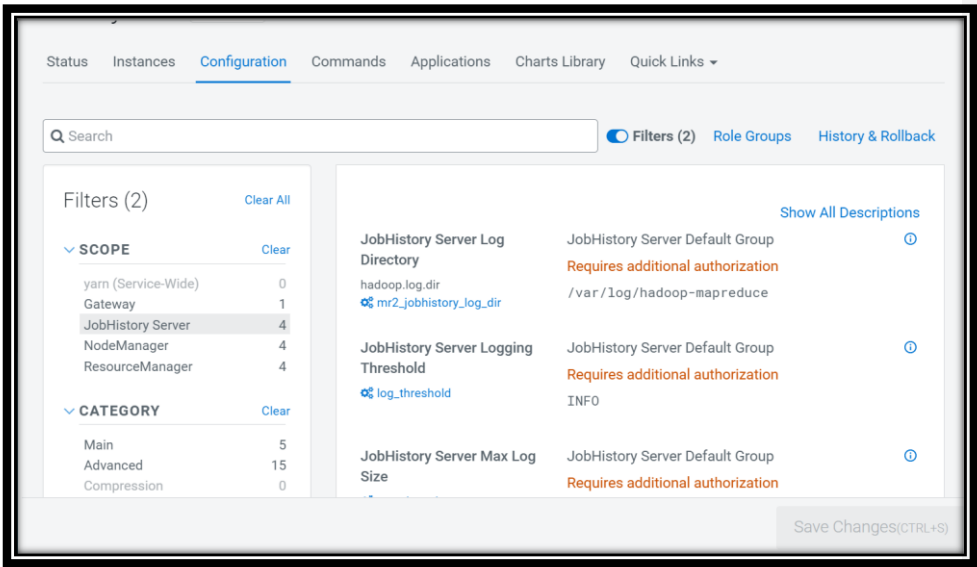
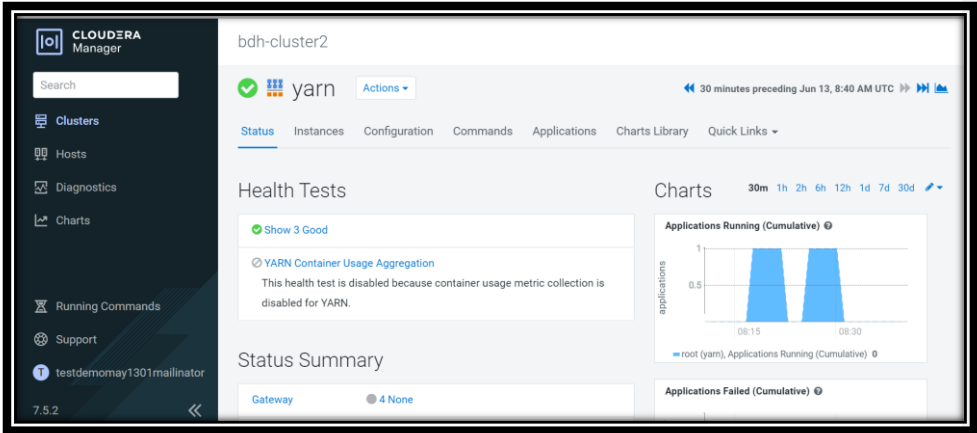
queries / second

08:15 08:30

Impala, Total Queries Across Impala Daemons 0

3.5 From the WebUI, click on “Job History Server”

This gives the list of all jobs history including details like the user, queue, start-end time, and logs.



**Note:** The running applications can be checked from the Applications tab in the Yarn cluster. The list of attributes can be selected as well.

The screenshot shows the Yarn Applications page in Hue. The 'Applications' tab is highlighted with a red box. The page displays a search bar for YARN applications, a 'Workload Summary' section for completed applications, and a table of running applications. The table shows details for a single application: ID: application\_1654575316638\_0456, Type: TEZ, User: hive, Pool: default, Duration: 5.2m, and Allocated Memory Seconds: 953K.

**Note:** Running jobs can be killed from the “Hue” job browser.

The screenshot shows the Hue Job Browser page. The 'Jobs' tab is selected. The page displays a search bar, a filter for 'user:centos', and a table of jobs. The table has columns for Id, Name, User, Type, Status, Progress, Group, Started, and Duration. No jobs are currently listed.

