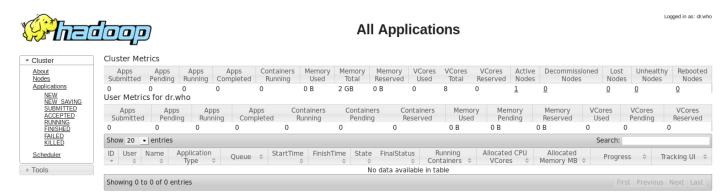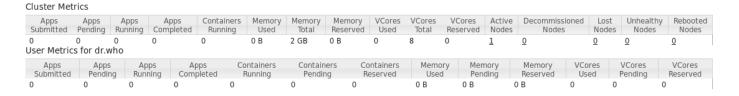BUAN 6346 Big Data

Homework 2

Run a YARN job

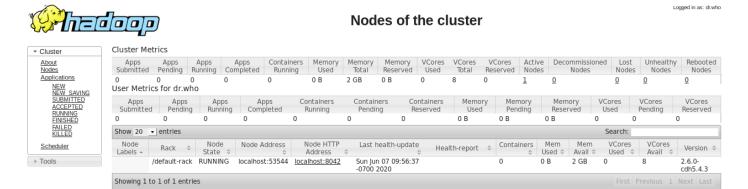**Step 1:** Visit YARN resource Manager(RM)

Step2: Cluster Metrics

Step3: Nodes of the Cluster

Step4: Node Manager UI

Logged in as: dr.who

	NodeManager information
Total Vmem allocated for Containers	4.20 GB
Vmem enforcement enabled	false
Total Pmem allocated for Container	2 GB
Pmem enforcement enabled	true
Total VCores allocated for Containers	8
NodeHealthyStatus	true
LastNodeHealthTime	Sun Jun 07 10:00:37 PDT 2020
NodeHealthReport	
Node Manager Version:	2.6.0-cdh5.4.3 from 4cd9f51a3f1ef748d45b8d77d0f211ad44296d4b by jenkins source checksum 8bacc650a3181bc0cfbae816dc81eddd on 2015-06-25T02:41Z
Hadoop Version:	2.6.0-cdh5.4.3 from 4cd9f51a3f1ef748d45b8d77d0f211ad44296d4b by jenkins source checksum 4acea6ac185376e0b48b33695e88e7a7 on 2015-06-25T02:34Z

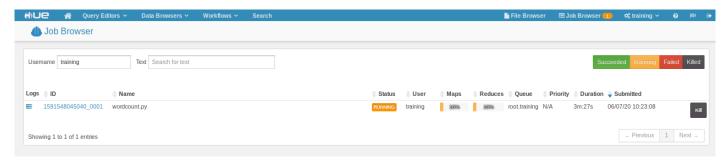## Step5: Back to Resource Manager Home

Step6: Change to exercise directory

```
[training@localhost ~]$ cd $DEV1/exercises/yarn
[training@localhost yarn]$ ls -l
total 4
-rwxr-xr-x 1 training training 1320 Aug 27  2015 wordcount.py
[training@localhost yarn]$ █
```
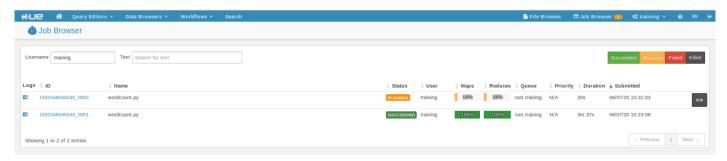
Step7: Run the example wordcount.py program on the YARN cluster to count the frequency of words

```
[training@localhost yarn]$ spark-submit --master yarn-cluster \
> wordcount.py /loudacre/kb/*
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/lib/zookeeper/lib/slf4j-log4j12-1.7.5.jar
!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/lib/flume-ng/lib/slf4j-log4j12-1.7.5.jar!
/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]
20/06/07 10:22:43 WARN util.NativeCodeLoader: Unable to load native-hadoop libra
ry for your platform... using builtin-java classes where applicable
20/06/07 10:22:46 INFO client.RMProxy: Connecting to ResourceManager at /0.0.0.0
:8032
20/06/07 10:22:47 INFO yarn.Client: Requesting a new application from cluster wi
th 1 NodeManagers
20/06/07 10:22:47 INFO yarn.Client: Verifying our application has not requested
more than the maximum memory capability of the cluster (1024 MB per container)
20/06/07 10:22:47 INFO yarn.Client: Will allocate AM container, with 896 MB memo
ry including 384 MB overhead
20/06/07 10:22:47 INFO yarn.Client: Setting up container launch context for our
AM
20/06/07 10:22:47 INFO yarn.Client: Preparing resources for our AM container
20/06/07 10:22:49 WARN shortcircuit.DomainSocketFactory: The short-circuit local
 reads feature cannot be used because libhadoop cannot be loaded.
20/06/07 10:22:49 INFO yarn.Client: Uploading resource file:/usr/lib/spark/lib/s
park-assembly-1.3.0-cdh5.4.3-hadoop2.6.0-cdh5.4.3.jar -> hdfs://localhost:8020/u
ser/training/.sparkStaging/application_1591548045040_0001/spark-assembly-1.3.0-c
dh5.4.3-hadoop2.6.0-cdh5.4.3.jar
█

20/06/07 10:24:15 INFO yarn.Client:
        client token: N/A
        diagnostics: N/A
        ApplicationMaster host: 192.168.206.128
        ApplicationMaster RPC port: 0
        queue: root.training
        start time: 1591550588022
        final status: UNDEFINED
        tracking URL: http://localhost:8088/proxy/application_1591548045040_0001/
        user: training
20/06/07 10:24:16 INFO yarn.Client: Application report for application_1591548045040_0001 (state: RUNNING)
20/06/07 10:24:17 INFO yarn.Client: Application report for application_1591548045040_0001 (state: RUNNING)
20/06/07 10:24:18 INFO yarn.Client: Application report for application_1591548045040_0001 (state: RUNNING)
```
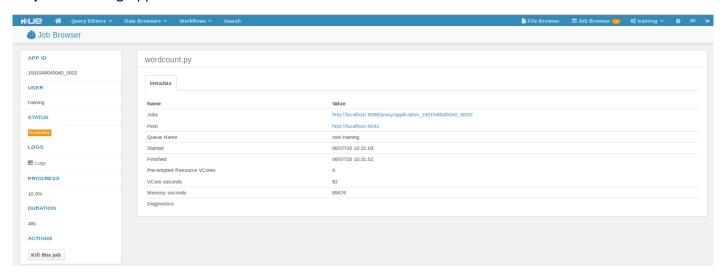
Step8: Job Browser

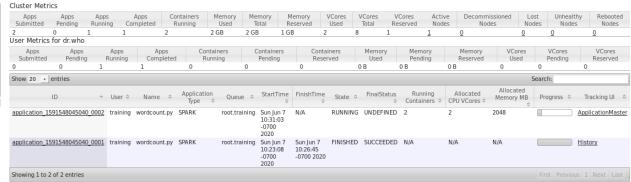**Step9:** Job Browser

Step 10: Running Application

Step 11: YARN RM

Step12: Nodes

Step13: Statistics about node

Step14: List of Applications

Applications running on this node

Logged in as: dr.who

▼ ResourceManager
  RM Home
▸ NodeManager
▸ Tools

Show 20 ▾ entries   Search:

ApplicationId	ApplicationState
application_1591548045040_0002	RUNNING

Showing 1 to 1 of 1 entries   First Previous 1 Next Last

Step15: List of Containers

Logged in as: dr.who

▼ ResourceManager
  RM Home
▸ NodeManager
▸ Tools

Application's information

ApplicationId	application_1591548045040_0002
ApplicationState	RUNNING
User	training

Show 20 ▾ entries   Search:

container_1591548045040_0002_01_000001
container_1591548045040_0002_01_000002

Showing 1 to 2 of 2 entries   First Previous 1 Next Last