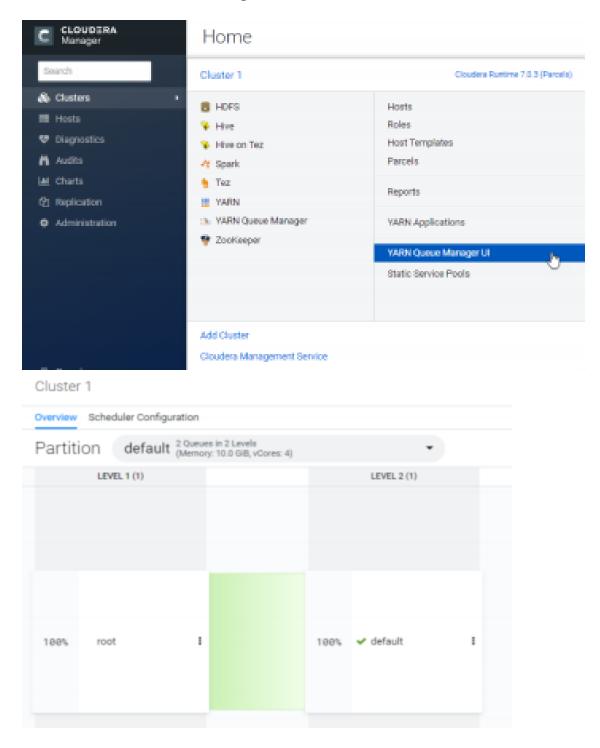
Exercise 03: Working with YARN Containers & Queues



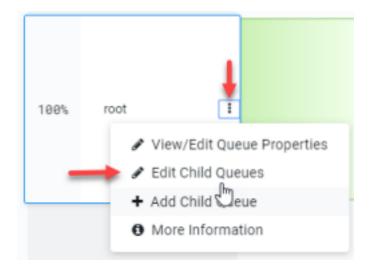
Understanding and Configuring YARN Queues

YARN queue settings and state behavior.

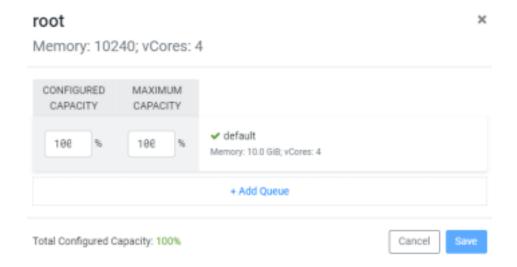
- 1. Open Cloudera Manager http://localhost:7180
- 2. In Cloudera Manager, click Clusters.
- 3. Select the YARN Queue Manager UI.



4. Click on "Edit Child Queues"



5. Click on "Add Queue"



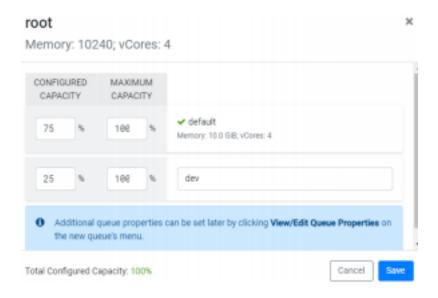
6. In Queue Name. Specify dev as its name Configure the following settings in dev:

Configured Capacity: 25%Maximum Capacity: 100%

In Default Queue

Configure the following settings:

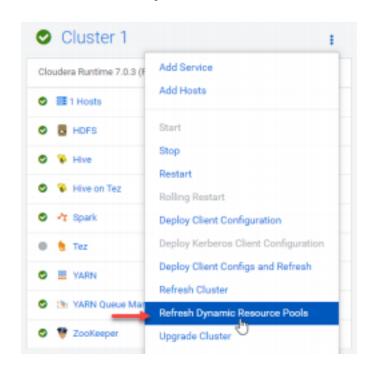
Configured Capacity: 75%Maximum Capacity: 100%



7. Click on save.

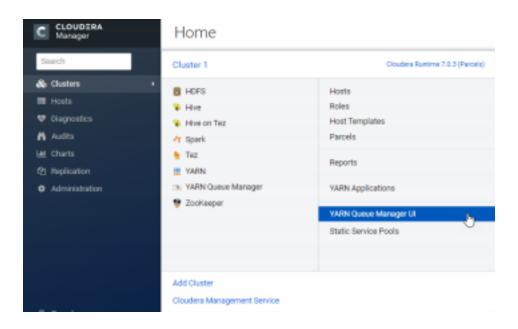
Note: Ignore the error "Failed to add queue"

8. Click on Refresh Dynamic Resource Pools



Note: Ignore warning if there is any.

- 9. In Cloudera Manager, click on Clusters.
- 10. Select the YARN Queue Manager UI.





Submit Spark Applications in different Queues

Launch two instances of wordcount YARN job in two terminal windows.

 Open Terminal, Change to the /opt/cloudera/parcels/SPARK3/lib/spark3/examples/jars directory.

```
cd /opt/cloudera/parcels/SPARK3/lib/spark3/examples/jars
```

2. In terminal window, Import sherlock.txt file into HDFS

```
hdfs dfs -put ~/training/data/sherlock.txt data/
```

3. Type the Spark Submit command to execute wordcount and pass --queue dev.

Terminal

```
spark3-submit --queue dev --class
org.apache.spark.examples.JavaWordCount --master yarn --deploy-mode
cluster spark-examples_2.12-3.0.0.2.99.0.0-23.jar data/sherlock.txt
output
```

End of Exercise