Module 3– Hadoop Cluster: Planning and Managing

Hands On Guide

Version 2.1

Module 3 – Hadoop Cluster: Planning and Managing

Configuring Scheduler and Queue

Table of Contents

1. Allocation

```
<allocations>
<pool name="ads">
<minMaps>10</minMaps>
<minReduces>5</minReduces>
</pool>
</allocations>
```

```
<allocations>
<pool name="ads">
<minMaps>10</minMaps>
<minReduces>5</minReduces>
<maxRunningJobs>3</maxRunningJobs>
</pool>
<user name="matei">
<maxRunningJobs>1</maxRunningJobs>
</user>
</allocations>
```


==========

```
<allocations>

 <!-- Example element for configuring a pool -->
 <pool name="pool1">
  <!-- Minimum shares of map and reduce slots. Defaults to 0. -->
  <minMaps>10</minMaps>
  <minReduces>5</minReduces>

  <!-- Limit on running jobs in the pool. If more jobs are submitted,
only the first <maxRunningJobs> will be scheduled at any given time.
Defaults to infinity or the global poolMaxJobsDefault value below. -->
  <maxRunningJobs>5</maxRunningJobs>

  <!-- Number of seconds after which the pool can preempt other pools'
tasks to achieve its min share. Requires preemption to be enabled in
mapred-site.xml by setting mapred.fairscheduler.preemption to true.
Defaults to infinity (no preemption). -->
  <minSharePreemptionTimeout>300</minSharePreemptionTimeout>

  <!-- Pool's weight in fair sharing calculations. Defaulti is 1.0. -->
```

```
        <weight>1.0</weight>
 </pool>

 <!-- Example element for configuring a user -->
 <user name="user1">
   <!-- Limit on running jobs for the user across all pools. If more
jobs than this are submitted, only the first <maxRunningJobs> will
be scheduled at any given time. Defaults to infinity or the
userMaxJobsDefault value set below. -->
   <maxRunningJobs>10</maxRunningJobs>
 </user>

 <!-- Default running job limit pools where it is not explicitly set. -->
 <poolMaxJobsDefault>20</poolMaxJobsDefault>

 <!-- Default running job limit users where it is not explicitly set. -->
 <userMaxJobsDefault>10</userMaxJobsDefault>

 <!-- Default min share preemption timeout for pools where it is not
explicitly configured, in seconds. Requires mapred.fairscheduler.preemption
to be set to true in your mapred-site.xml. -->
 <defaultMinSharePreemptionTimeout>600</defaultMinSharePreemptionTimeout>

 <!-- Preemption timeout for jobs below their fair share, in seconds.
If a job is below half its fair share for this amount of time, it
is allowed to kill tasks from other jobs to go up to its fair share.
Requires mapred.fairscheduler.preemption to be true in mapred-site.xml. -->
 <fairSharePreemptionTimeout>600</fairSharePreemptionTimeout>

</allocations>
```

2. Capacity

```
<property>
<name>mapred.jobtracker.taskScheduler</name>
<value>org.apache.hadoop.mapred.CapacityTaskScheduler</value>
</property>
```

3. Fair Scheduler

```
<property>
<name>mapred.jobtracker.taskScheduler</name>
<value>org.apache.hadoop.mapred.FairScheduler</value>
</property>

<property>
<name>mapred.fairscheduler.allocation.file</name>
```

```
<value>/home/hadoop/hadoop/conf/fair-scheduler.xml</value>
</property>
```

```
hadoop jar hadoop/hadoop-mapred-examples-0.22.0.jar wordcount -Dpool.name=pool1 /input
/output
```

```
<property>
<name>mapred.jobtracker.taskScheduler</name>
<value>org.apache.hadoop.mapred.FairScheduler</value>
</property>

<property>
<name>mapred.fairscheduler.allocation.file</name>
<value>/home/hadoop/hadoop/conf/fair-scheduler.xml</value>
</property>
```

4. Mapred - site

```
<property>
 <name>mapred.tasktracker.map.tasks.maximum</name>
 <value>8</value>
  <description>the number of available cores on the tasktracker machines
for map tasks
 </description>
</property>
<property>
 <name>mapred.tasktracker.reduce.tasks.maximum</name>
 <value>8</value>
  <description>the number of available cores on the tasktracker machines
for reduce tasks
 </description>
</property>
```

```
<property>
   <name>mapred.cluster.max.map.memory.mb</name>
   <value>4096</value>
 </property>

<property>
```

```
  <name>mapred.cluster.max.reduce.memory.mb</name>
  <value>4096</value>
</property>
```

5. Queue

```
<?xml version="1.0"?>

<configuration>

<!-- system limit, across all queues -->

<property>
 <name>mapred.capacity-scheduler.maximum-system-jobs</name>
 <value>3000</value>
<description>Maximum number of jobs in the system which can be initialized,
concurrently, by the CapacityScheduler.
</description>
</property>

<!-- queue: queueA -->
<property>
 <name>mapred.capacity-scheduler.queue.queueA.capacity</name>
 <value>8</value>
</property>

<property>
 <name>mapred.capacity-scheduler.queue.queueA.supports-priority</name>
 <value>false</value>
</property>
<property>
 <name>mapred.capacity-scheduler.queue.queueA.minimum-user-limit-percent</name>
 <value>20</value>
</property>
<property>
 <name>mapred.capacity-scheduler.queue.queueA.user-limit-factor</name>
 <value>10</value>
</property>
<property>
 <name>mapred.capacity-scheduler.queue.queueA.maximum-initialized-active-tasks</name>
 <value>200000</value>
</property>
<property>
 <name>mapred.capacity-scheduler.queue.queueA.maximum-initialized-active-tasks-per-
user</name>
 <value>100000</value>
</property>
<property>
 <name>mapred.capacity-scheduler.queue.queueA.init-accept-jobs-factor</name>
```

```
 <value>100</value>
</property>

<!-- queue: queueB -->
<property>
 <name>mapred.capacity-scheduler.queue.queueB.capacity</name>
 <value>2</value>
</property>
<property>
 <name>mapred.capacity-scheduler.queue.queueB.supports-priority</name>
 <value>false</value>
</property>
<property>
 <name>mapred.capacity-scheduler.queue.queueB.minimum-user-limit-percent</name>
 <value>20</value>
</property>
<property>
 <name>mapred.capacity-scheduler.queue.queueB.user-limit-factor</name>
 <value>1</value>
</property>
<property>
 <name>mapred.capacity-scheduler.queue.queueB.maximum-initialized-active-tasks</name>
 <value>200000</value>
</property>
<property>
 <name>mapred.capacity-scheduler.queue.queueB.maximum-initialized-active-tasks-per-
user</name>
 <value>100000</value>
</property>
<property>
 <name>mapred.capacity-scheduler.queue.queueB.init-accept-jobs-factor</name>
 <value>10</value>
</property>

<!-- queue: queueC -->
<property>
 <name>mapred.capacity-scheduler.queue.queueC.capacity</name>
 <value>30</value>
</property>
<property>
 <name>mapred.capacity-scheduler.queue.queueC.supports-priority</name>
 <value>false</value>
</property>
<property>
 <name>mapred.capacity-scheduler.queue.queueC.minimum-user-limit-percent</name>
 <value>20</value>
</property>
<property>
 <name>mapred.capacity-scheduler.queue.queueC.user-limit-factor</name>
 <value>1</value>
</property>
```

```
<property>
 <name>mapred.capacity-scheduler.queue.queueC.maximum-initialized-active-tasks</name>
 <value>200000</value>
</property>
<property>
 <name>mapred.capacity-scheduler.queue.queueC.maximum-initialized-active-tasks-per-
user</name>
 <value>100000</value>
</property>
<property>
 <name>mapred.capacity-scheduler.queue.queueC.init-accept-jobs-factor</name>
 <value>10</value>
</property>

<!-- queue: queueD -->
<property>
 <name>mapred.capacity-scheduler.queue.queueD.capacity</name>
 <value>1</value>
</property>
<property>
 <name>mapred.capacity-scheduler.queue.queueD.supports-priority</name>
 <value>false</value>
</property>
<property>
 <name>mapred.capacity-scheduler.queue.queueD.minimum-user-limit-percent</name>
 <value>20</value>
</property>
<property>
 <name>mapred.capacity-scheduler.queue.queueD.user-limit-factor</name>
 <value>20</value>
</property>
<property>
 <name>mapred.capacity-scheduler.queue.queueD.maximum-initialized-active-tasks</name>
 <value>200000</value>
</property>
<property>
 <name>mapred.capacity-scheduler.queue.queueD.maximum-initialized-active-tasks-per-
user</name>
 <value>100000</value>
</property>
<property>
 <name>mapred.capacity-scheduler.queue.queueD.init-accept-jobs-factor</name>
 <value>10</value>
</property>

<!-- queue: queueE -->
<property>
 <name>mapred.capacity-scheduler.queue.queueE.capacity</name>
 <value>31</value>
</property>
<property>
```

```
 <name>mapred.capacity-scheduler.queue.queueE.supports-priority</name>
 <value>false</value>
</property>
<property>
 <name>mapred.capacity-scheduler.queue.queueE.minimum-user-limit-percent</name>
 <value>20</value>
</property>
<property>
 <name>mapred.capacity-scheduler.queue.queueE.user-limit-factor</name>
 <value>1</value>
</property>
<property>
 <name>mapred.capacity-scheduler.queue.queueE.maximum-initialized-active-tasks</name>
 <value>200000</value>
</property>
<property>
 <name>mapred.capacity-scheduler.queue.queueE.maximum-initialized-active-tasks-per-
user</name>
 <value>100000</value>
</property>
<property>
 <name>mapred.capacity-scheduler.queue.queueE.init-accept-jobs-factor</name>
 <value>10</value>
</property>

<!-- queue: queueF -->
<property>
 <name>mapred.capacity-scheduler.queue.queueF.capacity</name>
 <value>28</value>
</property>
<property>
 <name>mapred.capacity-scheduler.queue.queueF.supports-priority</name>
 <value>false</value>
</property>
<property>
 <name>mapred.capacity-scheduler.queue.queueF.minimum-user-limit-percent</name>
 <value>20</value>
</property>
<property>
 <name>mapred.capacity-scheduler.queue.queueF.user-limit-factor</name>
 <value>1</value>
</property>
<property>
 <name>mapred.capacity-scheduler.queue.queueF.maximum-initialized-active-tasks</name>
 <value>200000</value>
</property>
<property>
 <name>mapred.capacity-scheduler.queue.queueF.maximum-initialized-active-tasks-per-
user</name>
 <value>100000</value>
</property>
```

```
<property>
  <name>mapred.capacity-scheduler.queue.queueF.init-accept-jobs-factor</name>
  <value>10</value>
</property>

</configuration>
```