

Extras/Monitoring YARN

Monitoring YARN

READY

This notebook shows you some tricks for monitoring YARN, the resource scheduling system for Hadoop. These can be useful on very small BDCS-CE clusters (especially the 2OCPU OC2M shape) where YARN processes can get stuck waiting for resources.

Here are 3 techniques for monitoring YARN

- Connect to the Resource Manager web UI
- Run the “yarn top” command via SSH
- Run various yarn commands via notebook

Connect to Resource Manager web UI

READY

The Resource Manager web UI runs on port 8088 on the BDCS-CE Master server. You can either open this port via an Access Rule or connect to it via SSH Tunneling.

To do so, follow the instructions in the note “xtra Connecting to Ambari” but use port 8088 instead of 8080 and http instead of https.

Ambari - bdcscce-dcb...


All Applications

+

127.0.0.1:8088/cluster

Search

☆ | 📁 | ⬇️ | 🏠 | 🛡️ | ☰



All Applications

Cluster

About

Nodes

Node Labels

Applications

NEW

NEW SAVING

SUBMITTED

ACCEPTED

RUNNING

FINISHED

FAILED

KILLED

Scheduler

Tools

Cluster Metrics

Apps Submitted	Apps Pending	Apps Running	Apps Completed	Containers Running	Memory Used	Memory Total	Memory Reserved	VCores Used	VCores Total	VCores Reserved	Active Nodes	...
10	0	2	8	3	3 GB	15.45 GB	0 B	3	3	0	1	0

Scheduler Metrics

Scheduler Type	Scheduling Resource Type	Minimum Allocation
Capacity Scheduler	[MEMORY, CPU]	<memory:512, vCores:1>

Show 20 entries

ID	User	Name	Application Type	Queue	StartTime	FinishTime	State	FinalStatus	Running Containers	Allocated CPU VCoers	Alloca Mem MB
application_1503512150236_0010	zeppelin	Zeppelin	SPARK	interactive	Wed Aug 23 16:03:29	N/A	RUNNING	UNDEFINED	2	2	2048

						-0400					
						2017					
application_1503512150236_0009	zeppelin	HIVE-63191098-0263-4bc6-981b-cce70314facf	TEZ	default	Wed Aug 23	Wed Aug 23	FINISHED	SUCCEEDED	N/A	N/A	N/A

Run the “yarn top” command via SSH

READY

Another way you can monitor yarn is via logging into BDCS-CE via SSH and running the “yarn top” command. Refer to the note “xtra Connecting via SSH” to connect via SSH. Then type “yarn top” to launch the tool. Type “h” followed by enter to view help.

```
opc@bdcscs-dcb-aug23-bdcscs-1:~
YARN top - 21:04:54, up 0d, 2:49, 0 active users, queue(s): root
NodeManager(s): 1 total, 1 active, 0 unhealthy, 0 decommissioned, 0 lost, 0 rebooted
Queue(s) Applications: 2 running, 10 submitted, 0 pending, 8 completed, 0 killed, 0 failed
Queue(s) Mem(GB): 12 available, 3 allocated, 0 pending, 0 reserved
Queue(s) Vcores: -1 available, 3 allocated, 0 pending, 0 reserved

APPLICATIONID USER      TYPE      QUEUE  #CONT  #RCONT  VCORES  RVCORES  MEM  RMEM  VCORESECS  MEMSECS  %PROGR  TIME NAME
application_1503512150236_0010 zeppelin  spark interactive  2      0      2      0      2G   0G      7351      7352   10.00   00:01:01 Zeppelin
application_1503512150236_0001 hive      spark  default    1      0      1      0      1G   0G      9976      9976   10.00   00:02:46 Thrift JDBC/ODBC Server
```

Run various yarn commands via notebook

READY

Another way to monitor yarn is via a series of yarn commands run in the notebook shell interpreter. The following paragraphs provide a few examples you can use and modify.

Script to see the Used versus Capacity for YARN CPU and Memory

FINISHED

```
%sh
export YARN_ROOT_LOGGER=WARN
export NODEID=`yarn node -list -all | tail -1 | cut -f1`
echo $NODEID
yarn node -status $NODEID

dcbnov15c-bdcscs-1.compute-gse00002281.oraclecloud.internal:45454
Node Report :
Node-Id : dcbnov15c-bdcscs-1.compute-gse00002281.oraclecloud.internal:45454
Rack : /default-rack
Node-State : RUNNING
Node-Http-Address : dcbnov15c-bdcscs-1.compute-gse00002281.oraclecloud.internal:8042
Last-Health-Update : Thu 16/Nov/17 04:22:56:328UTC
Health-Report :
Containers : 2
Memory-Used : 2048MB
Memory-Capacity : 15821MB
CPU-Used : 2 vcores
CPU-Capacity : 3 vcores
```

Node-Labels :

Took 5 sec. Last updated by anonymous at November 16 2017, 11:23:28 AM. (outdated)

Script to see the current YARN applications and State

FINISHED

```
%sh
export YARN_ROOT_LOGGER=WARN
yarn application -list
```

Total number of applications (application-types: [] and states: [SUBMITTED, ACCEPTED, RUNNING]):2

Application-Id	Application-Name	Application-Type	User	Queue	State	Final-State	Progress
Tracking-URL							
application_1510769091693_0011	Thrift JDBC/ODBC Server	SPARK	hive	default	RUNNING	UNDEFINED	10%
http://10.196.177.198:4040							
application_1510769091693_0007	Zeppelin	SPARK	zeppelin	interactive	RUNNING	UNDEFINED	10%
http://10.196.177.198:4041							

Took 3 sec. Last updated by anonymous at November 16 2017, 11:23:47 AM. (outdated)

```
%sh
```

READY