Spark Cluster setup

Scripts Installation

1. Download and extract the scripts tarball from git.

```
hwleong@jyc1:~/test/spark>
https://github.com/hwleong/spark-bw-scripts/archive/0.1.1.tar.gz
hwleong@jyc1:~/test/spark> tar xvf 0.1.1
spark-bw-scripts-0.1.1/
spark-bw-scripts-0.1.1/README
spark-bw-scripts-0.1.1/bin/
spark-bw-scripts-0.1.1/bin/alps-run-example.sh
spark-bw-scripts-0.1.1/bin/alps-start-master.sh
spark-bw-scripts-0.1.1/bin/alps-start-slaves.sh
spark-bw-scripts-0.1.1/conf/
spark-bw-scripts-0.1.1/conf/spark-env.sh
spark-bw-scripts-0.1.1/conf/spark-env.sh.template
spark-bw-scripts-0.1.1/install.sh
spark-bw-scripts-0.1.1/modulefiles/
spark-bw-scripts-0.1.1/modulefiles/spark-2.1.1.template
spark-bw-scripts-0.1.1/sbin/
spark-bw-scripts-0.1.1/sbin/start-master.sh
spark-bw-scripts-0.1.1/sbin/start-slave.sh
hwleong@jyc1:~/soft/inst/spark> cd spark-bw-scripts-0.1.1/
```

2. Change directory into spark-bw-scripts-0.1 directory, and execute the "./install.sh" script.

```
hwleong@jyc1:~/test/spark/spark-bw-scripts-0.1.1> ./install.sh --help
usage: ./install.sh install_prefix_path
Example: ./install.sh /usr/local
hwleong@jyc1:~/test/spark/spark-bw-scripts-0.1.1> ./install.sh
/u/staff/hwleong/soft/apps/spark/2.1.1
```

3. (Optional) One may need to copy the spark module file ("$INSTALL_DIR/modulefiles/spark-2.1.1) into an existing directory within existing $MODULEPATH, or manually add the new modulefiles path into $MODULEPATH.

```
hwleong@jyc1:~/test/spark/spark-bw-scripts-0.1.1> module use
/u/staff/hwleong/soft/apps/spark/2.1.1/modulefiles
```

Quick Start

The following example assumes 1 master, 16 slaves and 1 submit host (total 18 compute nodes). At least three compute nodes are required for this example.

First host: Spark master

2nd to 2nd last host: Spark slaves

Last host: Spark submit host

1. Submit the following job.

```
hwleong@jyc1:~> qsub -I -l nodes=18:ppn=32,gres=shifter16,walltime=02:00:00 -v
UDI=lgerhardt/spark-2.1.1:v1
INFO: Job submitted to account: fyy
qsub: waiting for job 229180.jyc to start
qsub: job 229180.jyc ready
Start time = Mon Feb 12 14:30:42 CST 2018
In Torque Shifter prologue batchID: 229180
Starting munge service on compute nodes
Successfully started munge service on compute nodes
Initializing udiRoot, please wait.
Retrieving Docker Image
udiRoot Start successful
----------------------------------------
Begin Torque Prologue Mon Feb 12 14:31:01 CST 2018
Job Id:                 229180.jyc
Username:               hwleong
Group:                  bw_staff
Job name:               STDIN
Requested resources:
gres=shifter16,neednodes=18:ppn=32,nodes=18:ppn=32,walltime=02:00:00
Queue:                  normal
Account:                fyy
End Torque Prologue:  18.321 elapsed
----------------------------------------
Directory: /u/staff/hwleong
Mon Feb 12 14:31:12 CST 2018
hwleong@nid00010:~>
```

Submission options:

```
-I                              -> interactive job
-l nodes=18:ppn=32              -> request for 18 compute nodes with 32
processors per node
   gres=shifter16               -> request for Shifter environment on compute
nodes
   walltime=02:00:00            -> request for 2 hours job walltime
-v UDI=lgerhardt/spark-2.1.1:v1 -> request to setup "lgerhardt/spark-2.1.1:v1"
container on compute node, this container contains Spark ver. 2.1.1 environment.
```

2. Load "spark-2.1.1" module. (Remember to add spark modulefiles path to $MODULEPATH environment)

```
hwleong@nid00010:~> module load spark-2.1.1
hwleong@nid00010:~> module show spark-2.1.1
-------------------------------------------------------------------
/u/staff/hwleong/modulefiles/spark-2.1.1:
module-whatis    Spark implementation for Blue Waters via Shifter.
setenv           SPARK_MASTER_HOST nid00002
setenv           SPARK_HOME /usr/local/bin/spark-2.1.1-bin-hadoop2.7
setenv           SPARK_SCRIPTS /u/staff/hwleong/soft/apps/spark/2.1.1
setenv           SPARK_CONF_DIR /u/staff/hwleong/soft/apps/spark/2.1.1/conf
prepend-path     PATH /u/staff/hwleong/soft/apps/spark/2.1.1/bin
-------------------------------------------------------------------
```

3. Start Spark master.

```
hwleong@nid00010:~> alps-start-master.sh
Starting Spark Master on nid00002 (URL: http://10.128.0.3:8080)...
starting org.apache.spark.deploy.master.Master, logging to
/scratch/staff/hwleong/spark/spark-hwleong-org.apache.spark.deploy.master.Master-
1-nid00002.out
```

4. Start Spark slaves.

```
hwleong@nid00010:~> alps-start-slaves.sh
Starting Spark slaves...
starting org.apache.spark.deploy.worker.Worker, logging to
/scratch/staff/hwleong/spark/spark-hwleong-org.apache.spark.deploy.worker.Worker-
1-nid00024.out
starting org.apache.spark.deploy.worker.Worker, logging to
/scratch/staff/hwleong/spark/spark-hwleong-org.apache.spark.deploy.worker.Worker-
1-nid00018.out
starting org.apache.spark.deploy.worker.Worker, logging to
/scratch/staff/hwleong/spark/spark-hwleong-org.apache.spark.deploy.worker.Worker-
1-nid00021.out
starting org.apache.spark.deploy.worker.Worker, logging to
/scratch/staff/hwleong/spark/spark-hwleong-org.apache.spark.deploy.worker.Worker-
1-nid00015.out
starting org.apache.spark.deploy.worker.Worker, logging to
/scratch/staff/hwleong/spark/spark-hwleong-org.apache.spark.deploy.worker.Worker-
1-nid00011.out
starting org.apache.spark.deploy.worker.Worker, logging to
/scratch/staff/hwleong/spark/spark-hwleong-org.apache.spark.deploy.worker.Worker-
1-nid00028.out
starting org.apache.spark.deploy.worker.Worker, logging to
/scratch/staff/hwleong/spark/spark-hwleong-org.apache.spark.deploy.worker.Worker-
1-nid00029.out
starting org.apache.spark.deploy.worker.Worker, logging to
/scratch/staff/hwleong/spark/spark-hwleong-org.apache.spark.deploy.worker.Worker-
1-nid00012.out
starting org.apache.spark.deploy.worker.Worker, logging to
/scratch/staff/hwleong/spark/spark-hwleong-org.apache.spark.deploy.worker.Worker-
1-nid00016.out
starting org.apache.spark.deploy.worker.Worker, logging to
/scratch/staff/hwleong/spark/spark-hwleong-org.apache.spark.deploy.worker.Worker-
1-nid00006.out
starting org.apache.spark.deploy.worker.Worker, logging to
/scratch/staff/hwleong/spark/spark-hwleong-org.apache.spark.deploy.worker.Worker-
1-nid00013.out
starting org.apache.spark.deploy.worker.Worker, logging to
/scratch/staff/hwleong/spark/spark-hwleong-org.apache.spark.deploy.worker.Worker-
1-nid00020.out
starting org.apache.spark.deploy.worker.Worker, logging to
/scratch/staff/hwleong/spark/spark-hwleong-org.apache.spark.deploy.worker.Worker-
1-nid00007.out
starting org.apache.spark.deploy.worker.Worker, logging to
/scratch/staff/hwleong/spark/spark-hwleong-org.apache.spark.deploy.worker.Worker-
1-nid00025.out
starting org.apache.spark.deploy.worker.Worker, logging to
/scratch/staff/hwleong/spark/spark-hwleong-org.apache.spark.deploy.worker.Worker-
1-nid00003.out
starting org.apache.spark.deploy.worker.Worker, logging to
/scratch/staff/hwleong/spark/spark-hwleong-org.apache.spark.deploy.worker.Worker-
1-nid00019.out
```

5. Submit a Spark example job.

```
hwleong@nid00010:~> aprun -n 1 -b -- "${SPARK_HOME}/bin/run-example" --master
spark://$SPARK_MASTER_HOST:7077 SparkPi 10000 2>/dev/null
Pi is roughly 3.1418034714180347
```