

MPI & MAPREDUCE TUTORIALS

15-440 – Fall 2015

MPI TEST: HELLO WORLD

- 1) SSH into the cluster using the given document (login to the 01 node).
- 2) Copy the "HelloWorld.c" program from the webpage using the following command:

```
wget -N "http://qatar.cmu.edu/~mhhammou/15440-f15/recitations/HelloWorld.c"
```

(The `-N` flag is to overwrite any existing files with the name "HelloWorld.c")
- 3) Copy the **machinefile** from the webpage using the following command:

```
wget -N "http://qatar.cmu.edu/~mhhammou/15440-f15/recitations/machinefile"
```
- 4) Open the **machinefile** using **vim**:

```
vim machinefile
```
- 5) Edit the machine file so that it has the correct names for your cluster nodes. For example, if your Andrew ID is *ahmad*, replace

```
andrewid-n01.qatar.cmu.local:1
```

with:

```
ahmad-n01.qatar.cmu.local:1
```

Do so for the two nodes.
- 6) To save and exit, hit **escape**, then enter a **colon** (the ":" character), then type **"wq"**
- 7) Now, to compile your HelloWorld.c program, use the following command:

```
mpicc HelloWorld.c -o HelloWorld
```
- 8) For MPI to run, you must copy the **object file**, **"HelloWorld"** to the other machines. Simply run the following command to copy it to node 2 (we'll only be using two machines for this test):

```
scp -p "HelloWorld" andrewid-n02.qatar.cmu.local:/home/hadoop/
```
- 9) Finally, use the following command to run the MPI program:

```
mpiexec -f machinefile -n 2 ./HelloWorld
```

MAPREDUCE TEST: WORD COUNT

- 1) As before, log in to your 01 machine.
- 2) All the files are compiled and ready for you. All you have to do is use the following command to run the program:

```
hadoop jar WordCount.jar WordCount /user/hadoop/wordcount/input  
/user/hadoop/wordcount/output1
```

- 3) To view the output, run the following command:

```
hadoop dfs -text /user/hadoop/wordcount/output1/part-00000
```

- 4) You can use the MapReduce Web UI to view information about your job:

```
http://your\_andrewid-n01.qatar.cmu.local:50030/jobtracker.jsp
```