**Setup the Apache spark as standalone cluster**

**Machine Details:**

**Master** **Machine:**

IP: 10.113.59.74

Core: 8

RAM: 8 GB

**Slave Machine**

IP: 10.113.59.5

Core: 1

RAM: 1GB

**Configuration:**

**Modify the conf/spark-env.sh for both master and slave**

1) Add the following entry into the conf/spark-env.sh file

export SPARK\_MASTER\_IP=10.113.59.74 ->IP address of the Master

export SPARK\_MASTER\_PORT=7077

**Modify the conf/spark-defaults.conf for both master and slave**

spark.master spark://pradeepkm-ux:7077 ->URL of spark master you can get by http://10.113.59.74:8081/ where master is running

#spark.driver.memory 1g

spark.executor.meory 700m

spark.storage.memoryFraction 0.2

spark.storage.safetyFraction 0.4

spark.akka.heartbeat.pauses 6000

spark.akka.heartbeat.interval 1000

spark.default.parallelism 2

**Modify the /etc/hosts file for both master and slave**

1)Add the following entry into the /etc/hosts file

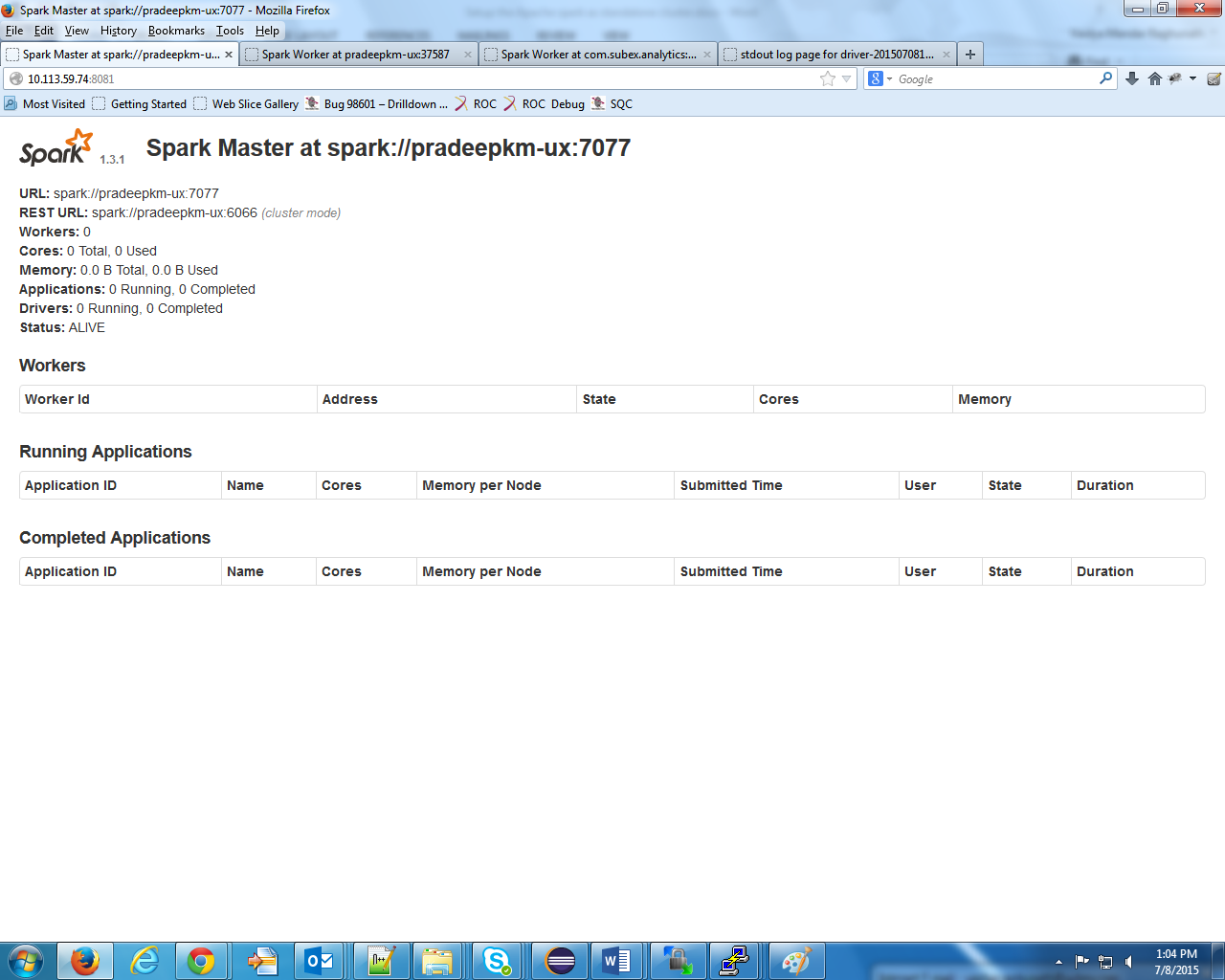
10.113.59.5 com.subex.analytics

10.113.59.74 pradeepkm-ux

**Run the program in spark standalone cluster environment**

1. **Start the master on Master machine using following command:**

./sbin/start-master.sh ->(Master) (Start the master on 10.113.59.74)



Note:

**To stop the master on master machine use following command:**

./sbin/stop-master.sh

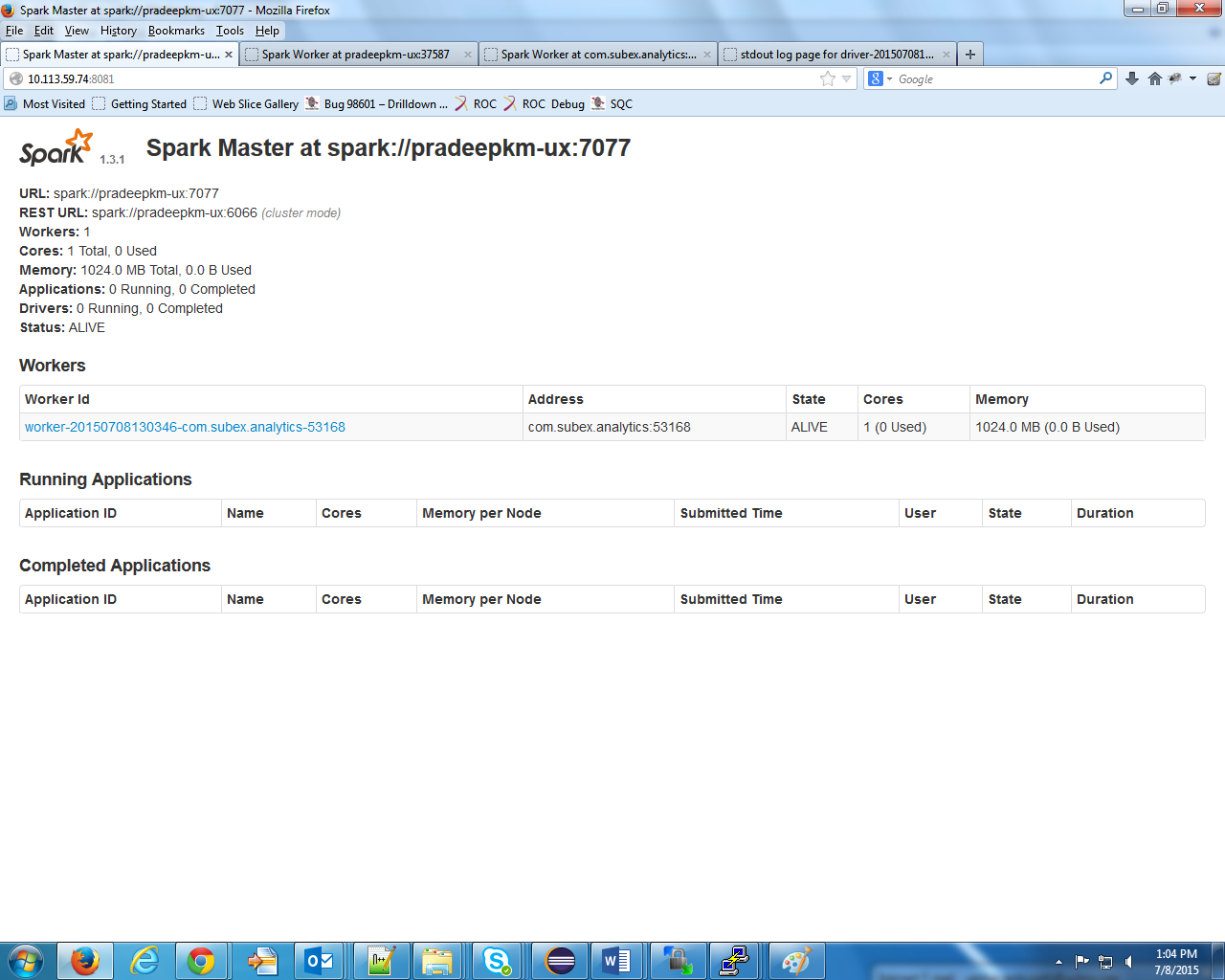
1. **Open the url to see the master is started or not**

http://10.113.59.74:8081/ or http://10.113.59.74:8080/

It will show as spark URL: spark://pradeepkm-ux:7077

1. **Start the slave(Worker) with following command on slave machine**

./bin/spark-class org.apache.spark.deploy.worker.Worker spark://pradeepkm-ux:7077 --cores 1 --memory 1g ->(Worker)(Start the worker on 10.113.59.5)



Note:

To stop the worker on worker machine press ctrl+c which will terminate the worker process.

1. **Check the worker started by using following url**

http://10.113.59.5:8081/ or http://10.113.59.5:8082/

1. **Check the worker added to master by following url**

http://10.113.59.74:8081/

On Url it will show the following details which indicates worker added to spark master

URL: spark://pradeepkm-ux:7077

REST URL: spark://pradeepkm-ux:6066 (cluster mode)

Workers: 1

Cores: 1 Total, 0 Used

Memory: 1024.0 MB Total, 0.0 B Used

Applications: 0 Running, 0 Completed

Drivers: 0 Running, 0 Completed

Status: ALIVE

1. **Now submit the job from worker node(10.113.59.5) with following command**:

./bin/spark-submit --master spark://pradeepkm-ux:7077 --class com.sparkexpert.OracleDBConnect --deploy-mode cluster ./spark-load-from-db-1.0-SNAPSHOT.jar

Parameter

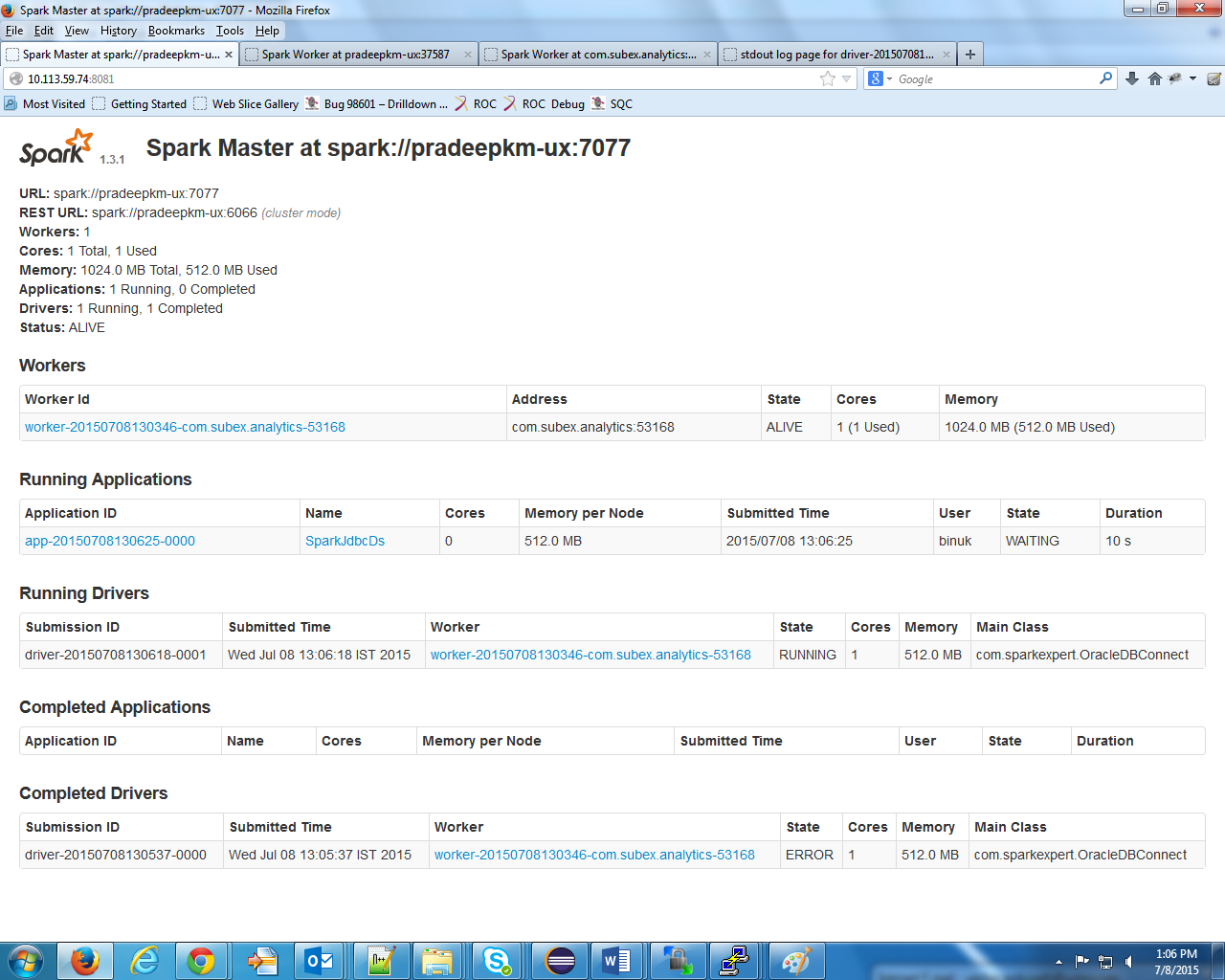
spark://pradeepkm-ux:7077 :Master url where it is running

com.sparkexpert.OracleDBConnect :Name of the class which we need to run which is present in jar

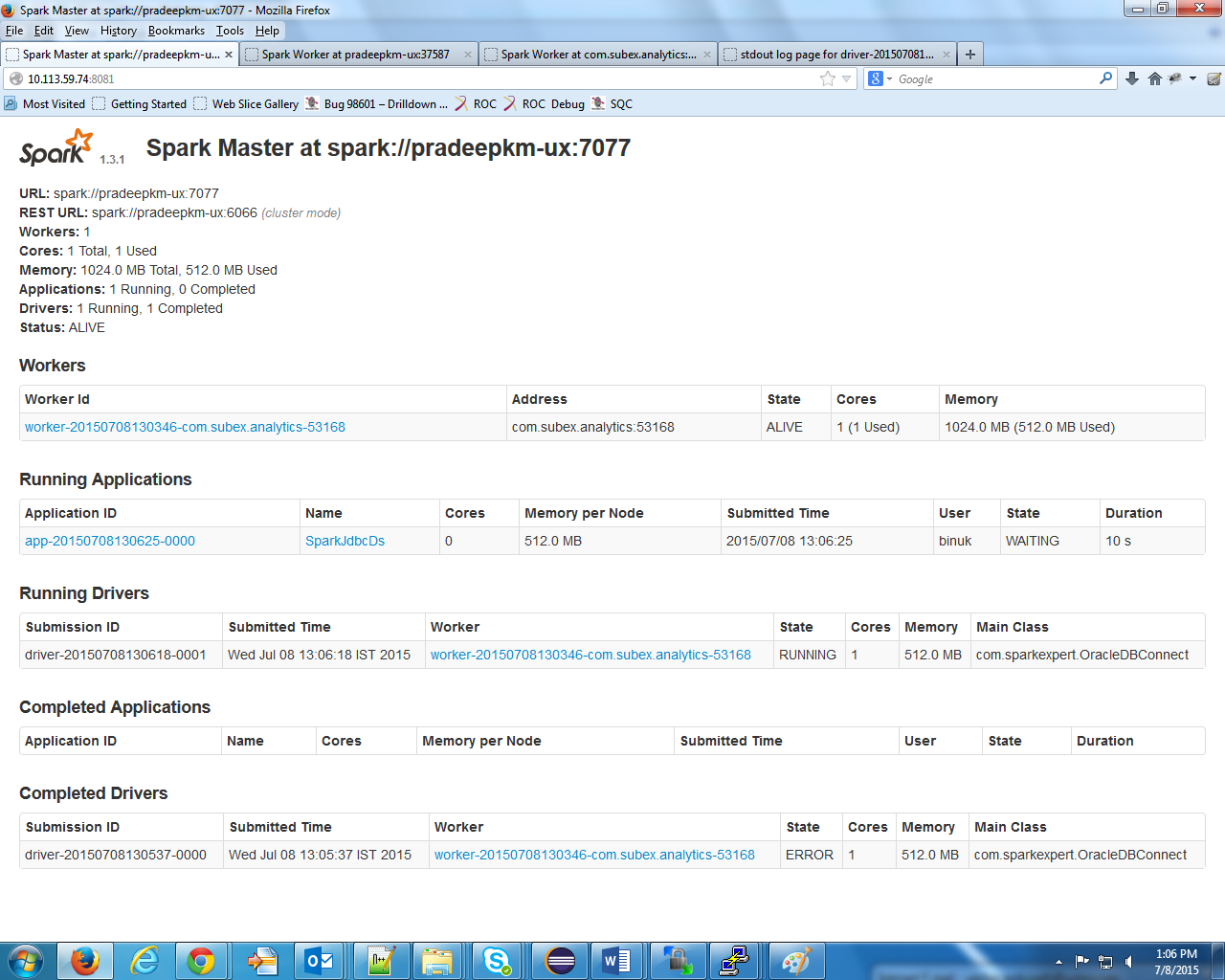
--deploy-mode: deployment mode for spark job submission

spark-load-from-db-1.0-SNAPSHOT.jar :Jar path

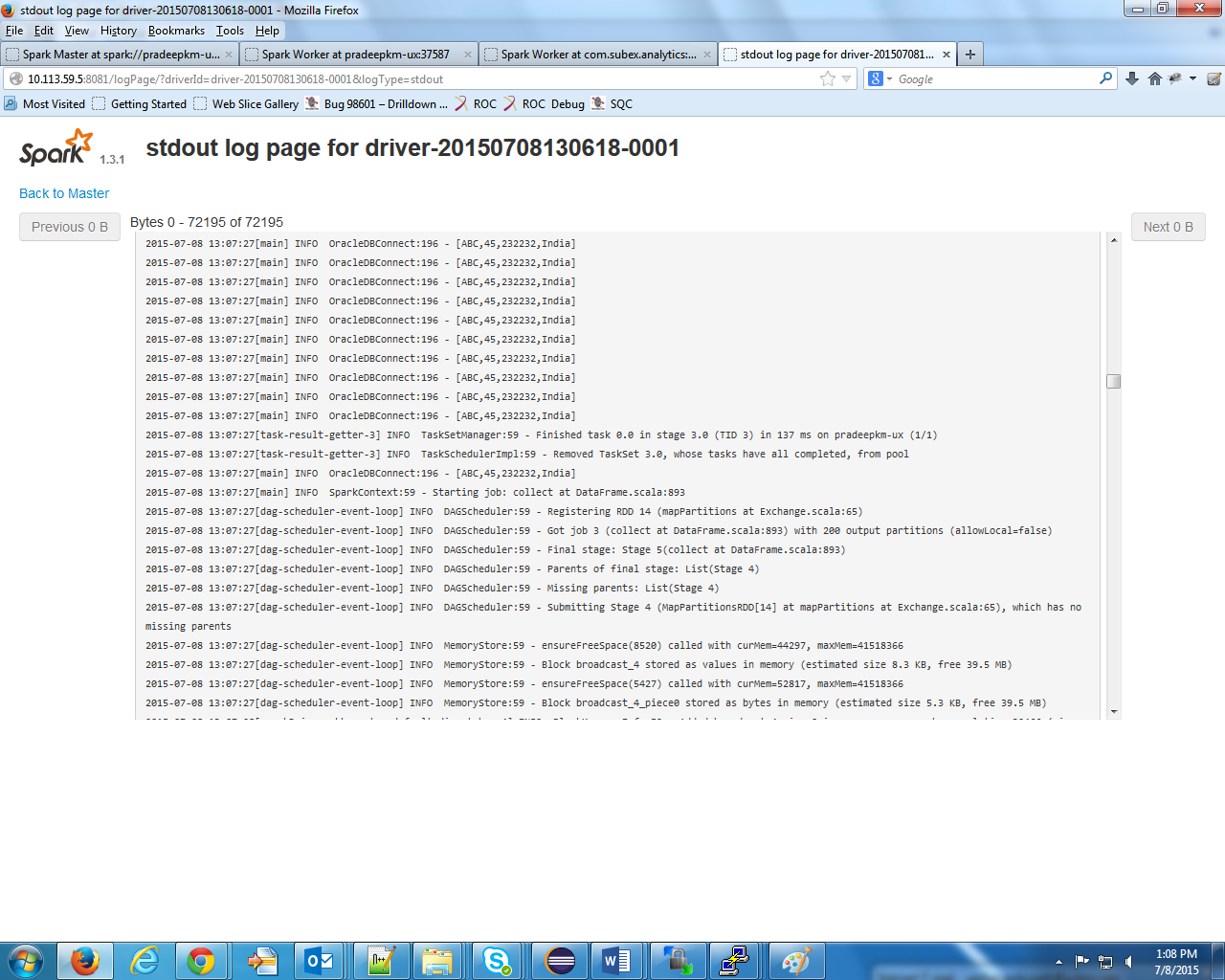
**Job status on Master Node:**



**Job Status on Worker Node:**



**Job Status Log:**

****

**Setup the password less connection between spark master and worker**

**Run the following command from master node (10.113.59.74)**

Tejas :: ~/apache\_spark/spark-1.3.1-bin-hadoop2.6/logs--> ssh-keygen

Generating public/private rsa key pair.

Enter file in which to save the key (/home/raghu/.ssh/id\_rsa):

Enter passphrase (empty for no passphrase):

Enter same passphrase again:

Your identification has been saved in /home/raghu/.ssh/id\_rsa.

Your public key has been saved in /home/raghu/.ssh/id\_rsa.pub.

The key fingerprint is:

6c:ea:44:75:f3:67:ca:60:42:36:74:66:01:a8:5a:80 raghu@pradeepkm-ux

The key's randomart image is:

+--[ RSA 2048]----+

| . .o.=. |

|E . .. + |

| . . = o |

| o = o o |

| o . S o . o |

| . . o o o + |

| o o |

| o |

| . |

+-----------------+

Tejas :: ~/apache\_spark/spark-1.3.1-bin-hadoop2.6/logs--> whoami

raghu

Tejas :: ~/apache\_spark/spark-1.3.1-bin-hadoop2.6/logs--> ssh-copy-id raghu@pradeepkm-ux

The authenticity of host 'pradeepkm-ux (10.113.59.74)' can't be established.

RSA key fingerprint is b9:26:6f:c6:d8:b8:63:c8:07:33:da:61:45:dc:7b:f5.

Are you sure you want to continue connecting (yes/no)? yes

Warning: Permanently added 'pradeepkm-ux' (RSA) to the list of known hosts.

raghu@pradeepkm-ux's password:

Now try logging into the machine, with "ssh 'raghu@pradeepkm-ux'", and check in:

.ssh/authorized\_keys

to make sure we haven't added extra keys that you weren't expecting.

Tejas :: ~/apache\_spark/spark-1.3.1-bin-hadoop2.6/logs-->

Copy the key to client machine(10.113.59.5) so server no need to give password while connecting to the worker:

Tejas :: ~/apache\_spark/spark-1.3.1-bin-hadoop2.6--> ssh-copy-id -i ~/.ssh/id\_rsa.pub binuk@10.113.59.5

binuk@10.113.59.5's password:

Now try logging into the machine, with "ssh 'binuk@10.113.59.5'", and check in:

.ssh/authorized\_keys

to make sure we haven't added extra keys that you weren't expecting.

Tejas :: ~/apache\_spark/spark-1.3.1-bin-hadoop2.6-->