

CSCI4180 Tutorial

Update on Cloud Platform

XU, Min (xum@cse)

Oct. 17, 2013

Outline

- Instance flavor
- User data
- Configure Hosts
- Check datanode status at master
- Check job history

Instance Flavor

- When launching instance, DO NOT use the flavor “m1.tiny (1VCPU/0GB Disk/512MB Ram)”
- Recommend “flavor.tiny” & “flavor.mini”

User Script Issue

- Issue with previous user data script caused by improper line break

```
#!/bin/bash
echo '-----'
if [ "`cat /etc/shadow | grep root | cut -d: -f2`" == '!*' ]; then
password=`dd if=/dev/urandom bs=1 count=320 2> /dev/null | tr -dc a-z2-9 | head -c8`
echo "root:$password"|chpasswd
sed -i 's/PasswordAuthentication no/PasswordAuthentication yes/g' /etc/ssh/sshd_config
/etc/init.d/ssh restart > /dev/null
echo "Your root account is ready."
echo "Username: root"
echo "Password: $password"
echo ''
echo 'Please keep it well/change it, or you will risk losing your assignment work!'
echo '-----'
fi
```

Configure Hosts

- If you ever encountered this error, you need to configure your /etc/hosts file on both master and slave nodes

java.lang.IllegalArgumentException: Wrong FS

- Procedures

1. Check hostname of all node in file /etc/hostname
2. On each node, add following content to /etc/hosts with root priviledge

<ip_of_master> <hostname_of_master>

<ip_of_slave1> <hostname_of_slave1>

...

<ip_of_slaveN> <hostname_of_slaveN>

Datanode issues

- If you ever encounter this error when trying to “hadoop dfs -put” or “hadoop dfs -copyFromLocal”, it means there is some issue with your datanode setup

...datanode could only be replicated to 0 nodes, instead of 1...

1. No datanodes running ⇒ ‘jps’ on datanode to check service status
2. Datanodes cannot communicate with namenode ⇒ Check datanode log on one of the datanodes
3. You can use “hadoop dfsadmin -report” to check datanodes status from namenode

```
$ hadoop dfsadmin -report
```

```
... Datanodes available: 4 ( 4 total, 0 dead) ...
```

4. Datanode has no harddisk ⇒ wrong flavor selected

Check Job History

- Check running situation of done MapReduce job
`hadoop job -history /your/output/directory`
- Task summary
 - # of mappers used
 - # of reducers used

Q & A