**`ISSUES and SOLUTIONS**

Issue: Not able to add the new partition as extended partition and change the partition id.

Soln: Add it as primary and change the partition id to desired one

Issue: Unable to delete a file even as root

Soln: delete file using inode number

ls -il --> to show inode number of file

find . -inum <inodenumber> -exec rm -i {} \; --> to delete the file with particular inode number

Que: How to find CPU details?

Ans: lscpu

Que: How to check link speed

Ans: ethtool –p eth0

Que: how to check NTP details?

Ans: ntpq –p , ntpupdate –u 10.1.2.2

Que how to check kernel version?

Ans: uname –r

echo "linuxpassword" | passwd --stdin linuxuser,In Unix echo $user:$password | /usr/sbin/chpasswd

Que how to check RAM in Linux?

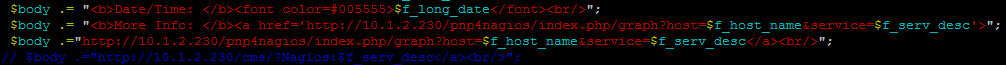
Ans: cat /proc/meminfo …….........in HPUNIX #####machinfo | grep Memory#### , /sbin/init.d/xntpd stop to stop ntpd

Que: How to check OS version details

Ans: lsb\_release –a , cat /etc/redhat –release

Que: URL in Nagios alert is not working

Make changes in file “/usr/local/nagios/libexec/nagios\_service\_mail” whose location is mentioned in “/usr/local/nagios/etc/objects/commands.cfg”..



Setting Path for python

export PATH=/usr/local/bin:$PATH

Mounting Unix on CentOS

share -F nfs -o rw=10.1.2.42 /Hadoop\_mig ---on Unix

In /etc/dfs/dfstab ---🡪 share -F nfs -o rw=10.1.2.42 /Homedepo

chmod -R 777 /Hadoop\_mig1

mount -o rw 10.2.2.16:/Hadoop\_mig /Mig\_Hadoop ---- on Centos

stoping ntp service in Unix 🡪 /sbin/init.d/xntpd stop

changing date in unix 🡪date -u 122205562015

Appache http configurtn giving paasswd authentication for webpage

htpasswd –cm /etc/httpd/userauthfile shyam

Installing PIP package but unable to import as it was getting installed under /root/anaconda2/ so copy

cp /root/anaconda2/\* /opt/anaconda2/

Issue : Tesklink tool enabling port

Soln : /opt/testlink-1.9.13-0/ctlscript.sh start

IME no.352356071315154

SSH tunneling

Ssh –L 0.0.0.0:80:10.1.11.92:8080 10.1.11.92 wrong

ssh -L 0.0.0.0:4455:127.0.0.1:4454 10.1.11.202 correct one

**New Things**

1.Generate public key using app puttygen

2.Open putty-->ssh-->auth

3.Just browse the key file

4.Then login

5.Will be able to login without password

6.Provided we should be having access.

7. capturing history completely -> script history\_log.txt the type exit to stop recording

8 .dos2unix is used to convert windows file format to linux –error was extra chareter at end of line “^M”

**Other commands**

1. **smem –tuk**
2. **netstat –tuln**
3. **sysctl –p** --> free the memory
4. **/sbin/init.d/syslog start/stop** --> to restart syslog service in unix
5. **service --status-all** --> to list all service
6. **vgchange –a y vgname** --> to activate any vg which gets deactivated after removing the device where y denotes yes n denotes no i.e. deactivate
7. **usermod –aG shyam.mohan** --> a is for appending so that other group of the user still remains there
8. **edquota -p shyam.mohan kishore.m** --> setting same quota limit to Kishore
9. **edquota -p shyam.mohan `awk -F: '$3 > 499 {print $1}' /etc/passwd**` -->setting same quota for all user assuming 500 is the starting uuid of users
10. **lvextend -L+120G /dev/mapper/data--volume-argos –r** -->to extend lv online
11. **Fsck –a /dev/sd1** -->to fix error automatically if errors are found
12. **yum --disablerepo=epel update -y ca-certificates** -->bug in centos
13. **yum install epel-release -->** basic pacakges
14. **yum install lvm2 smem vim mloacte nmap**
15. **scp –rp** --->r for folder and p for permission
16. **saconfig /dev/ciss3**
17. **sautil /dev/ciss3**
18. **pip install pythonpackage -->** installing python package
19. **umount –l** --> for lazy mount
20. **pkill -9 -u USERNAME** -->for killing all process belonging to a user
21. **w | sed '1,2d' | cut -f1 -d' ' | sort | uniq –c –to find sessions user wise**
22. **cat /proc/cpuinfo | grep 'model name' | wc –l --to find no. of cores**
23. **id -gn shyam.mohan to find whether domain user or not**
24. **mount -t cifs -o user=<uid>,pass=<password>,dom=mu-sigma.local //10.1.2.25/it /NetSight**
25. **mount -t cifs -o user=<uid>,pass=<pass>,dom=mu-sigma.local "//10.1.2.25/it/1. IT-Corporate/1.2 Servers-NON Windows/Windows-PWD-Expiry-Details" /WindowsPWDExpiry**
26. **//10.1.70.12/coop\_temp /Coop\_CAM\_TEMP/ cifs username=vignesh.n,password=justviky@123,iocharset=utf8,file\_mode=0777,dir\_mode=0777 0 0 to mount cifs with rw**
27. **ssh -fN -L 0.0.0.0:8080:127.0.0.1:6080 10.1.2.48 -i /root/.ssh/id\_rsa**
28. **to get usernames from Mysql --> use information\_schema; ,** **select GRANTEE,TABLE\_SCHEMA from SCHEMA\_PRIVILEGES; ,** **select distinct GRANTEE,TABLE\_SCHEMA from SCHEMA\_PRIVILEGES order by TABLE\_SCHEMA ;**
29. **Downloading package using yum** 
    1. **Yum install yum-plugin-downloadonly**
    2. **yum install --downloadonly --downloaddir=/opt/Dirty\_COW/ yum-utils**

**Points to remember**

1. Delete key wont work in unix.
2. We cannot extend live partition in unix...bt in linux its possible
3. Using “less” option we can scroll up and down while “more” option cannot
4. Swap will be 1000 times slower than RAM
5. When you are deleting data it just unlinks the data doesn’t delete it
6. Always enter in fstab after mounting something
7. Never stop iptables for any instance? No we can
8. Free –g gives memory details
9. Why temp takes 8Gb??
10. Lost+found contains the files which was deleted but was in use while deletion or if any software error or hardware error ooccurs to the system
11. Always check fstab entry before restarting a server
12. Always unmount before running fsck(file system consistency check)
13. Uid above 500 are normal user,root uid will be 0 and uid below 500 is system users and will not be able to login to Rstudio
14. Pam pluggable authentication module
15. Sftp server name --> get “filename”--> to do download file --> put “filename”--> to upload
16. 15 Tb for SAS from EMC
17. 1TB for R server from EMC
18. Non dfs will be always 10% of total size
19. Exclude a Repository from Yum Update  - yum - - disablerepo=mongodb-org-3.0 update
20. To check Kerberos validity check /etc/krb5.conf
21. 1 CPU means one core….one CPU can have only one core but on core can have more than one CPU

**Works Done**

1. configured web server so that clients can push data through Json also installed php and enabled json for it
2. Added Linux and Unix systems to KIWI Syslog systems---by modifying /etc/rsyslog.conf file

###/etc/rsys.conf configuration on server side###

$template TEMPLATE,"/var/log/%fromhost-ip%/syslog.log"

\*.\*?TEMPLATE

###/etc/rsys.conf configuration on client side###

\*.\* @@serverip:514

1. Added windows system to Nagios
2. Install NSclient ++ on windows and add port = 12489 in nsclient.ini file in c-->prgmfils-->nsclient/
3. In Nagios server---Created windows-test.cfg template under usr/local/Nagios/etc/objects/
4. Then add windows-testcfg files path to Nagios.cfg
5. usr/local/Nagios/etc/objects/Nagios –v /path to Nagios.cfg/-->to check for error in Nagios.cfg file
6. vi usr/local/Nagios/etc/objects/contact.cfg --- File where all contact groups are
   * 1. In Linux
7. Yum install Nagios Nagios-plugins nrpe Nagios-plugins-all Nagios-common
8. apt-get install nagios-nrpe-server(for ubuntu)
9. Change allow host as 127.0.0.1 to 127.0.0.1,172.31.0.146(Nagios server ip) inside nrpe.conf
10. service nagios-nrpe-server restart (for ubuntu)
11. Port no. 5666 shud be opened.
12. Plugins path will be --> /usr/lib64/nagios/plugins/
13. R server crashed – tried to troubleshoot found that everything is mounted using uuid which is related to harddisk and when harddisk changed uuid got changed and we got kernel panic error

Ran fsck on root file system but still that doesn’t get fixed, hence formatted the OS partition.

1. Created credentials for 3 users for AWS instance and enable auditd for monitoring login logs, install R server,
2. Configured R server and AD integration has been done.

**<<<<<<<<<<<<<<<<R server>>>>>>>>>>>>>>>>>>>>**

Our configuration

256 Ram 250 gb swap

Max we can run 16Gb file, we have 32 core

We can run 35 million rows

Package installation In Linux is easy

Port for Rstudio 6311

For 3..4 GB data set we need 64 gb min RAM and 16cpu cores

For 5gb data set 5\*5=25 gb ram is required if 10 users using 5gb data set then 250 gb ram and 250\*10 =2500gb hard disk size…ram is 1000 times faster than hard disk

Backups to be taken before formatting

/etc/fstab , /etc/groups,df –kh,/etc/pam.d-->passwd-auth(fr AD authntcn),system-auth,rstudio,/etc/smb.conf,/etc/krb5.conf,/etc/security/pam-windbind.conf

**<<<<<<<<<<<<<<Basic troubleshooting in R>>>>>>>>>>>>>>>>>>>>>**

1. Check home dir. Size

2. Check account lock

3. Check memory usage

4. Check home dir. permission

5. Restart rstudio

6. Check for pam

7. Check for Zombie process

8. install.packages(“zoo”,lib=”/C:/Imp/shyam”)

9.library(“ggplot2”,lib.loc=”/home/shyam.mohan/R/packages”) -->loading package from specific location

**<<<<<<<<<<<<<<Basic troubleshooting in SAS>>>>>>>>>>>>>>>>>>>>>**

There can be multiple reasons for slow SAS services,

1. If there is not enough space in the project folder
2. The SAS workspace is completely utilized (This holds all your intermediate tables and you are mostly not aware of this)
3. The SAS session exit is not clean due to network issues or closing the SAS session forcefully.
4. Sometimes the SAS code written is not efficient and it creates huge load, other users start feeling the slowness.

Sudo –u sasadmin ./sas server service down/up start

./sas server is a script

**<<<<<<<<<<<<<<Basic troubleshooting in Hadoop>>>>>>>>>>>>>>>>>>>>>**

/etc/init.d/cloudera-scm-server status

/etc/init.d/cloudera-scm-agent

service cloudera-scm-server-db status

To check size of projects in hdfs --> hdfs hdfs dfs -du -h /user/hive/warehouse/

`If size is 5 tb that multiply by replication factor will give size which hdfs should have, better use 5 DN,f 100gb that into 3 for space on Master node, what is the data volume or and velocity…..which all services like hive impala etc…cloud or local??

Assume 500GB --> 500g \*3repl =1.5TB + 25% of 1.5TB() + 30% to store intermediate(can go beyond)

RAM & CPU -->500GB one file -->500GB (optimum)(250GB slow) required n 125-250 cores…..if we have 32gb ram…then cores shud be in multiples of 32….recommended 32 core CPU n 80GB RAM

125GB RAM n 60 cores each…… maximum 48 cores for a system

Bench mark test in Hadoop

hadoop jar /usr/lib/hadoop-0.20-mapreduce/hadoop-test.jar nnbench -operation create\_write -maps 100 -reduces 25 -blockSize 2 -bytesToWrite 0 -numberOfFiles 4000 -replicationFactorPerFile 2 -readFileAfterOpen true -baseDir /benchmarks/NNBench-`hostname -s`

Spark

Spark-shell Access provided till 12th Nov 2015

**<<<<<<<<<<<<<<Hardening>>>>>>>>>>>>>>**

Selinux disable

Chkconfig off

Iptables stop

yum --disablerepo=epel update ca-certificates

pvcreate /dev/xvdf

vgcreate vg01 /dev/xvdf

lvcreate -L 20G -n lv\_tmp vg01

mkfs.ext4 -j /dev/mapper/vg01-lv\_tmp

mount /dev/mapper/vg01-lv\_tmp /tmp

vim /etc/fstab

chmod -R 1777 tmp ---> giving sticky bit to tmp folder

chmod -R g+s /data

yum install wget

wget <http://dev.mysql.com/get/mysql-community-release-el6-5.noarch.rpm>

yum localinstall mysql-community-release-el6-5.noarch.rpm

yum install mysql

yum install mysql-community-client mysql-community-server

service mysqld start

chown -R mysql:mysql /data

**<<<<<<<<<<<<<Rough Commands>>>>>>>>>>>**

yum install R

rstudio-server

rstudio-server test-config

rstudio-server verify-installation

yum install epel-release

yum --disablerepo=epel -y update ca-certificates

fdisk -l

mkfs.etx4 /dev/xvdf1

mkfs -t etx4 /dev/xvdf1

Sarf commnds on R server

yum remove rstudio

yum remove rstudio\*

rpm -qa | grep rstudio

yum remove rstudio-server-0.99.473-1.x86\_64

nmap -sT -O localhost

wget <https://www.rstudio.com/products/rstudio/download-server>

yum localinstall rstudio-server-rhel-0.99.473-x86\_64.rpm

rstudio-server

rstudio-server test-config

rstudio-server verify-installation

Sarf cmmnds on ktpl server

sudo -u sasadmin ./sas.servers start

fsck -F vxfs -o full /dev/vg02/lvol3

ioscan -fnk disk --> to check connectd harddisk

**R server /etc/rstudio conf files**

# R Session Configuration File

session-timeout-minutes=30

r-libs-user=~/R/packages

# Server Configuration File

www-port=8787

Topics

Kill commands & various RAID Configuration

Security Features in Linux

I week

1. Different types of protocols(http://vlaurie.com/computers2/Articles/protocol.htm)
2. Basics of Linux Routings(<http://www.techrepublic.com/article/understand-the-basics-of-linux-routing/>)

II week

1. Different RAID Configurations
2. Configuring Alias IP
3. Various Kill commands

III week

1. Basics of Hadoop
   1. Architecture of Hadoop
   2. Idea of writing Map n Reduce codes
   3. Basics of Hive
2. Use of Squid Proxy server and configuring  the same.

MYSQL

**To reset mysql password SET PASSWORD FOR 'root'@'localhost' = PASSWORD('MyNewPass');**

1. Remount project mount 10.2.2.221:/dellConsumer /dellConsumer , mount 10.2.2.221:/DELL\_SCM/ /DELL\_SCM/
2. mount -o rw,remount /data
3. eximount -o remount,rw /dev/mapper/database\_vol-mysql\_vol /data
4. mount -t ext4 -o remount,rw /dev/mapper/database\_vol-mysql\_vol /data
5. umount /data

Creating MySQL credential

create user 'divya.chaurasia'@'%' identified by 'divya@3214';

grant all privileges on .\* to 'divya.chaurasia'@'%';

grant file on \*.\* to 'divya.chaurasia'@'%';

to verify -->> mysql  –u  username -p

create database kfc\_temp;

grant all privileges on muscle.\* to 'shravan.rao'@'%' with grant option;

Mu Sigma Buissness Solutions Pvt. Ltd

Level 13, Aviator Building,

Ascendas - ITPL SEZ,

International Tech Park Bangalore ,

Whitefield , Bangalore

Karnataka,India

Data Import ad Export

use innovation;

show tables;

CREATE EXTERNAL TABLE IF NOT EXISTS viewed\_products(geolocation\_country string,post\_evar12 string,freq bigint) location '/user/shyam.mohan/viewed\_products';

select \* from viewed\_products;

drop TABLE viewed\_products;

CREATE TABLE viewed\_products(geolocation\_country string,post\_evar12 string,freq bigint) location '/user/shyam.mohan/viewed\_products';

import TABLE nike from '/user/shyam.mohan/udit\_nike'

select \* from nike;

mapred job -list

mapred job -kill job id

hdf dfsadmin -list

hdfs dfsadmin -report

mapred -l

Hadoop

28TB

300Gb ram apprx

mappres takes less time and reducers takes much longer time why?

due to assymmetric data

embedded database??

container= ram+core+libraries

Integrated development tool(IDE) eg eclipse,rstudio....rstudio help to write the code and r packages help to execute the codes

If u delete something but still space is not released why?? buck some files are locked by some process use lsof

cmd

Check winbind service if authentication not working

Day to Day Activities

1. Handling tickets for SAS,R, Hadoop, MySQL servers
2. Following up the tickets
3. Troubleshooting issues related to R SAS n Hadoop
4. 3 SAS,1R,hadoop,2MySQL,NTP server, Nagios server
5. Cloud Lloyd, Dixon, Otsuka,
6. Sending Daily Nagios Snapshots from Nagios
7. Generating weekly report for all SAS servers during Alternate Mondays.

OPEN SOURCE TOOLS

Yum install psacct

/etc/init.d/psacct status

/etc/init.d/psacct enable

Ac –d –p to list users login time

Sa to find total number of process

Hadoop Upgradation

Challenges:

1. Hive metadata backup

Checklist:

1. Backups
2. Stop all services
3. Bring down the cluster

Suggestion:

1. Make facility to backup
2. Tho ughts on secondary name node

Queries:

1. Do we have to upgrade each database separately??
2. Difference between Rolling upgrade and Cloudera-Manager upgrade

DNS

Caching(forwarder),caching dns refreshes, if there is not a forwarder it goes to 13 root servers…

Authoritative

We will have access to domain controller if we buy a domain and we can add munet,mucab to that DC

First DNS will check on root server for IP addr and root server replies as to check with top level domain server (.com) and while querying with .com it reply as to check with name server of .example.com and dns will obtain a ‘A’ record

A special PTR-record type is used to store reverse DNS entries. The name of the PTR-record is the IP address with the segments reversed + ".in-addr.arpa"(The reverse DNS database of the Internet is rooted in the Address and Routing Parameter Area) if ip addr is 10.1.2.3 then PTR record will be 3.2.1.10.in-addr.arpa

Client(Stub resolver)

/etc/resolv.conf

Why reverse lookup zone required

Go to putty

Login to 10.1.2.41 using ur credential

cd /home/urname

du –sh .rstudio 🡪 if it shows more than 200MB then follow below steps

cd .rstudio

du –sh \* 🡪 check which folder is having more space and navigate to that particular folder

cd suspended-sessions

du –sh \* 🡪 again check which file is taking more space

for eg if file called environment is taking more space then make that file empty by using below command

cat > filename 🡪then press enter and press cntrl+D(to save the file)

and now try to

login

Zookeeper- master follower concept, odd number of zookeeper services  
Node Manager – Apptn Master and container……….application master requests containers

Resource manager –cheduler

Application master keeps track of all jobs….and it keeps talking to resource manager…and Rm asks dn if it is capable then it will give it to node manager

Name node is not a part of Yarn

Yarn – Nodemanager Resource manager Resource history

Job history and resource manager shud b on Master Node

Journal node takes data from master node to Standby node and syncs it

Journaling node used a concept of quoram disk

NameNOde –Fsimage and editlogs(Fsimage contains info of where chunkz are stored and edits logs contain modification made to that chunk)

Cloudera takes 128 as default and core Hadoop takes 64 mb as default

Copying data is in two way----ROUNDROBIN AND available space by default it takes roundrobin

Spacewalk Issues

1. Edit up2date from /etc/sysconfig/rhn path change the IP address to hostname
2. Edit up2date from /etc/sysconfig/rhn path change https to http
3. noSSLserverURL=http://INBAAVMYSQL01.mu-sigma.local/XMLRPC

ABBOTT ISSUES and SOLUTIONS

1. Log location - /tmp/hive.log,
2. Pre authentication problem as keytab was expired – Service called zookeeper was not able to authenticate

Solution- Regenerated keytab in missing node but….that didn’t help so regenerated keytab for all nodes

1. Ambari Metrics not getting restarted – Killed the procces used by ams(ambary metrics service ) and restarted

AWS

1. As of 26 Tuesday Dasbboard got refreshed but date (23) didn’t get refreshed – It cuz AMAPS final date is 22
2. journal node has epoch number
3. if u decommision host from back end ...ambari database will not be updated so not recommended
4. To remove make active node to stand by in HA (hdfs haadmin -failover <target standby (service id of namenode)> <target active (service id of namenode)>
5. hdfs haadmin -failover nn1 nn2

The above example, prior to the command executing, nn1 was active and nn2 was standby.  After the command is ran, nn2 will be active and nn1 will be stand by.    
NOTE:  If the above command is ran and the current state before the command is ran is nn1 is standby and nn2 is active, the command will execute and say the failover is successful but no failover occurs and nn1 and nn2 are still in the same state.

CMDB URL - \\oneabbott.com\misc\GIS\ITSM\_Public\_Reports\Scheduled Reports Current\GIS\Daily\CMDB

Vi connect.pem paste content of key in that file

Ssh-keygen –t rsa

Cd .ssh

Cat id\_rsa.pub > authorized\_keys

Ssh –L 0.0.0.0.:8080:172.31.27.25:9000 ip-172.31.27.25 –I id\_rsa

Perl and Python

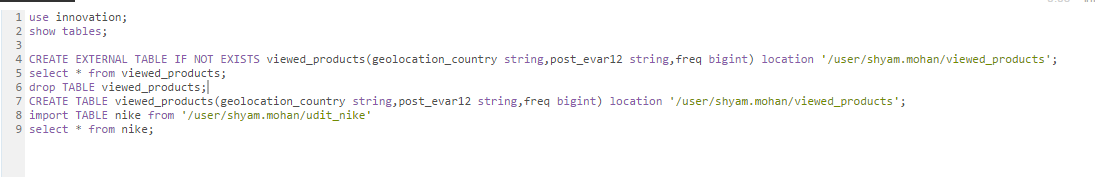
Perl in single codes doesn’t consider a special character as like that

To uninstall packages from data node

http://www.cloudera.com/documentation/enterprise/5-5-x/topics/cm\_ig\_uninstall\_cm.html

Interview Questions

1. Wat is logical name
2. What is fencing mechanism
3. Wat is JAR file
4. Wat is there is hive metadata --> Structure of table,date of creation, ID’s of tables,Database
5. Does reducers gets executed in select \* where
6. Diff between filezilla and winscp
7. Command not found why--> setting environment varibales
8. Where are all executables kept
9. Wat is heap space
10. How to decommission node from backend
11. Why jdk is required for Cloudera
12. NAS storage



Spark Cluster

To install java

apt-get install oracle-java7-installer

133 apt-add-repository ppa:webupd8team/java

134 apt-get install oracle-java7-installer

Untar spark-2.0.0-bin-hadoop2.7.tgz

**Issues History**

1. Cloudera manager service not starting even after couple of restarts – As it was HA enabled cluster, Secondary was down
2. wbinfo --domain-users - to obtain the all R users from group
3. R studio service keeps getting stopped

# service cloudera-scm-server-db stop

waiting for server to shut down............................................................... failed

pg\_ctl: server does not shut down

HINT: The "-m fast" option immediately disconnects sessions rather than

waiting for session-initiated disconnection.

Solution:

#cd /var/lib/cloudera-scm-server-db/data

# rm postmaster.pid

# service cloudera-scm-server-db status

pg\_ctl: no server running

# service cloudera-scm-server status

cloudera-scm-server dead but pid file exists

# rm /var/run/cloudera-scm-server.pid

# service cloudera-scm-server status

cloudera-scm-server is stopped

# service cloudera-scm-server-db start

DB initialization done.

waiting for server to start.... done

server started

# service cloudera-scm-server startx

Starting cloudera-scm-server: [ OK ]

Error: configure: error: no acceptable C compiler found in $PATH

Sln centos : yum groupinstall "Development tools"

Redhat : apt-get install build-essential

Pip3.5

yum install python3-pip

yum install python-setuptools

yum install python34-setuptools

pip3

easy\_install pip

easy\_install-3.5 pip

python3.5

python3.5 -m pip install sklearn

or: /usr/local/bin/python3 -m pip install sklearn