Cloudera CDH5 Setup

## \*For most parts of this installation, run as root.

# Version History

2015-12-08 Created Draft – Wenzong Zhao

# Introduction

This is a guide for walking through the steps of installing CDH5 on a Linux machine.

Core Steps Taken

1. Repo Installation
2. Cloudera Installation
3. MySql Installation
4. Starting Cloudera

# Repo Installation

Navigate to /etc/yum.repos.d

Create the following files in the directory:

soe-bigdata-5.repo

[soe-bigdata-5]

name=soe bigdata5 products

baseurl=http://gdsks1.nam.nsroot.net/SOELinux/repos/prod/soe6c5products-x86\_64

enabled=0

gpgcheck=0

soe-bigdata-c5.repo

[soe-bigdata-c5]

name=soe bigdata c5 products

baseurl=http://gdsks1.nam.nsroot.net/SOELinux/repos/prod/soe6c5bigdataproducts-x86\_64

enabled=0

gpgcheck=0

soe-bigdata-cm500.repo

[soe-bigdata-cm500]

name=soe bigdata cm500 products

baseurl=http://gdsks1.nam.nsroot.net/SOELinux/repos/prod/soe6cm500products-x86\_64

enabled=0

gpgcheck=0

soe-bigdata-cdh5.repo

[soe-bigdata-cdh5]

name=soe bigdata cdh5 products

baseurl=http://gdsks1.nam.nsroot.net/SOELinux/repos/prod/soe6cdh5products-x86\_64

enabled=0

gpgcheck=0

soe-bigdata-at.repo

[soe-bigdata-at]

name=soe bigdata analytical tools products

baseurl=http://gdsks1.nam.nsroot.net/SOELinux/repos/prod/soe6bigdataanalyticaltoolsproducts-x86\_64

enabled=0

gpgcheck=0

soe6iiproducts.repo

[soe6iiproducts]

name=soe6iiproducts

baseurl=http://gdsks1.nam.nsroot.net/SOELinux/repos/prod/soe6iiproducts-x86\_64/

enabled=0

gpgcheck=0

# Cloudera Installation

Run the below commands to install the repo

sudo yum install BD\_bigdata\_repo.x86\_64

Yum Refresh

yum clean all

yum makecache

Update JDK

yum remove jdk

yum install /net/stealth/export1/home1.localhost/sw/Linux/jdk-1.8.0\_65l64-linux.x86\_64.rpm

Cloudera Installation

yum -y install fuse redhat-lsb fuse-libs --enablerepo='soe6u5, soe6products, soe6local, soe-bigdata-c5';

yum -y install cyrus-sasl-gssapi libxslt --enablerepo='soe6u5, soe6products soe6local, soe-bigdata-c5';

yum install cloudera-manager-daemons cloudera-manager-agent enterprise-debuginfo cloudera-manager-server --enablerepo=soe-bigdata-5;

yum install BD\_cloudera\_cdh\_parcel\* --enablerepo=soe-bigdata-5;

# MySql Installation

Install MySql

yum install --disablerepo='\*' --enablerepo='soe-bigdata-c5, soe-bigdata-5, soe-bigdata-cdh5, soe6iiproducts' MySQL-server MySQL-client MySQL-shared-compat mysql-connector-java;

ln -s /usr/share/java/mysql-connector-java.jar /usr/share/cmf/lib/mysql-connector-java.jar

rm /var/lib/mysql/ib\_logfile1

rm /var/lib/mysql/ib\_logfile0

Create the file **/etc/my.cnf**

[client]

port = 3306

socket = /var/lib/mysql/mysql.sock

[mysqld]

key\_buffer = 16M

key\_buffer\_size = 32M

max\_allowed\_packet = 16M

thread\_stack = 256K

thread\_cache\_size = 64

query\_cache\_limit = 8M

query\_cache\_size = 64M

query\_cache\_type = 1

max\_connections = 600

read\_buffer\_size = 2M

read\_rnd\_buffer\_size = 16M

sort\_buffer\_size = 8M

join\_buffer\_size = 8M

# InnoDB settings

innodb\_file\_per\_table = 1

innodb\_flush\_log\_at\_trx\_commit = 2

innodb\_log\_buffer\_size = 64M

innodb\_buffer\_pool\_size = 2G

innodb\_thread\_concurrency = 8

innodb\_flush\_method = O\_DIRECT

innodb\_log\_file\_size = 512M

# log-bin=mysql-bin Commented out to save disk space in log file dir

# For MySQL version 5.1.8 or later

# binlog\_format=mixed

server-id = 1

[mysqldump]

quick

max\_allowed\_packet = 16M

[mysql]

no-auto-rehash

[myisamchk]

key\_buffer\_size = 20M

sort\_buffer\_size = 20M

read\_buffer = 2M

write\_buffer = 2M

[mysqlhotcopy]

interactive-timeout

[mysqld\_safe]

log-error=/var/log/mysqld.log

pid-file=/var/run/mysqld/mysqld.pid

skip-show-database=1

log-warnings=2

general-log=1

general-log-file==/var/log/general\_query.log

secure-auth=on

local-infile=0

Start MySql

service mysql start

/usr/bin/mysql\_secure\_installation

Change root password to: ‘m0nkeys’

Remove anonymous users? [Y/n] **Y**

Disallow root login remotely? [Y/n] **n**

Remove test database and access to it? [Y/n] **Y**

Reload privilege tables now? [Y/n] **Y**

mysql –u root -p

drop database servicemonitor;

drop database activitymonitor;

drop database hostmonitor;

drop database reportsmanager;

create database hostmonitor DEFAULT CHARACTER SET utf8;

create database servicemonitor DEFAULT CHARACTER SET utf8;

create database activitymonitor DEFAULT CHARACTER SET utf8;

create database reportsmanager DEFAULT CHARACTER SET utf8;

grant all on activitymonitor.\* TO 'CM\_Monitor'@'%' IDENTIFIED BY 'CM\_Monitor';

grant all on servicemonitor.\* TO 'CM\_Monitor'@'%' IDENTIFIED BY 'CM\_Monitor';

grant all on hostmonitor.\* TO 'CM\_Monitor'@'%' IDENTIFIED BY 'CM\_Monitor';

grant all on reportsmanager.\* TO 'CM\_Monitor'@'%' IDENTIFIED BY 'CM\_Monitor';

update mysql.user set Password=PASSWORD('CM\_Monitor') where user='CM\_Monitor';

create database scm;

grant all on scm.\* TO 'scm'@'%' IDENTIFIED BY 'scm';

update mysql.user set Password=PASSWORD('scm') where user='scm';

drop database scm;

create database nav;

grant all on nav.\* TO 'nav'@'%' IDENTIFIED BY 'nav';

update mysql.user set Password=PASSWORD('nav') where user='nav';

create database navms;

grant all on navms.\* TO 'navms'@'%' IDENTIFIED BY 'navms';

update mysql.user set Password=PASSWORD('navms') where user='navms';

create database sentry;

grant all on sentry.\* TO 'sentry'@'%' IDENTIFIED BY 'sentry';

update mysql.user set Password=PASSWORD('sentry') where user='sentry';

flush privileges;

exit

/usr/share/cmf/schema/scm\_prepare\_database.sh --host localhost --user root --password --scm-host localhost mysql scm scm scm

mysql –u root -p

CREATE USER 'hive'@'%' IDENTIFIED BY 'hive';

GRANT ALL PRIVILEGES ON metastore.\* TO 'hive'@'%' WITH GRANT OPTION;

update mysql.user set Password=PASSWORD('hive') where user='hive';

select Host, User, Password from mysql.user ;

FLUSH PRIVILEGES;

quit;

# Starting Cloudera

service cloudera-scm-server start

service cloudera-scm-agent start

Wait 2 minutes, then access <http://hostname:7180/cmf> with the credentials: admin/admin.

Click through each screen.

# If Previous Installation of Cloudera Exists

Use ‘yum remove’ to remove any associated packages that Cloudera Manager reports as in use.  
Additionally, remove the default HDFS directories, allowing them to be overwritten by the new installation.

rm -rf /var/log/hadoop-hdfs

rm -rf /dfs