Cloudera CHD5.7 Kerboros 인증 적용 매뉴얼

1. 설정 환경

* CentOS6.4 64bit
* Cloudera Manger 5.7
* External DataBase MySql(InnoDB engine)
* 설치계정 root

1. Cluster 구성

192.168.1.175 cm.localdomain cm (Cloudera Manger)  
192.168.1.171 name.localdomain name (NameNode)  
192.168.1.172 sname.localdomain sname (Secondary NameNode)  
192.168.1.173 data.localdomain data (DataNode)  
192.168.1.174 data2.localdomain data2 (DataNode)

1. Ntp 설치해야 된다 cluster 간의 시간이 5분이상 틀리면 kerberos가 동작하지 않는다.

# yum install ntp  
$>vi /etc/ntp.conf  
server kr.pool.ntp.org  
server time.bora.net  
server time.kornet.net  
# service ntpd start  
# chkconfig ntpd on

타임존 설정

$>vi /etc/sysconfig/clock  
ZONE="Aisa/Seoul"  
UTC=true  
ARC=false

KST로 타임존 변경

$> ln -sf /usr/share/zoneinfo/Asia/Seoul /etc/localtime

ln -sf /usr/share/zoneinfo/GMT /etc/localtime

$> **date**

1. JCE Policy 파일 설치

# wget <http://download.oracle.com/otn-pub/java/jce/7/jce_policy-7.zip>

# tar xvfz jce\_policy-7.zip

다운로드 받은 파일을 각 호스트에 복사해준다. 이미 존재한다고 메시지가 뜨면 덥어씌운다.

# scp local\_policy.jar root@name.locadomain: /usr/java/jdk1.7.0\_67-cloudera/jre/lib/security/

# scp US\_export\_policy.jar root@name.locadomain: /usr/java/jdk1.7.0\_67-cloudera/jre/lib/security/

1. 설치 과정

Step 1) 모든 Cluster의 모든 서비스를 중지한다.

Step 2) Cloudera Manager가 설치된 host(cm.localdomain)에 Kerberos libs, Kerberos Server  
 Kerberos Client를 설치한다.

# yum -y install krb5-libs krb5-server krb5-workstation

# yum -y install krb5-server-ldap krb5-auth-dialog pam\_krb5

Step 3) NameNode와 DataNode에 Kerberos Client를 설치한다.

# yum -y install krb5-libs krb5-workstation

Step 4) Realm 명칭 변경(cm.localdomain)

# vi /etc/krb5.conf

[libdefaults]  
default\_realm = LOCALDOMAIN  
 dns\_lookup\_realm = false  
 dns\_lookup\_kdc = false  
 ticket\_lifetime = 24h  
 renew\_lifetime = 7d  
 forwardable = true

[realms]  
 LOCALDOMAIN = {  
 kdc =192.168.1.175:88  
 admin\_server = 192.168.1.175:749  
 default\_domain=localdomain  
 }

[domain\_realm]  
 cm.localdoamin = LOCALDOMAIN  
 name.localdoamin = LOCALDOMAIN  
 sname.localdomain = LOCALDOMAIN  
 data.localdomain= LOCALDOMAIN  
 data2.localdomain= LOCALDOMAIN

Step 5) vi /var/kerberos/krb5kdc/kdc.conf 수정

[kdcdefaults]  
 kdc\_ports = 88  
 kdc\_tcp\_ports = 88

[realms]  
 LOCALDOMAIN = {  
 # master\_key\_type = aes256-cts  
 max\_renewable\_life = 7d 0h 0m 0s  
 acl\_file = /var/kerberos/krb5kdc/kadm5.acl  
 dict\_file = /usr/share/dict/words  
 admin\_keytab = /var/kerberos/krb5kdc/kadm5.keytab  
 supported\_enctypes = aes256-cts:normal aes128-cts:normal des3-hmac-sha1:normal arcfour-hmac:normal des-hmac-sha1:normal des-cbc-md5:normal des-cbc-crc:normal  
 default\_principal\_flags = +renewable  
 }

Step 6) vi /var/kerberos/krb5kdc/kadm5.acl ACL(Access Control List) 수정한다.

\*/admin@LOCALDOMAIN \*  
cloudera-scm@LOCALDOMAIN \* flume/\*@LOCALDOMAIN  
cloudera-scm@LOCALDOMAIN \* hbase/\*@LOCALDOMAIN  
cloudera-scm@LOCALDOMAIN \* hdfs/\*@LOCALDOMAIN  
cloudera-scm@LOCALDOMAIN \* hive/\*@LOCALDOMAIN  
cloudera-scm@LOCALDOMAIN \* httpfs/\*@LOCALDOMAIN  
cloudera-scm@LOCALDOMAIN \* HTTP/\*@LOCALDOMAIN  
cloudera-scm@LOCALDOMAIN \* hue/\*@LOCALDOMAIN  
cloudera-scm@LOCALDOMAIN \* impala/\*@LOCALDOMAIN  
cloudera-scm@LOCALDOMAIN \* mapred/\*@LOCALDOMAIN  
cloudera-scm@LOCALDOMAIN \* oozie/\*@LOCALDOMAIN  
cloudera-scm@LOCALDOMAIN \* solr/\*@LOCALDOMAIN  
cloudera-scm@LOCALDOMAIN \* sqoop/\*@LOCALDOMAIN  
cloudera-scm@LOCALDOMAIN \* yarn/\*@LOCALDOMAIN  
cloudera-scm@LOCALDOMAIN \* zookeeper/\*@LOCALDOMAIN

Step 7) kdb5\_util을 이용해서 KDC가 티켓 발급 대상 데이터베이스 생성  
# /usr/sbin/kdb5\_util create -r LOCALDOMAIN -s  
# ls /var/kerberos/krb5kdc/principal  
\*\* 참고 : 데이터베이스 삭제 명령  
 kdb5\_util destroy-f LOCALDOMAIN

Step 8) Principal을 데이터베이스발급한다.

# kadmin.local

kadmin.local: addprinc root/admin  
kadmin.local: addprinc cloudera-scm/admin  
kadmin.local: addprinc cloudera-scm/admin@LOCALDOAMIN

\*\* 참고 principal 관련명령어  
principal 조회명령어 kadmin.local: listprincs  
principal 삭제명령어 kadmin.local: delprinc  
principal 비밀번호변경 kadmin.local: cpw

Step 9) kadmin 키탭파일 추가

# kadmin.local

kadmin.local : ktadd -k /var/kerberos/krb5kdc/kadm5.keytab kadmin/admin  
kadmin.local : ktadd -k /var/kerberos/krb5kdc/kadm5.keytab kadmin/changepw

Step 11) 호스트 principals 등록

Kadmin.local: addprinc -randkey host/cm.localdomain  
kadmin.local: addprinc -randkey host/name.localdomain  
kadmin.local: addprinc -randkey host/sname.localdomain  
kadmin.local: addprinc -randkey host/data.localdomain  
kadmin.local: addprinc -randkey host/data2.localdomain

Step 12) 호스트 keytab 등록

kadmin.local: ktadd -k /etc/krb5.keytab host/cm.localdomain  
kadmin.local: ktadd -k /etc/krb5.keytab host/name.localdomain  
kadmin.local: ktadd -k /etc/krb5.keytab host/sname.localdomain  
kadmin.local: ktadd -k /etc/krb5.keytab host/data.localdomain  
kadmin.local: ktadd -k /etc/krb5.keytab host/data2.localdomain

Step 13) Kerboros 서비스를 시작한다.  
 #service krb5kdc start  
 #service kadmin start

서비스 프로세스를 확인한다.

netstat –nat | grep 88로 포트 Listen을 확인한다.

netstat –nat | grep 749로 포트 Listen을 확인한다.

Step 14) keytab file 확인

# ktutil   
ktutil : rkt /var/kerberos/krb5kdc/kadm5.keytab   
ktutil : list

~~Step15) /etc/kdc5.conf 파일을 각 호스트에 (클라이언트)에 전송한다.~~

~~# scp /etc/krb5.conf root@name.localdomain:/etc/krb5.conf~~

CM에서 Kerberos 클라이언트 구성배포를 선택하면 자동배포

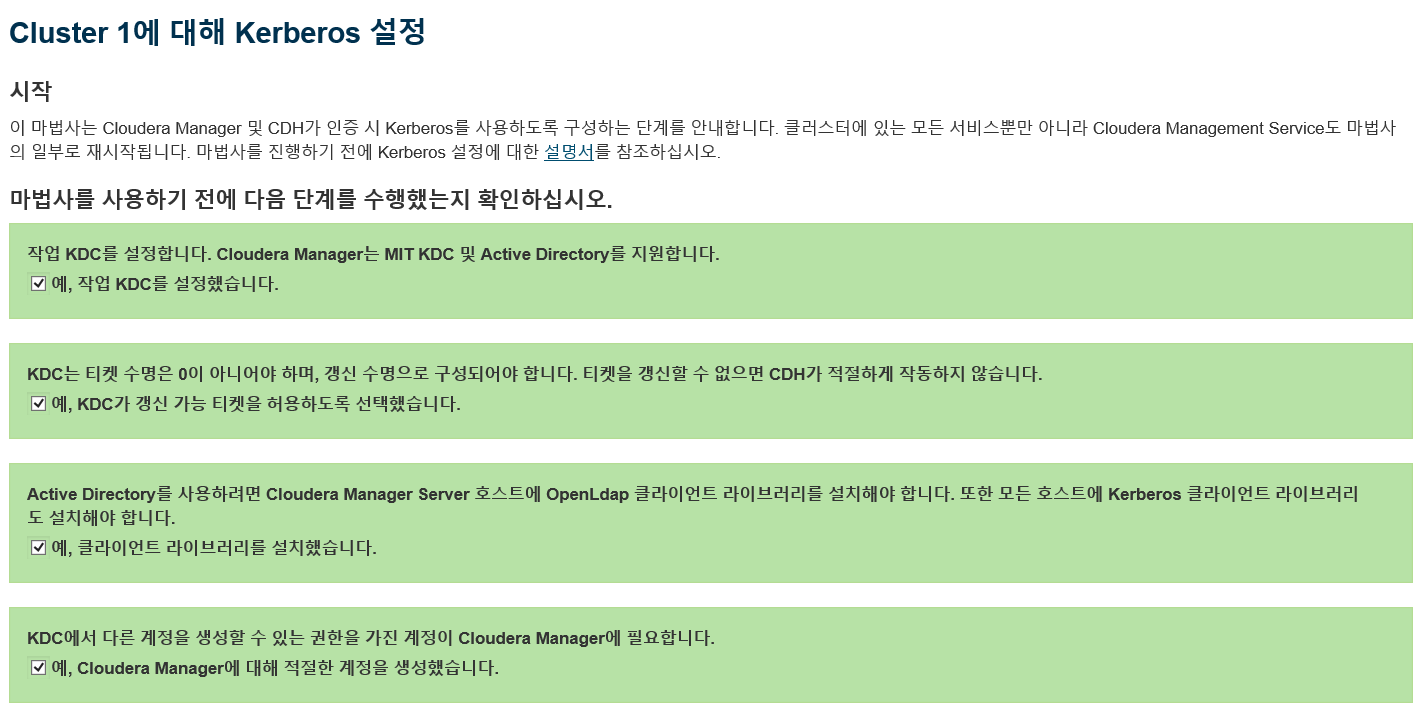
되므로 위 작업은 필요업음

Step16) 각 클라이언트 컴퓨터에서 서버 컴퓨터로 접속을 확인한다.

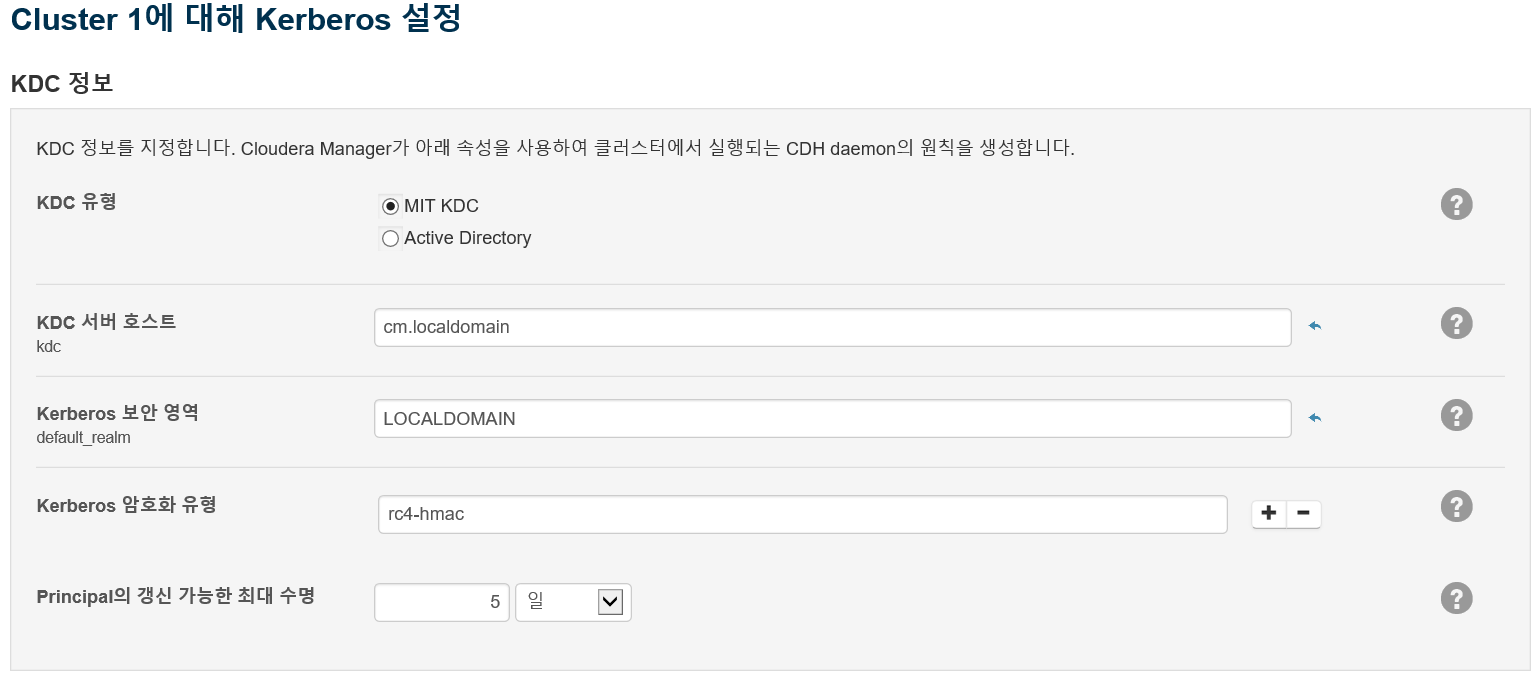
# kadmin -p root/admin

Step 18) Cloudera Manger 설정  
 #service cloudera-scm-manager start

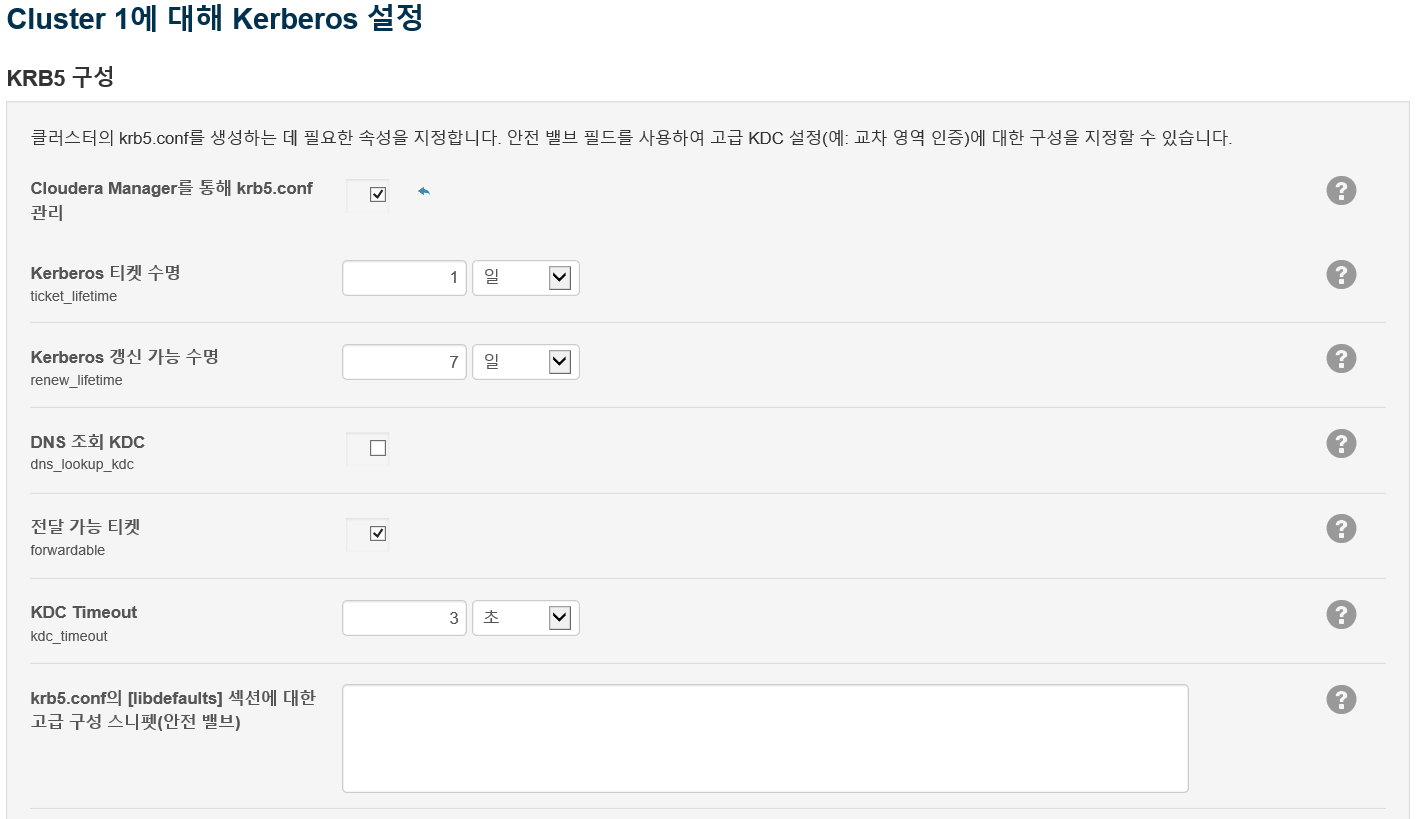
Step 19) 관리 > 보안 > kerberos 설정  
사전작업 진행여부 체크박스에 모두 체크하고 계속을 클릭한다.



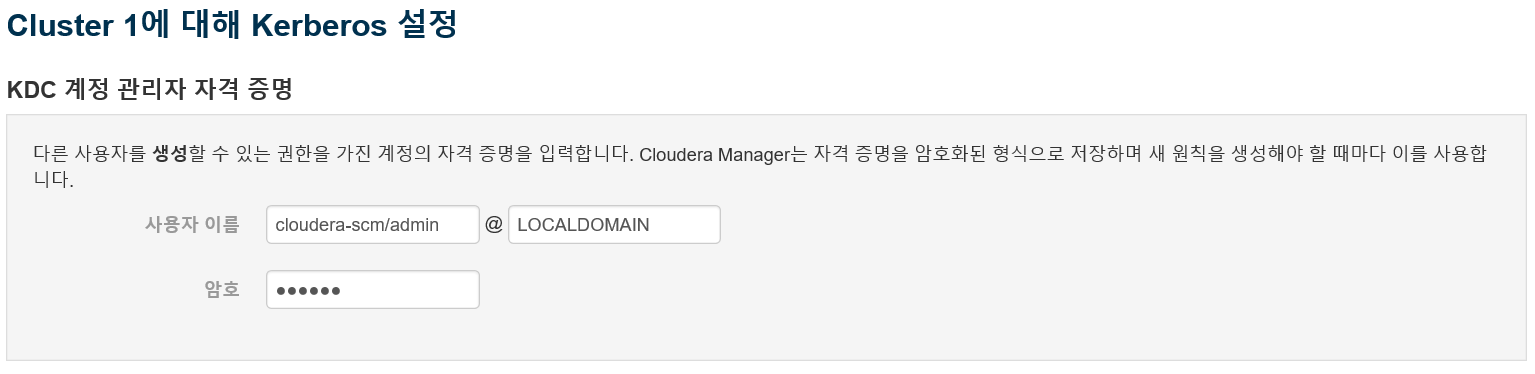
Step 20) KDC 서버 지정 및 Realm 설정



Step 21) 티켓수명과 갱신가능 수명 지정

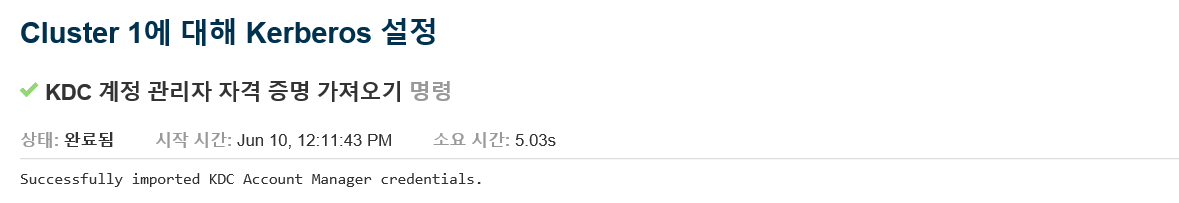


Step 22) 계정관리자격증명



Principal 생성 시 입력했던 암호를 입력한다.

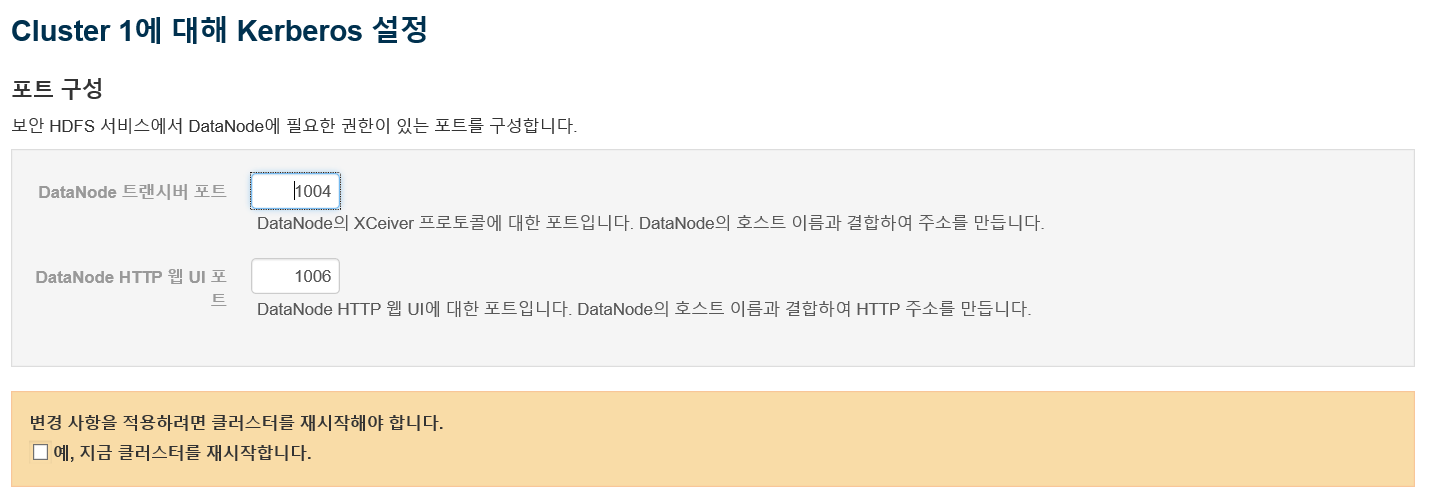
Step 23) 계정관리자격 가져오기



Step 24) Kerberos 설정 서비스 지정



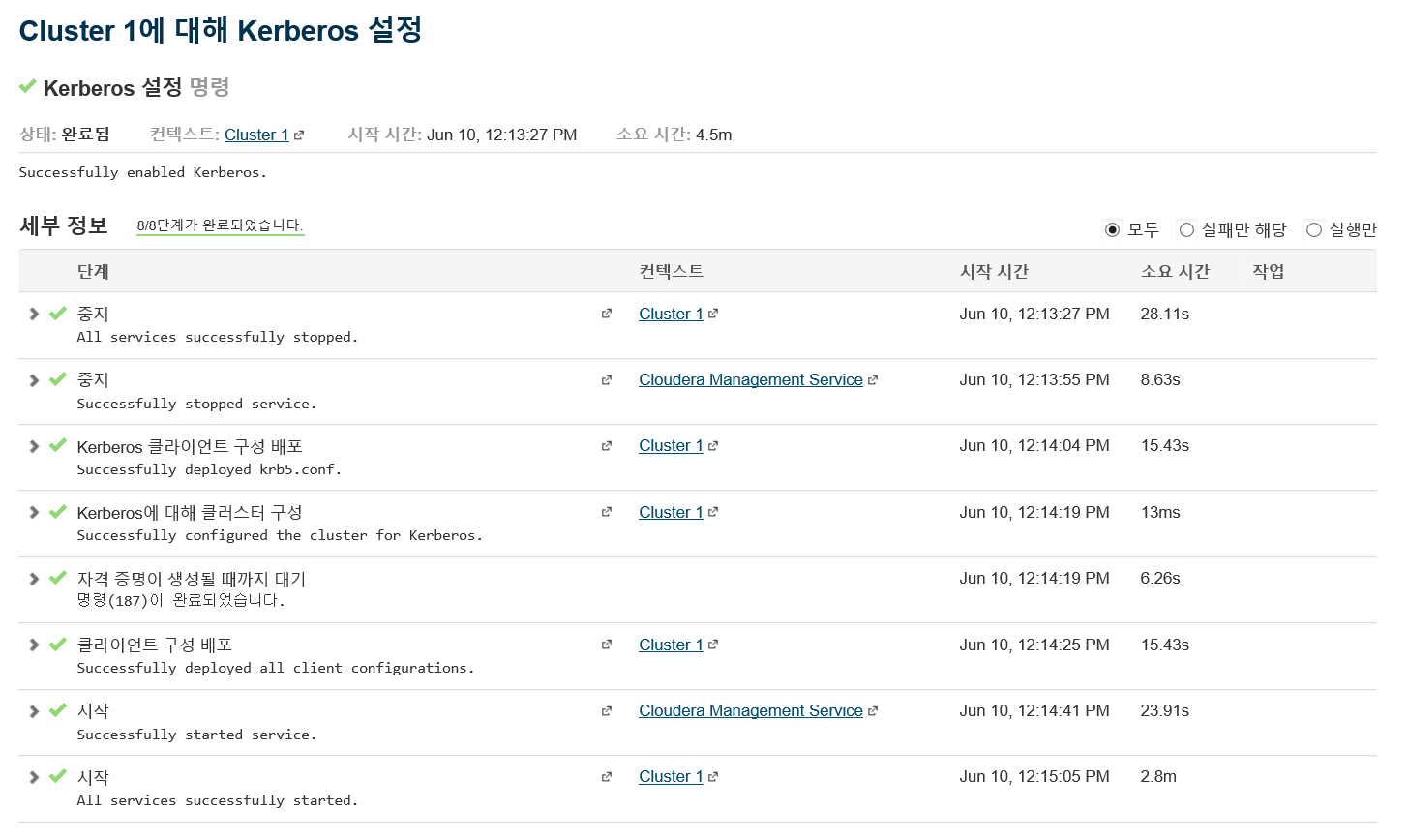
Step 25) port 설정



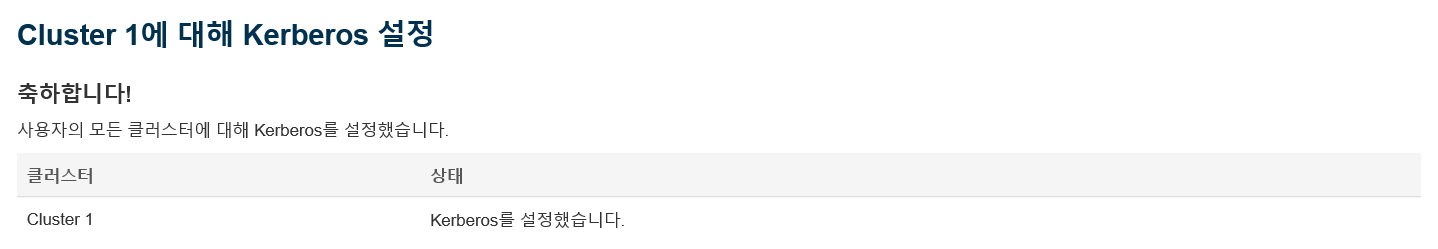
클러스터를 재시작하면 서비스 principal과 keytab이 자동 등록된다.

Step 26) 클러스터 재시작





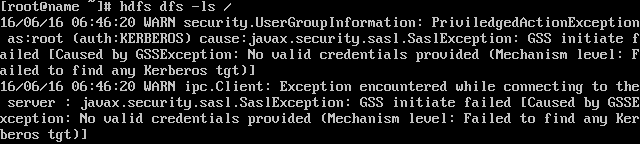
Step 27) 설정완료



설치 테스트

Test 1) 설치 테스트를 위해 HDFS 명령어를 수행해 본다.

# hdfs dfs –ls /



\*\* kerberos 티켓 발생이 안되었기 때문에 권한 에러가 발생한다.

kinit를 통해 티켓발행 발행후 수행해야 된다

# kinit hdfs@LOCALDOMAIN



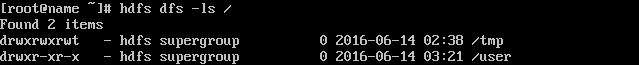
위와 같은 에러가 발생하는 경우 kdc 서버에서 hdfs에 대한 principal을 등록해준다.

kadmin.local: addprinc hdfs

다시 namenode로 와서 티켓을 발생하고 하둡 명령을 수행해 본다.

# kinit hdfs@LOCALDOMAIN

# hdfs dfs –ls /



정상적으로 수행되는 것을 볼 수 있다.

Test 2) 설치 테스트를 위해 MapReduce example을 테스트 한다.

# hadoop jar /opt/cloudera/parcels/CDH/lib/hadoop-0.20-mapreduce/hadoop-examples.jar pi 10 10000

아래와 같은 에러가 발생한다.

Requested user hdfs is not whitelisted and has id 496,which is below the minimum allowed 1000  
Failing this attempt. Failing the application.  
16/06/16 06:55:52 INFO mapreduce.Job: Counters: 0  
Job Finished in 5.631 seconds  
java.io.FileNotFoundException: File does not exist: hdfs://name.localdomain:8020/user/hdfs/QuasiMonteCarlo\_1466060143436\_1132364948/out/reduce-out

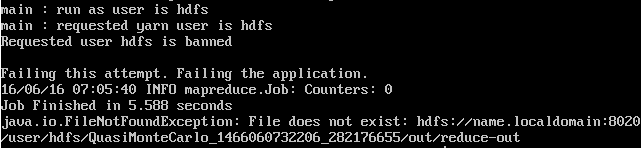
Admin Console >> Yarn >> 구성 >> 보안

최소사용자 ID를 450으로 낮춰준다.

Yarn 서비스를 재 시작 후 다시 실행해 본다.



# hadoop jar /opt/cloudera/parcels/CDH/lib/hadoop-0.20-mapreduce/hadoop-examples.jar pi 10 10000

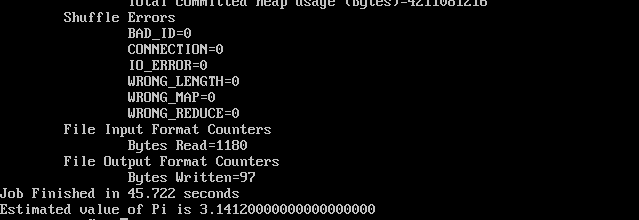


Admin Console >> Yarn >> 구성 >> 보안

Yarn 허용되지 않은 사용자에 hdfs가 있다 hdfs를 삭제하고 허용된 시스템 사용자에 등록 해준다.



Yarn 서비스를 재 시작 후 다시 실행해 본다.



정상적으로 작업이 종료되었다.

트러블슈팅

Problem 1) Hue service 시작시 아래와 같은에러가 발생하는 경우

Couldn't renew kerberos ticket in order to work around Kerberos 1.8.1 issue. Please check that the ticket for 'hue/name.localdomain@LOCALDOMAIN' is still renewable

Solution ) 아래 명령어를 수행해 준다.

kadmin.local: modprinc -maxrenewlife 90day krbtgt/LOCALDOMAIN  
kadmin.local: modprinc -maxrenewlife 90day +allow\_renewable hue/name.localdomain@LOCALDOMAIN

Problem 1)

SIMPLE authentication is not enabled. Available:[TOKEN, KERBEROS]

위와 같은 에러가 발생하는 경우는 CORE-SITE.XML 파일에 아래와 같은 내용이 변경되지 않아서 이다. CM에 재로그인 하여 오래된 구성을 재배포 해준다.

다른 서버에서는 모두 잘되는데 특정서버만 안되는 경우는 게이트웨이가 설치가 않되서 그런 경우가 있으니 게이트웨이 서버는 모두 설치해줘야 한다.

manage배포가 안되는 경우는 수동으로 해준다.( /etc/hadoop/conf.cloudera.hdfs/core-site.xml)

수동으로 수정을 해줘도 않된는 경우는 gateway가 없는 서버에는 cm이 client configuration을 배포하지 않는다.

<property>

<name>hadoop.security.authentication</name>

<value>kerberos</value> <!-- A value of "simple" would disable security. -->

</property>

<property>

<name>hadoop.security.authorization</name>

<value>true</value>

</property>

커버로스 적용후 hue를 통해 oozie작업을 위한 설정

1. Yarn-site.xml 수정(cm을 통해 작업)

yarn.scheduler.maximum-allocation-mb 전체메모리의 5/6

yarn.nodemanager.resource.memory-mb 전체메모리의 1/6

1. mysql이 root 또는 접속 계정으로 원격접속이 가능하도록 조치해준다.

GRANT ALL PRIVILEGES ON \*.\* TO 'root'@'%' identified by ‘kangjh’;

FLUSH PRIVILEGES;

3. admin 프린시펄을 생성해준다.

addprinc admin/name.localdomain@LOCALDOMAIN

addprinc admin/sname.localdomain@LOCALDOMAIN

addprinc admin/data.localdomain@LOCALDOMAIN

addprinc admin/data2.localdomain@LOCALDOMAIN

addprinc admin@LOCALDOMAIN

4. 각 호스트에서 admin 계정을 생성해준다.

$>adduser admin -G hadoop

5. JDBC 드라이버를 복사한다.

sqoop사용시 로컬 jdbc 드라이버 위치

/opt/cloudera/parcels/CDH/lib/sqoop/lib

sqoop사용시 hdfs jdbc 드라이버 위치

/user/oozie/share/lib/lib\_타입스템프/sqoop

1. hue에서 oozie work-flow를 생성한다.

import --connect jdbc:mysql://cm.localdomain/test --username root   
--password kangjh --table test\_table --target-dir hdfs://name.localdomain:8020/user/admin/test\_table -m 1

1. 주의 할 점은 target-dir에 hue로 로그인한 계정이 쓰기 권한이 있어야 한다.

hive-server2 사용하기

Enabling Kerberos Authentication for HiveServer2

/etc/hive/conf/hive-site.xml file 수정

네임노드 히스토리서버 url 및 포트

http://name.localdomain:19888/jobhistory

네임노드 리소스매니저 url 및 포트

http://name.localdomain:8088/cluster

- CDH kerberos 적용후에는 webhdfs의 파일브라우저기능을 사용하지 못하고

hue의 파일브라우저 기능을 사용해야 한다.

- yarn-site.xml파일과 기타 설정파일의 위치 및 내용점검

http://www.cloudera.com/documentation/enterprise/5-2-x/topics/cdh\_ig\_yarn\_tuning.html

yarn.nodemanager.resource.memory-mb => 92106

호튼웍스 홈페이지에는 권고치는 리눅스 사용메모리를 1/6을 제외한

전체메모를 할당하라고 나와 있습니다.

yarn.scheduler.minimum-allocation-mb => 1024

mapreduce.map.memory.mb

mapreduce.reduce.memory.mb

mapreduce.map.java.opts

mapreduce.reduce.java.opts

각로그파일의 위치 파악하기

-CDH kerberos 적용후에는 webhdfs의 파일브라우저기능을 사용하지 못하고

hue의 파일브라우저 기능을 사용해야 한다.

-CDH 5.7.1의 하둡버전은 2.6 이다.

- 커널파라메터 수정

vi /etc/sysctl.conf

vm.swappiness=0(0이 권장값)

vm.overcommit\_memory=0(메모리값이 충분할경우 0을 권장)

- 하둡 디렉토리의 권한설정 확인

hadoop operation 89페이지 확인

- ntp 설정 확인

- 디스크설정 확인

하둡시스템은 LVM(Logical Volume Manager를 사용하면 절대 안 된다.

LVM은 여러 디바이스를 하나로 묶어 큰 디바이스를 구성한다.

그래서 디스크중 하나라도 장애가 발생하면 모두 망가진다.

df -k 명령어로 디스크를 확인했을때 /dev/sda가 아니고 /deb/vg\*이면 설정이

잘못된 것이다.