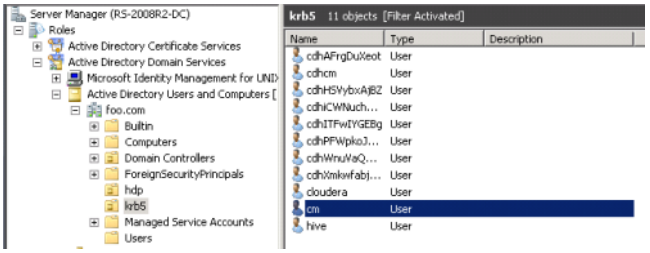
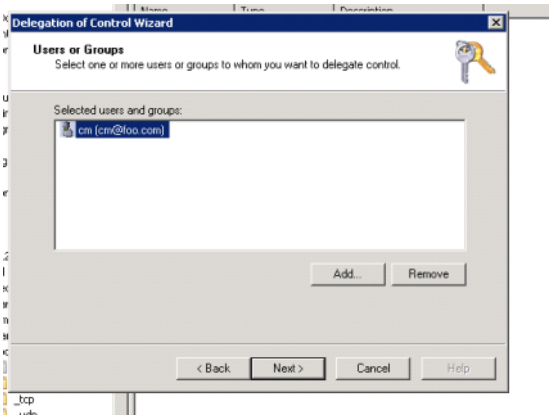
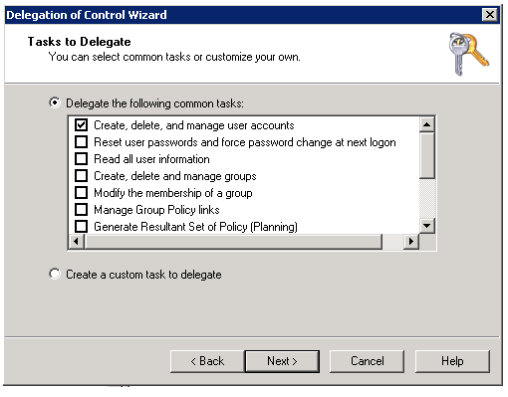
CDM_Isilon_KRB5_v1.0

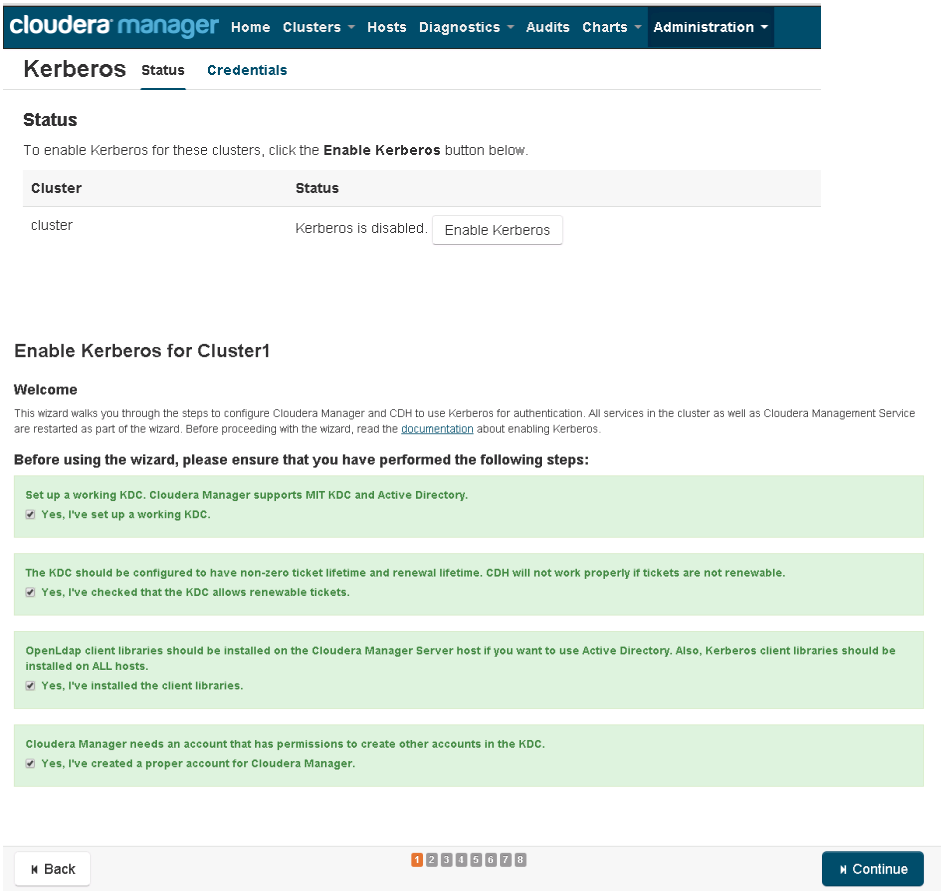
Tuesday, September 08, 2015
9:40 AM

russ stevenson
Cloudera Manager- Isilon Kerberos v1.0

http://www.cloudera.com/content/cloudera/en/documentation/core/latest/topics/cm_sg_intro_kerb.html

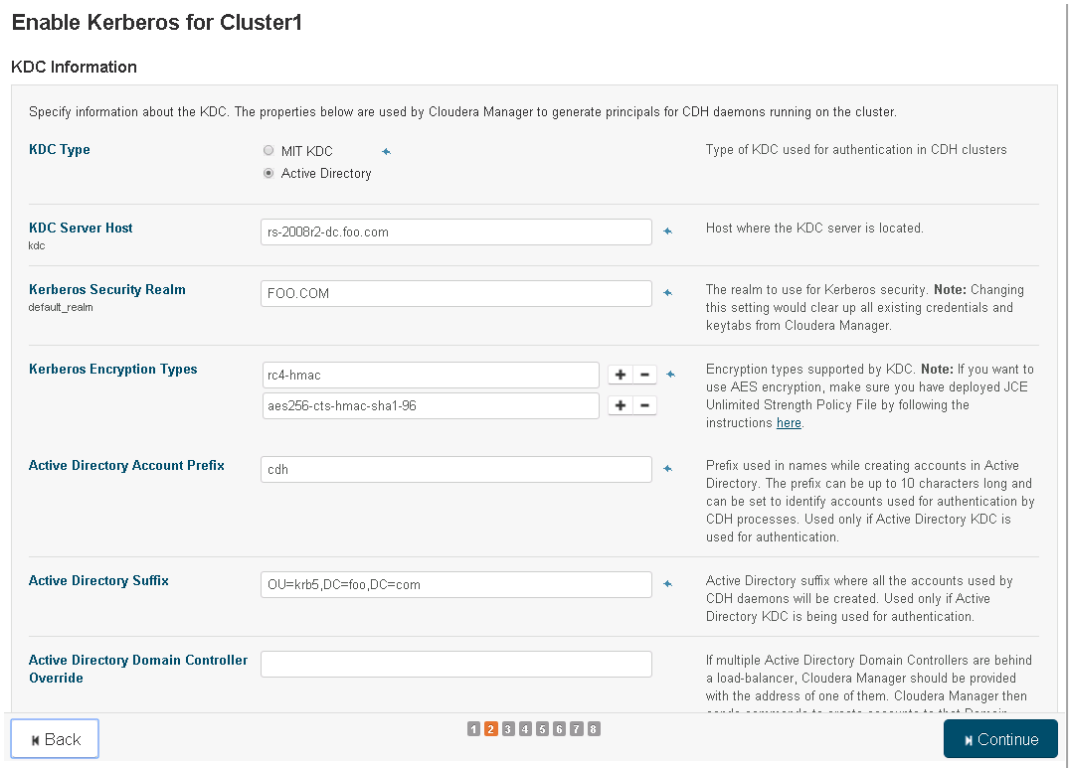
1. prep host for Kerberos if not done, see https://inside.emc.com/blogs/russ_stevenson/2015/12/16/finalizing-the-kerberization-of-a-linux-host for more info install prerequisites on linux hosts if needed to support Kerborized OS	https://technet.microsoft.com/en-us/library/bb742516.aspx yum install krb5-workstation yum install krb5-libs yum install openldap-clients If you are using Active Directory, make sure LDAP over SSL (LDAPS) is enabled for the Domain Controllers AD Certificate Services for LDAPS http://www.cloudera.com/content/cloudera/en/documentation/cloudera-manager/v5-1-x/Configuring-Hadoop-Security-with-Cloudera-Manager/cm5chs_s4_kerb_wizard.html
2. test kerberos from compute node kinit <hdfs account name> klist -e	[root@cdhcm ~]# kinit user1 Password for user1@FOO.COM: [root@cdhcm ~]# klist -e Ticket cache: FILE:/tmp/krb5cc_0 Default principal: user1@FOO.COM Valid starting Expires Service principal 09/08/15 14:57:46 09/09/15 00:57:47 krbtgt/FOO.COM@FOO.COM renew until 09/15/15 14:57:46, Etype (skey, tkt): des-cbc-crc, aes256-cts-hmac-sha1-96
3. if using Active Directory: Create an Organizational Unit (OU) in your AD setup where all the principals used by your CDH cluster will reside. eg: krb5 Add a new user account to Active Directory OU, for example, <username>@YOUR-REALM.COM. The password for this user should be set to never expire. eg: cm@foo.com Use AD's Delegate Control wizard to allow this new user to Create, Delete and Manage User Accounts.	   

4. start the kerberos wizard from Cloudera manager Validate the requirements

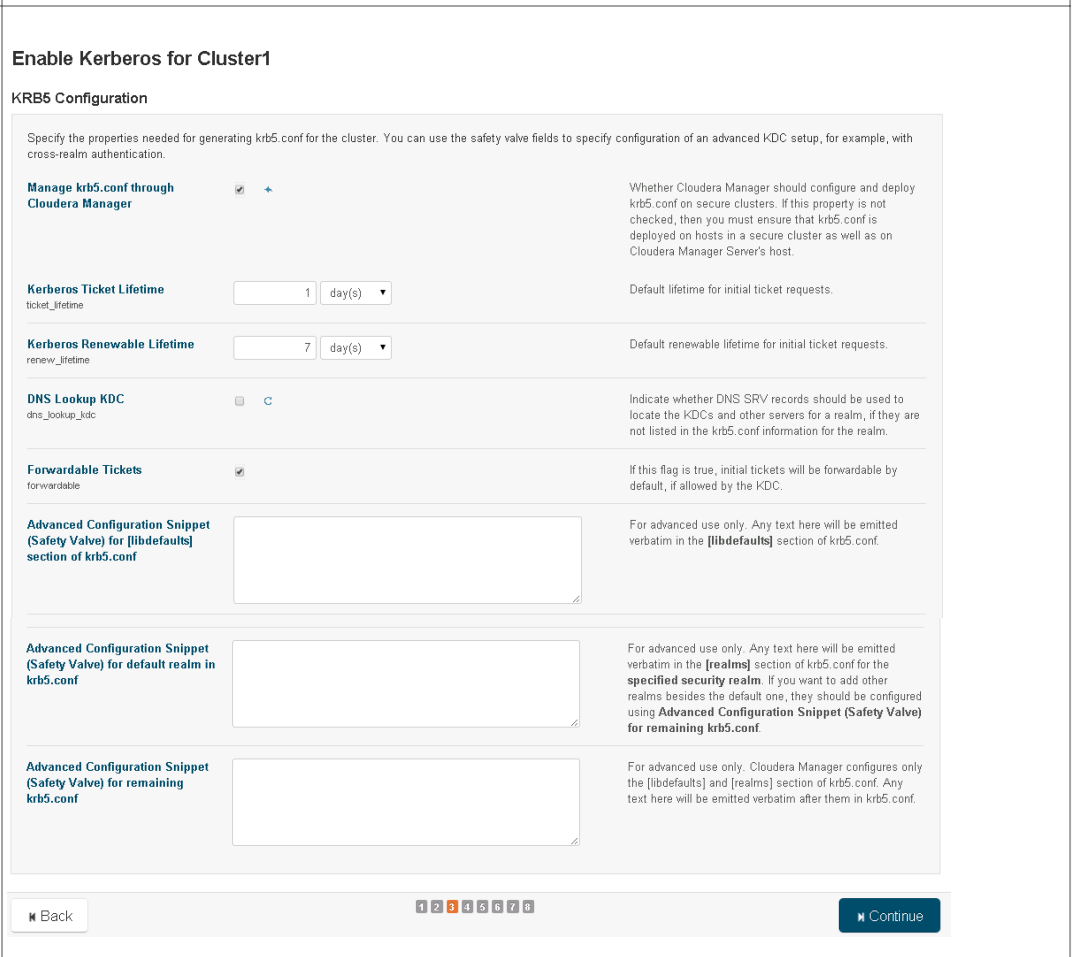


5. Complete the KDC information

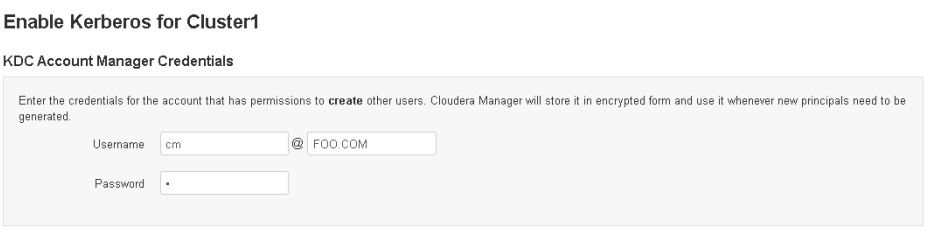
   set encryption type to rc4-hmac, (isilon 7.x no aes)



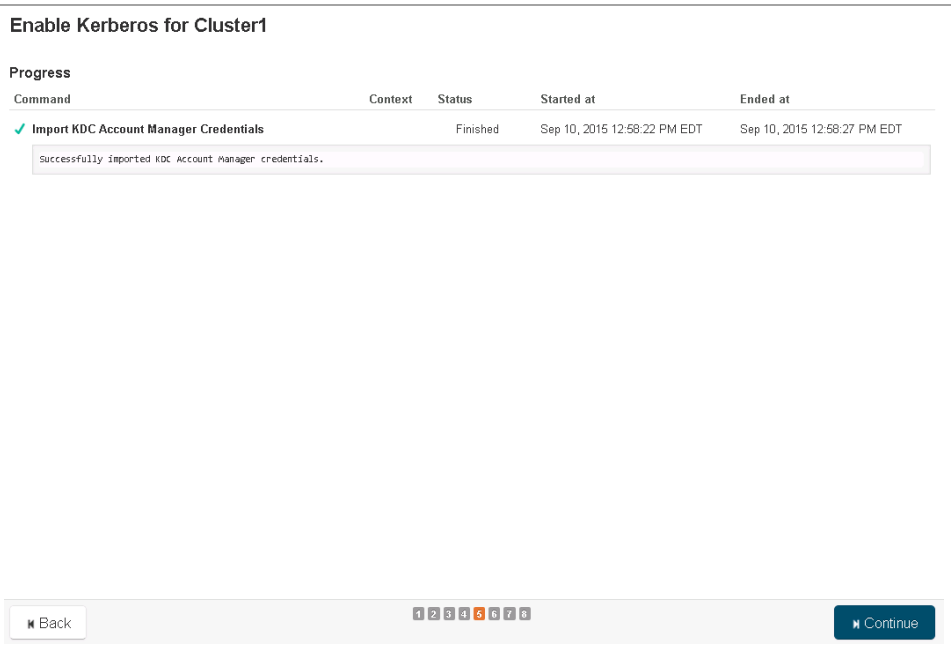
6. continue, let Cloudera Manager manage krb5.conf



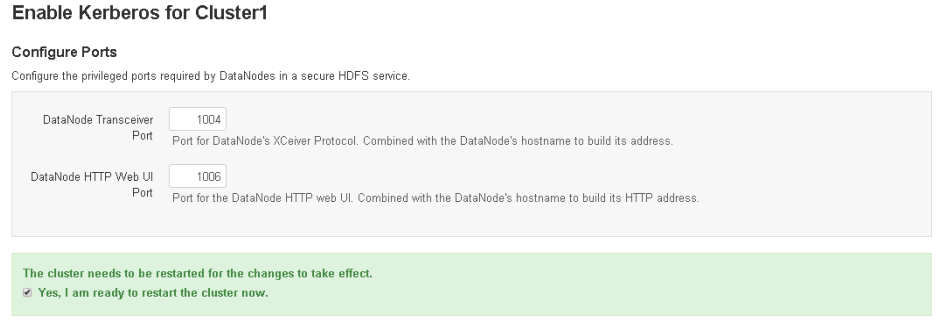
7. add the user created in AD for Cloudera Manager with delegated access, this account will create and manage all the CM principals in AD but not normal user accts.



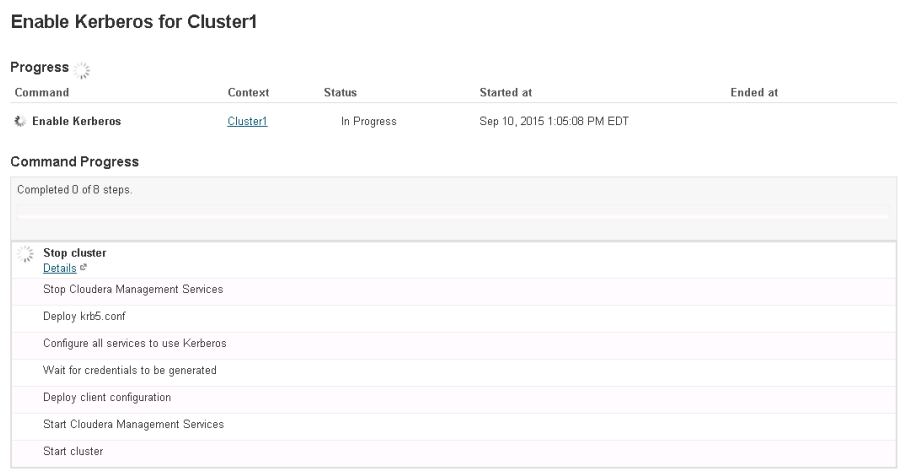
8. continue

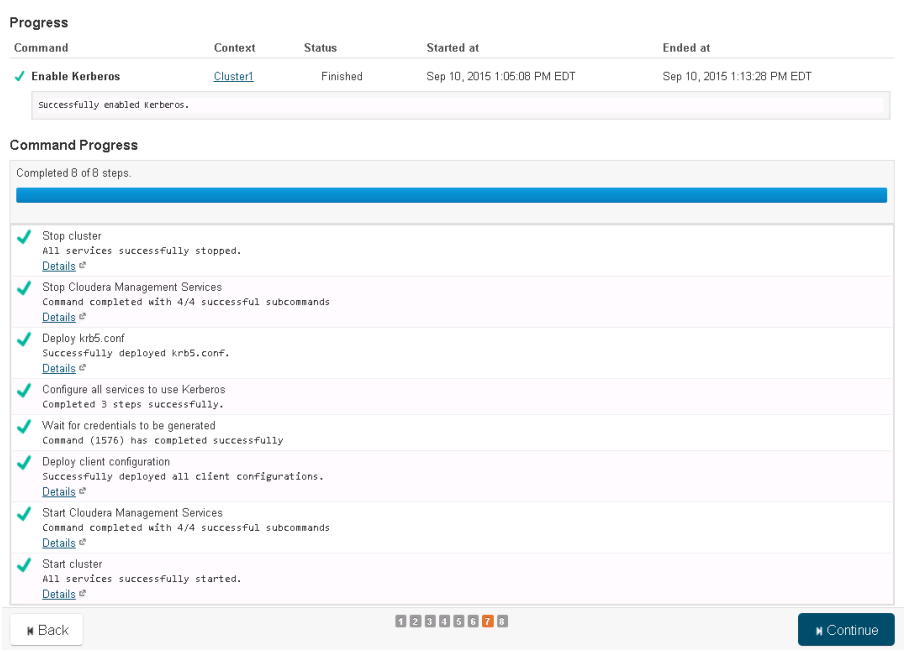


9. continue, leave as default

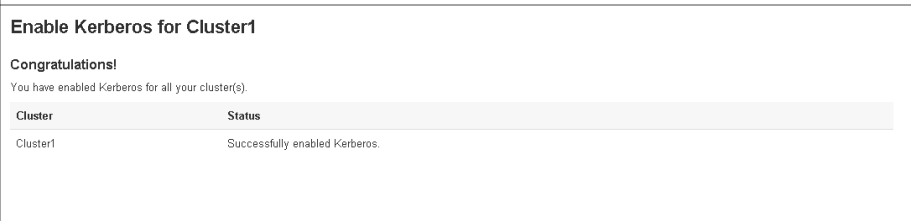


10. let the wizard complete



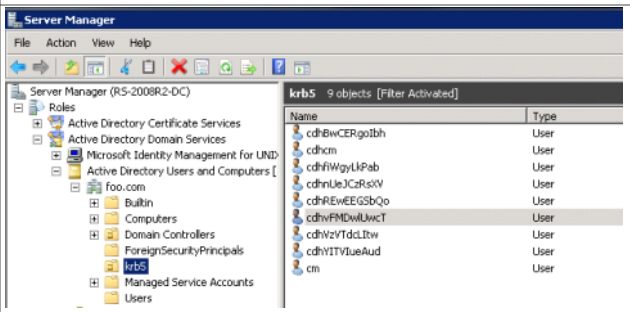


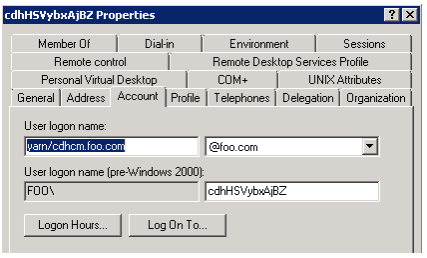
11. complete

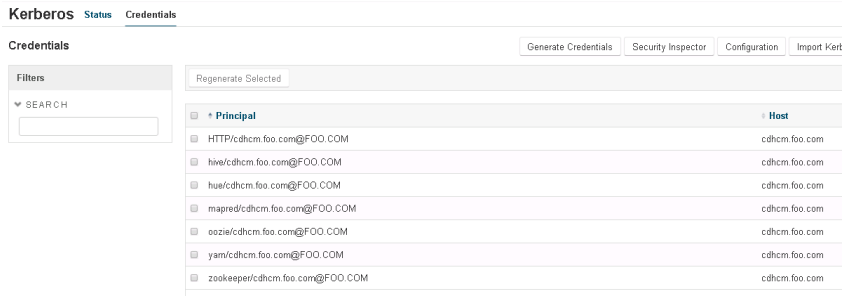


12. review AD, you will see Cloudera Manager has created all the required principles

    review the user principal names and how they match back up to cloudera manager







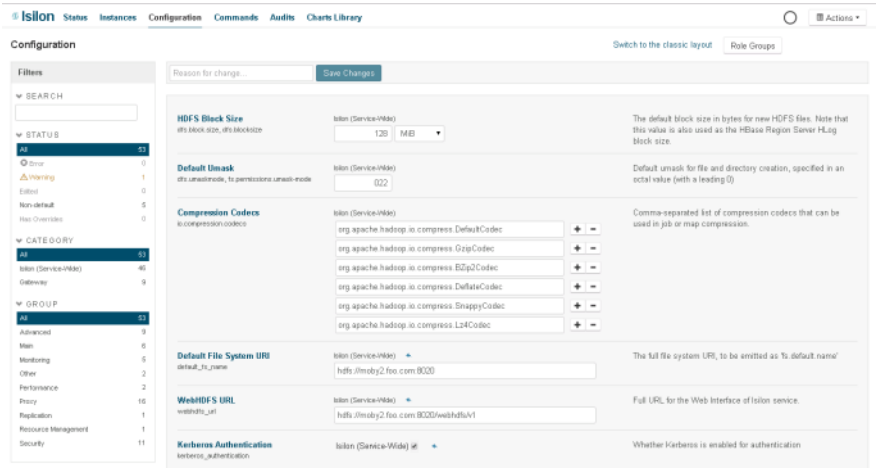
13. in Coudera Manager, review the isilon settings

    Add the following to:
    HDFS Client Advanced Configuration Snippet
    (Safety Valve) for hdfs-site.xml

    where:
    hdfs/moby2.foo.com@FOO.COM= hdfs/SCZonename@domain.com

    <property>
    <name>dfs.namenode.kerberos.principal</name>
    <value>hdfs/moby2.foo.com@FOO.COM</value>
    </property>
    <property>
    <name>dfs.datanode.kerberos.principal</name>
    <value>hdfs/moby2.foo.com@FOO.COM</value>
    </property>

Screen clipping taken: 12/23/2015 10:55 AM



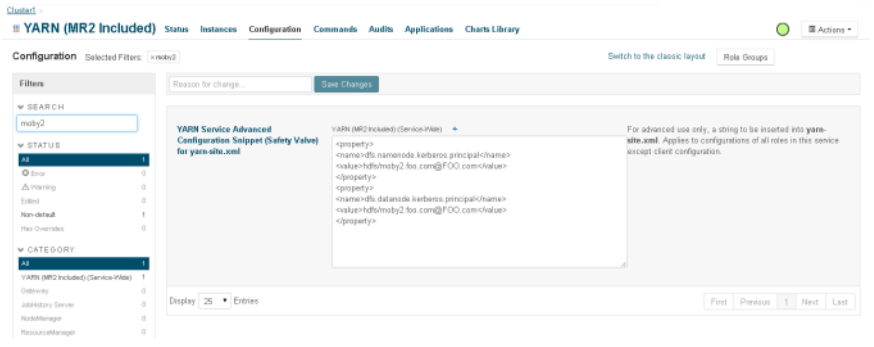
14. in Coudera Manager, review the yarn settings

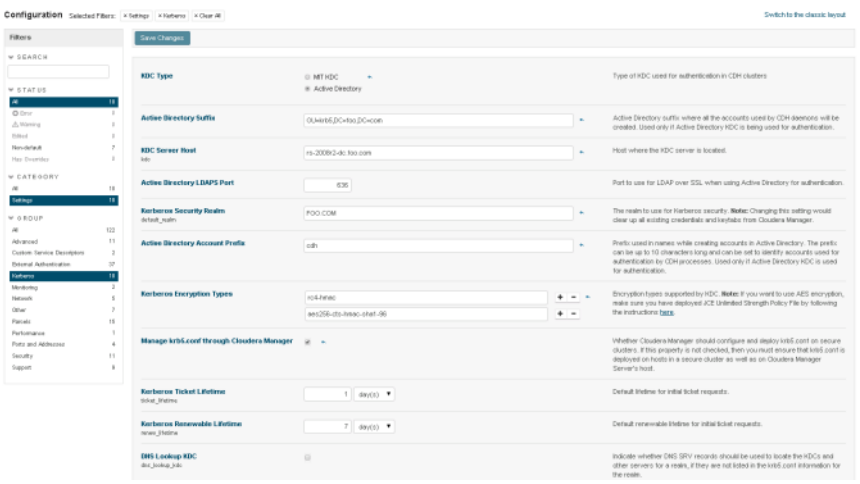
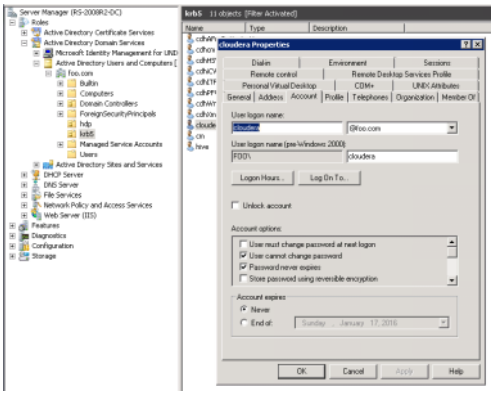
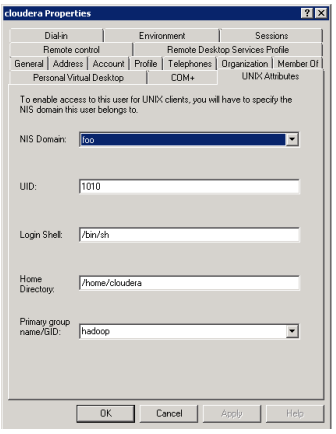
    Add the following to:
    HDFS Client Advanced Configuration Snippet
    (Safety Valve) for yarn-site.xml

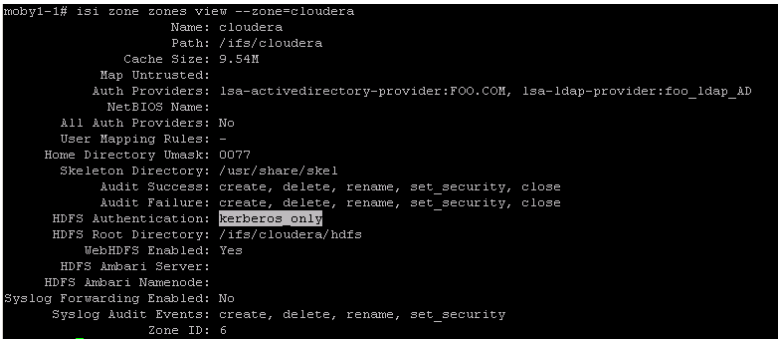
    where:
    hdfs/moby2.foo.com@FOO.COM= hdfs/SCZonename@domain.com

    yes, hdfs NOT yarn/

    <property>
    <name>dfs.namenode.kerberos.principal</name>
    <value>hdfs/moby2.foo.com@FOO.COM</value>
    </property>
    <property>
    <name>dfs.datanode.kerberos.principal</name>
    <value>hdfs/moby2.foo.com@FOO.COM</value>
    </property>



15. Review the Kerberos setting in Cloudera Manager	
16. In AD, create all the relevant users & group with the required UID & GID that will run jobs create a user cloudera to run hadoop jobs as give a uid & join to hadoop group, to set permissions as needed on isilon	 
17. Isilon Configuration valid for Kerborized Access	Validate Isilon for Kerberos SPN's are present and valid Zone Configuration is good users permissions < --- how to do this not covered, normal isilon permissions, use users and groups as need, now the users and groups coming from AD as Kerborized user, can use these identities to permissions hdfs root as needed
18. SPN - create hdfs/ spn's and additional	moby1-1# isi auth ads spn create --domain=foo.com --spn=hdfs/moby2.foo.com --user=foo\\administrator Enter password for administrator@FOO.COM: Successfully added SPN(s). Successfully added SPN(s). moby1-1# isi auth ads spn list --domain=foo.com nfs/moby2.foo.com hdfs/moby2.foo.com HOST/moby2.foo.com HOST/moby2

19. Set the Access Zone HDFS Authentication to: kerberos_only	
20. Configure and set the permissions on the hdfs root to be consistant with what is required to enabled the correct user access	<beyond scope of doc>
21. On compute node, create home directory for the kerboros AD user: cloudera	#mkdir /home/cloudera #chown 1010:root /home/cloudera/ using uid added into AD
22. become user on compute node	su - cloudera
23. set required java envt, add to .bash_profile	export JAVA_HOME=/usr/java/jdk1.7.0_67-cloudera export PATH=$JAVA_HOME/bin:$PATH vi .bash_profile
24. su to user and validate the java version	 [root@cdhcm ~]# **su - cloudera** -sh-4.1$ **bash** bash-4.1$ **java -version** bash: java: command not found bash-4.1$ **source .bash_profile** bash-4.1$ **java -version** java version "1.7.0_67" Java(TM) SE Runtime Environment (build 1.7.0_67-b01) Java HotSpot(TM) 64-Bit Server VM (build 24.65-b04, mixed mode)
25. kinit the user and validate user ticket and groups from AD	bash-4.1$ **kinit cloudera@FOO.COM** Password for cloudera@FOO.COM: bash-4.1$ **klist -e** Ticket cache: FILE:/tmp/krb5cc_1010 Default principal: cloudera@FOO.COM Valid starting Expires Service principal 09/15/15 18:49:29 09/16/15 04:49:26 krbtgt/FOO.COM@FOO.COM renew until 09/22/15 18:49:29, Etype (skey, tkt): arcfour-hmac, aes256-cts-hmac-sha1-96 bash-4.1$ **id** uid=1010(cloudera) gid=497(hadoop) groups=497(hadoop),10000(Domain Users)
26. run a job to validate	run a simple teragen job bash-4.1$ hadoop jar /opt/cloudera/parcels/CDH/lib/hadoop-mapreduce/hadoop-mapreduce-examples.jar teragen 10000 /teragenOUT2 15/12/28 17:14:53 INFO client.RMProxy: Connecting to ResourceManager at cdhcm.foo.com/172.16.201.100:8032 15/12/28 17:14:53 INFO hdfs.DFSClient: Created HDFS_DELEGATION_TOKEN token 67 for cloudera on 172.16.201.90:8020 15/12/28 17:14:53 INFO security.TokenCache: Got dt for hdfs://moby2.foo.com:8020; Kind: HDFS_DELEGATION_TOKEN, Service: 172.1 6.201.90:8020, Ident: (HDFS_DEL EGATION_TOKEN token 67 for cloudera) 15/12/28 17:14:54 INFO terasort.TeraSort: Generating 10000 using 2 15/12/28 17:14:54 INFO mapreduce.JobSubmitter: number of splits:2 15/12/28 17:14:54 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1448999039760_0004 15/12/28 17:14:54 INFO mapreduce.JobSubmitter: Kind: HDFS_DELEGATION_TOKEN, Service: 172.16.201.90:8020, Ident: (HDFS_DELEGAT ION_TOKEN token 67 for cloudera) 15/12/28 17:14:55 INFO impl.YarnClientImpl: Submitted application application_1448999039760_0004 15/12/28 17:14:55 INFO mapreduce.Job: The url to track the job: http://cdhcm.foo.com:8088/proxy/application_1448999039760_0004/ 15/12/28 17:14:55 INFO mapreduce.Job: Running job: job_1448999039760_0004 15/12/28 17:15:09 INFO mapreduce.Job: Job job_1448999039760_0004 running in uber mode : false 15/12/28 17:15:09 INFO mapreduce.Job: map 0% reduce 0% 15/12/28 17:15:16 INFO mapreduce.Job: map 50% reduce 0% 15/12/28 17:15:23 INFO mapreduce.Job: map 100% reduce 0% 15/12/28 17:15:23 INFO mapreduce.Job: Job job_1448999039760_0004 completed successfully 15/12/28 17:15:24 INFO mapreduce.Job: Counters: 31 File System Counters FILE: Number of bytes read=0 FILE: Number of bytes written=223316 FILE: Number of read operations=0 FILE: Number of large read operations=0 FILE: Number of write operations=0 HDFS: Number of bytes read=164 HDFS: Number of bytes written=1000000 HDFS: Number of read operations=8 HDFS: Number of large read operations=0 HDFS: Number of write operations=4 Job Counters Launched map tasks=2 Other local map tasks=2 Total time spent by all maps in occupied slots (ms)=10877 Total time spent by all reduces in occupied slots (ms)=0 Total time spent by all map tasks (ms)=10877 Total vcore-seconds taken by all map tasks=10877 Total megabyte-seconds taken by all map tasks=11138048 Map-Reduce Framework Map input records=10000

```
        Map output records=10000
        Input split bytes=164
        Spilled Records=0
        Failed Shuffles=0
        Merged Map outputs=0
        GC time elapsed (ms)=124
        CPU time spent (ms)=1940
        Physical memory (bytes) snapshot=358191104
        Virtual memory (bytes) snapshot=3055304704
        Total committed heap usage (bytes)=377487360
    org.apache.hadoop.examples.terasort.TeraGen$Counters
        CHECKSUM=21555350172850
    File Input Format Counters
        Bytes Read=0
    File Output Format Counters
        Bytes Written=1000000
```

27. with no valid kerberos ticket, job fails

```
bash-4.1$ kdestroy

bash-4.1$ hadoop jar /opt/cloudera/parcels/CDH/lib/hadoop-mapreduce/hadoop-mapreduce-examples.jar teragen 10000 /teragenOUT
15/12/28 17:16:34 WARN security.UserGroupInformation: PriviledgedActionException as:cloudera (auth:KERBEROS) cause:javax.secu rity.sasl.SaslException: GSS initiate failed
[Caused by GSSException: No valid credentials provided (Mechanism level: Failed to find any Kerberos tgt)]
15/12/28 17:16:34 WARN ipc.Client: Exception encountered while connecting to the server : javax.security.sasl.SaslException:  GSS initiate failed [Caused by GSSException: No valid credentials provided (Mechanism level: Failed to find any Kerberos tgt)]
15/12/28 17:16:34 WARN security.UserGroupInformation: PriviledgedActionException as:cloudera (auth:KERBEROS) cause:java.io.IO Exception: javax.security.sasl.SaslException:
GSS initiate failed [Caused by GSSException: No valid credentials provided (Mechanism level: Failed to find any Kerberos tgt) ]
java.io.IOException: Failed on local exception: java.io.IOException: javax.security.sasl.SaslException: GSS initiate failed [ Caused by GSSException: No valid credentials provided
(Mechanism level: Failed to find any Kerberos tgt)]; Host Details : local host is: "cdhcm.foo.com/172.16.201.100"; destinatio n host is: "moby2.foo.com":8020;
    at org.apache.hadoop.net.NetUtils.wrapException(NetUtils.java:772)
    at org.apache.hadoop.ipc.Client.call(Client.java:1472)
    at org.apache.hadoop.ipc.Client.call(Client.java:1399)
    at org.apache.hadoop.ipc.ProtobufRpcEngine$Invoker.invoke(ProtobufRpcEngine.java:232)
    at com.sun.proxy.$Proxy14.getFileInfo(Unknown Source)
    at org.apache.hadoop.hdfs.protocolPB.ClientNamenodeProtocolTranslatorPB.getFileInfo(ClientNamenodeProtocolTranslatorPB.java:7 52)
    at sun.reflect.NativeMethodAccessorImpl.invoke0(Native Method)
    at sun.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorImpl.java:57)
    at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImpl.java:43)
    at java.lang.reflect.Method.invoke(Method.java:606)
    at org.apache.hadoop.io.retry.RetryInvocationHandler.invokeMethod(RetryInvocationHandler.java:187)
    at org.apache.hadoop.io.retry.RetryInvocationHandler.invoke(RetryInvocationHandler.java:102)
    at com.sun.proxy.$Proxy15.getFileInfo(Unknown Source)
    at org.apache.hadoop.hdfs.DFSClient.getFileInfo(DFSClient.java:1982)
    at org.apache.hadoop.hdfs.DistributedFileSystem$18.doCall(DistributedFileSystem.java:1128)
    at org.apache.hadoop.hdfs.DistributedFileSystem$18.doCall(DistributedFileSystem.java:1124)
    at org.apache.hadoop.fs.FileSystemLinkResolver.resolve(FileSystemLinkResolver.java:81)
    at org.apache.hadoop.hdfs.DistributedFileSystem.getFileStatus(DistributedFileSystem.java:1124)
    at org.apache.hadoop.fs.FileSystem.exists(FileSystem.java:1400)
    at org.apache.hadoop.examples.terasort.TeraGen.run(TeraGen.java:292)
    at org.apache.hadoop.util.ToolRunner.run(ToolRunner.java:70)
    at org.apache.hadoop.examples.terasort.TeraGen.main(TeraGen.java:309)
    at sun.reflect.NativeMethodAccessorImpl.invoke0(Native Method)
    at sun.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorImpl.java:57)
    at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImpl.java:43)
    at java.lang.reflect.Method.invoke(Method.java:606)
    at org.apache.hadoop.util.ProgramDriver$ProgramDescription.invoke(ProgramDriver.java:71)
    at org.apache.hadoop.util.ProgramDriver.run(ProgramDriver.java:144)
    at org.apache.hadoop.examples.ExampleDriver.main(ExampleDriver.java:74)
    at sun.reflect.NativeMethodAccessorImpl.invoke0(Native Method)
    at sun.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorImpl.java:57)
    at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImpl.java:43)
    at java.lang.reflect.Method.invoke(Method.java:606)
    at org.apache.hadoop.util.RunJar.run(RunJar.java:221)
    at org.apache.hadoop.util.RunJar.main(RunJar.java:136)
Caused by: java.io.IOException: javax.security.sasl.SaslException: GSS initiate failed [Caused by GSSException: No valid cred  entials provided (Mechanism level: Failed to find any
Kerberos tgt)]
    at org.apache.hadoop.ipc.Client$Connection$1.run(Client.java:680)
    at java.security.AccessController.doPrivileged(Native Method)
    at javax.security.auth.Subject.doAs(Subject.java:415)
    at org.apache.hadoop.security.UserGroupInformation.doAs(UserGroupInformation.java:1671)
    at org.apache.hadoop.ipc.Client$Connection.handleSaslConnectionFailure(Client.java:643)
    at org.apache.hadoop.ipc.Client$Connection.setupIOstreams(Client.java:730)
    at org.apache.hadoop.ipc.Client$Connection.access$2800(Client.java:368)
    at org.apache.hadoop.ipc.Client.getConnection(Client.java:1521)
    at org.apache.hadoop.ipc.Client.call(Client.java:1438)
    ... 33 more
Caused by: javax.security.sasl.SaslException: GSS initiate failed [Caused by GSSException: No valid credentials provided (Mec hanism level: Failed to find any Kerberos tgt)]
    at com.sun.security.sasl.gsskerb.GssKrb5Client.evaluateChallenge(GssKrb5Client.java:212)
    at org.apache.hadoop.security.SaslRpcClient.saslConnect(SaslRpcClient.java:413)
    at org.apache.hadoop.ipc.Client$Connection.setupSaslConnection(Client.java:553)
    at org.apache.hadoop.ipc.Client$Connection.access$1800(Client.java:368)
    at org.apache.hadoop.ipc.Client$Connection$2.run(Client.java:722)
    at org.apache.hadoop.ipc.Client$Connection$2.run(Client.java:718)
    at java.security.AccessController.doPrivileged(Native Method)
    at javax.security.auth.Subject.doAs(Subject.java:415)
    at org.apache.hadoop.security.UserGroupInformation.doAs(UserGroupInformation.java:1671)
    at org.apache.hadoop.ipc.Client$Connection.setupIOstreams(Client.java:717)
    ... 36 more
Caused by: GSSException: No valid credentials provided (Mechanism level: Failed to find any Kerberos tgt)
    at sun.security.jgss.krb5.Krb5InitCredential.getInstance(Krb5InitCredential.java:147)
    at sun.security.jgss.krb5.Krb5MechFactory.getCredentialElement(Krb5MechFactory.java:121)
    at sun.security.jgss.krb5.Krb5MechFactory.getMechanismContext(Krb5MechFactory.java:187)
    at sun.security.jgss.GSSManagerImpl.getMechanismContext(GSSManagerImpl.java:223)
    at sun.security.jgss.GSSContextImpl.initSecContext(GSSContextImpl.java:212)
    at sun.security.jgss.GSSContextImpl.initSecContext(GSSContextImpl.java:179)
    at com.sun.security.sasl.gsskerb.GssKrb5Client.evaluateChallenge(GssKrb5Client.java:193)
    ... 45 more
```