ISILON ONEFS WITH AMBARI MULTITENANT ACTIVE DIRECTORY

Integration Guide

ABSTRACT

The following whitepaper outlines the implementation approach and considerations that are required to implement multiple Isilon-connected Ambari instances against a single Active Directory.

17 May 2018

WHITE PAPER

Publication History

Version	Date	Description
1.00	17 May 2018	Initial version.

The information in this publication is provided “as is.” Dell Inc. makes no representations or warranties of any kind with respect to the information in this publication, and specifically disclaims implied warranties of merchantability or fitness for a particular purpose.

Use, copying, and distribution of any software described in this publication requires an applicable software license.

Dell believes the information in this document is accurate as of its publication date. The information is subject to change without notice.

TABLE OF CONTENTS

Executive Summary

It is well documented that Isilon OneFS is the enterprise solution for multitenant Hadoop cluster support and integration. This paper will illustrate the approaches and configuration integrations required to support secure multitenant Hadoop clusters with a single Isilon OneFS cluster against a single Active Directory domain. The goal is to highlight the design decisions and implementation modifications that are required to successfully deploy multitenancy in a single domain. It is critical to assess the deployment of multiple Hadoop clusters against a single Isilon cluster before the Hadoop cluster is deployed or secured, as the deployment process may require modification to ensure a successful and supported integration. It is also suggested that this approach is used even if no multitenancy is to be implemented, as it provides isolation of user accounts regardless of how many clusters will be deployed. The use of Isilon-based mapping rules will simplify the deployment of Ambari-based HDP Kerberos deployments.

Audience

This guide is intended for Hadoop systems administrators, storage administrators, IT architects, and IT managers who will be running Isilon OneFS with Cloudera CDH or Ambari Hortonworks HDP-based Hadoop distributions.

Introduction

When implementing a Hadoop cluster with Isilon OneFS, you need to make some initial decisions regarding how access control and management will be implemented. OneFS supports integrating multiple approaches to Hadoop security. The focus of this whitepaper is to address multitenant-specific implementations and the different approaches that are needed when designing and setting up Isilon and the Hadoop cluster. This document does not address the specific procedure of setting up Hadoop – Isilon security, as you can read about those procedures in the following installation guides: Isilon and Hadoop Cluster Install Guides.

This document is an addition to the following Isilon implementation guides, which review and highlight the approaches to Isilon and Hadoop Kerberos and Identity Management. Both should be reviewed for additional information on multitenant integrations.

- http://www.emc.com/collateral/white-papers/docu87984-isilon-hadoop-kerberos-identity-management.pdf
- http://www.emc.com/collateral/white-papers/multitenant-installation-and-integration-guide.pdf

The goal of this document is to specifically illustrate the implementation of multiple Ambari-based HDP clusters against a single Active Directory with Isilon OneFS. Although the approach can be used with a KDC-based Kerberos deployment, that process is not illustrated in this document.

Hortonworks – Ambari

Kerberization for Ambari and HDP services is usually deployed using the **Cluster Kerberization** wizard that completes many of the steps needed to Kerberize the cluster. The same general approach as is illustrated in the EMC Isilon OneFS with Hadoop and Hortonworks for Kerberos Installation Guide should be followed to deploy and set up the clusters with Kerberization.

This document will outline the modifications required to allow a multitenant installation.

Requirements

The following requirements are still valid for the integration of Isilon and Hadoop clusters:

- Use an Access Zone and Authentication Providers on each HDP cluster install
- A per Access Zone Local Provider is used for Service Accounts identities
- User Account UPNs are created in the KDC by the wizard
- All Hadoop service account SPNs are unique and machine-specific, which provides uniqueness for all Hadoop cluster services
- Per Access Zone SmartConnect SPNs are required in the KDC; SPNs for Isilon OneFS must meet Isilon provider requirements

Note:

These behaviors should be well understood when Kerberizing Hortonworks clusters against Isilon OneFS.

The primary challenge with Ambari Kerberized clusters with Isilon is the integration of the Identity Management, as well as how Ambari, Active Directory, and Isilon all handle the Ambari-created Smoketest user UPNs specifically. With multitenant deployments, how we architect and modify each deployment for the Ambari auto-generated smoke test users (for example, hdfs, ambari-qa, spark, storm, hbase etc.) is the critical factor in a successful deployment.

Shared Active Directory between multiple Ambari HDP Hadoop clusters

In order to deploy multiple Hadoop clusters with a single shared Active Directory, we have to account for the Ambari smoke test user UPNs. In a standard Isilon integration, we would modify the Ambari-based smoke test user UPNs and map these with the SAMAccount names which correspond to the Isilon service account users as shown in the installation guides. However, in order to support multiple Ambari smoke test user UPNs in the same Active Directory, we need to modify our deployment approach to support the creation of unique UPNs with a corresponding unique SAMAccount name and then map these back to Isilon for Identity management. You have the following options to enable this type of deployment:

Option 1: Use the '-clustername' suffix based Ambari UPNs with Isilon mapping rules [the suggested and recommended approach]

Or

Option 2: Install using custom usernames and Kerberize as described in the installation guides (not described in this document)

Mapping rules with Ambari-appended ClusterName suffix to UPN approach – Option 2

The recommended approach is to leverage Ambari's own internal mapping rules to append a clustername suffix to the UPN with an Isilon-specific mapping rule to map the modified SAMAccount name back to the Isilon identity as follows:

- Deploy each Ambari Hortonworks HDP with standard accounts, for example: hdfs, yarn, ambari-qa
- Use a per Access Zone Isilon deployment model
- Deploy Isilon with standard accounts, for example: hdfs, yarn, ambari-qa, hbase, storm
- Kerberize using the standard wizard; allow the wizard to `append cluster_name|toLower()` to Ambari UPNs*
- Add new Mapping Rules to Isilon OneFS

*This is where the process deviates from prior installation methods

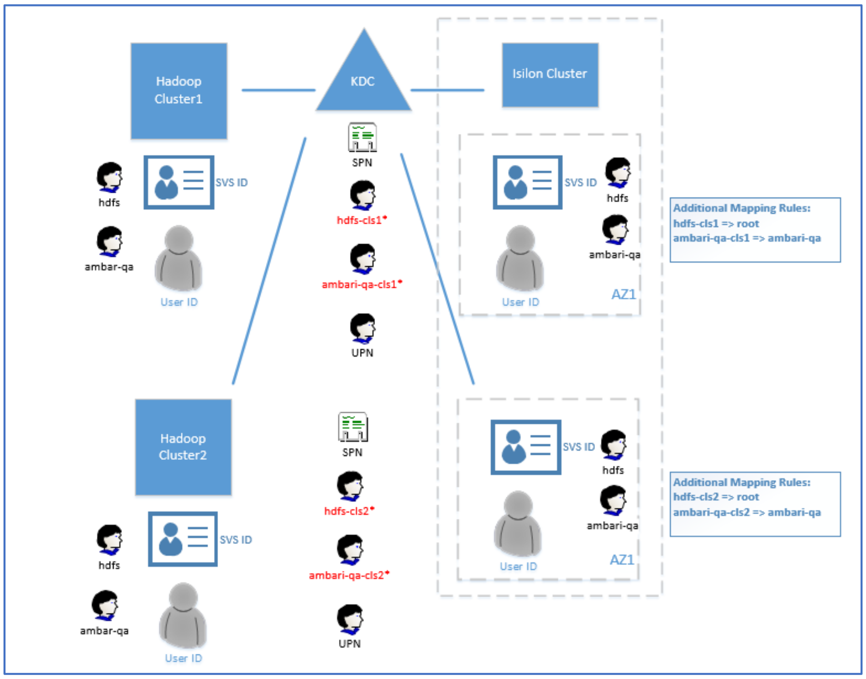
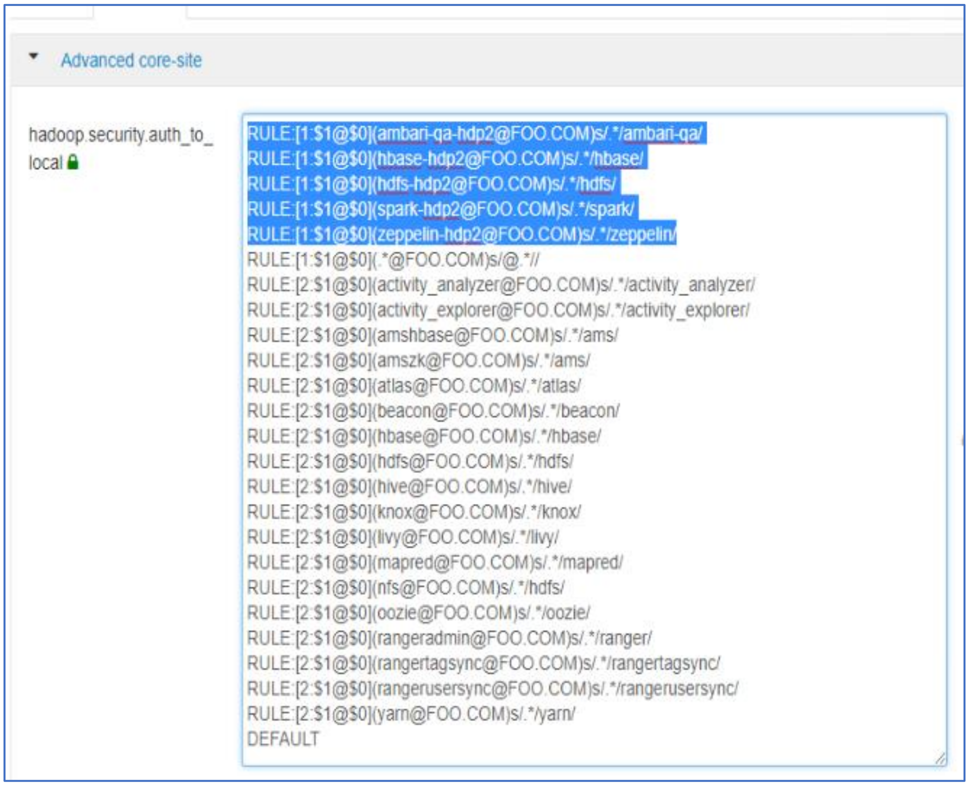


Figure 1. Additional mapping rules to support clustername Ambari UPNs

This approach allows the existing install to integrate into a shared AD/KDC without having to reinstall or use custom user accounts.

Ambari already does something similar with its own local user accounts, where it maps the local user to the appropriate principal as shown below:



hadoop.security.auth_to_local

```
RULE:[1:$1@$0](ambari-qa-hdp2@FOO.COM)s/.*/ambari-qa/
RULE:[1:$1@$0](hbase-hdp2@FOO.COM)s/.*/hbase/
RULE:[1:$1@$0](hdfs-hdp2@FOO.COM)s/.*/hdfs/
RULE:[1:$1@$0](spark-hdp2@FOO.COM)s/.*/spark/
RULE:[1:$1@$0](zeppelin-hdp2@FOO.COM)s/.*/zeppelin/
RULE:[1:$1@$0](.*@FOO.COM)s/@.*//
RULE:[2:$1@$0](activity_analyzer@FOO.COM)s/.*/activity_analyzer/
RULE:[2:$1@$0](activity_explorer@FOO.COM)s/.*/activity_explorer/
RULE:[2:$1@$0](amshbase@FOO.COM)s/.*/ams/
RULE:[2:$1@$0](amszk@FOO.COM)s/.*/ams/
RULE:[2:$1@$0](atlas@FOO.COM)s/.*/atlas/
RULE:[2:$1@$0](beacon@FOO.COM)s/.*/beacon/
RULE:[2:$1@$0](hbase@FOO.COM)s/.*/hbase/
RULE:[2:$1@$0](hdfs@FOO.COM)s/.*/hdfs/
RULE:[2:$1@$0](hive@FOO.COM)s/.*/hive/
RULE:[2:$1@$0](knox@FOO.COM)s/.*/knox/
RULE:[2:$1@$0](livy@FOO.COM)s/.*/livy/
RULE:[2:$1@$0](mapred@FOO.COM)s/.*/mapred/
RULE:[2:$1@$0](nfs@FOO.COM)s/.*/hdfs/
RULE:[2:$1@$0](oozie@FOO.COM)s/.*/oozie/
RULE:[2:$1@$0](rangeradmin@FOO.COM)s/.*/ranger/
RULE:[2:$1@$0](rangertagsync@FOO.COM)s/.*/rangertagsync/
RULE:[2:$1@$0](rangerusersync@FOO.COM)s/.*/rangerusersync/
RULE:[2:$1@$0](yarn@FOO.COM)s/.*/yarn/
DEFAULT
```

Figure 2. Ambari local service account to Kerberos Principals mapping

Implementation

To facilitate this approach, additional mapping rules are required to map the Ambari UPNs back to the standard service account names with which the cluster was initially installed. As mentioned earlier in this whitepaper, Ambari already has similar rules; by adding them to Isilon, we can operate in a similar manner.

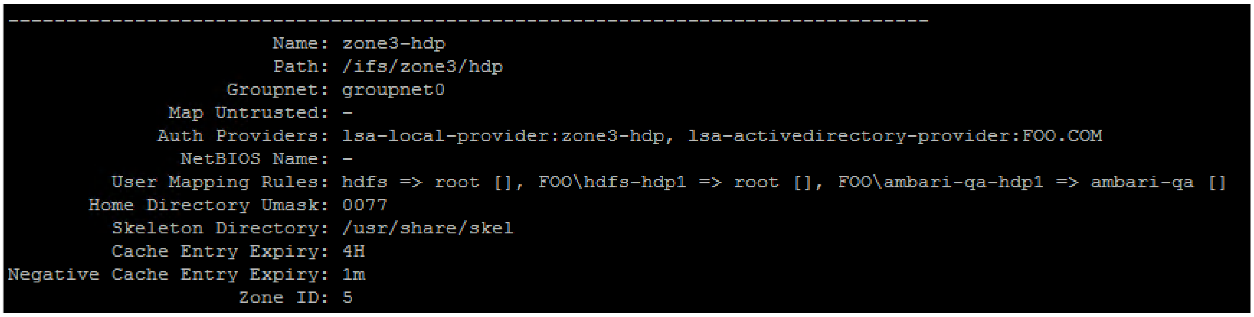
The table below outlines the UPNs that need mapping rules added to Isilon; if other services are used, additional rules should be added to map those accounts as well.

Hortonworks Service Account	AD/KDC principal without removing `cluster_name|toLower()*`	SAMAccount Name *modified	Isilon Service Account	Mapping Rule Needed
hdfs	hdfs-hdp1	hdfs-hdp1	hdfs	<domain>\hdfs-hdp1 => root []
ambari-qa	ambari-qa-hdp1	ambari-qa-hdp1	ambari-qa	<domain>\\ambari-qa-hdp1 => ambari-qa []

Table 1. Ambari UPNs and required mapping rules – clustername hdp1

The additional Ambari-generated smoke user UPNs may be present and rules should be added to Isilon OneFS if they are in use.

- hbase-clustername
- spark-clustername
- storm-clustername
- zeppelin-clustername



```
--------------------------------------------------------------------------------
                       Name: zone3-hdp
                       Path: /ifs/zone3/hdp
                   Groupnet: groupnet0
              Map Untrusted: -
             Auth Providers: lsa-local-provider:zone3-hdp, lsa-activedirectory-provider:FOO.COM
               NetBIOS Name: -
         User Mapping Rules: hdfs => root [], FOO\hdfs-hdp1 => root [], FOO\ambari-qa-hdp1 => ambari-qa []
       Home Directory Umask: 0077
         Skeleton Directory: /usr/share/skel
         Cache Entry Expiry: 4H
Negative Cache Entry Expiry: 1m
                    Zone ID: 5
```

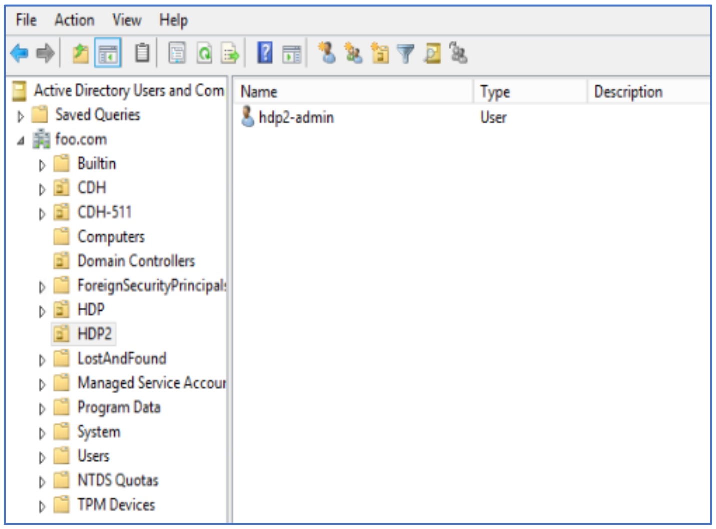
Figure 3. Additional mapping rules required on Isilon

As stated earlier, the general approach still follows the procedure as described in the EMC Isilon OneFS with Hadoop and Hortonworks for Kerberos Installation Guide.

The following steps illustrate the modifications and additional configurations required to support multitenant AD deployments. In the following deployment, we will Kerberize the HDP2 Ambari cluster; HDP1 is already Kerberized against AD using the same approach.

Multitenant AD deployment

1. Create the dedicated cluster OU in the AD and add the cluster-specific admin account for Ambari,

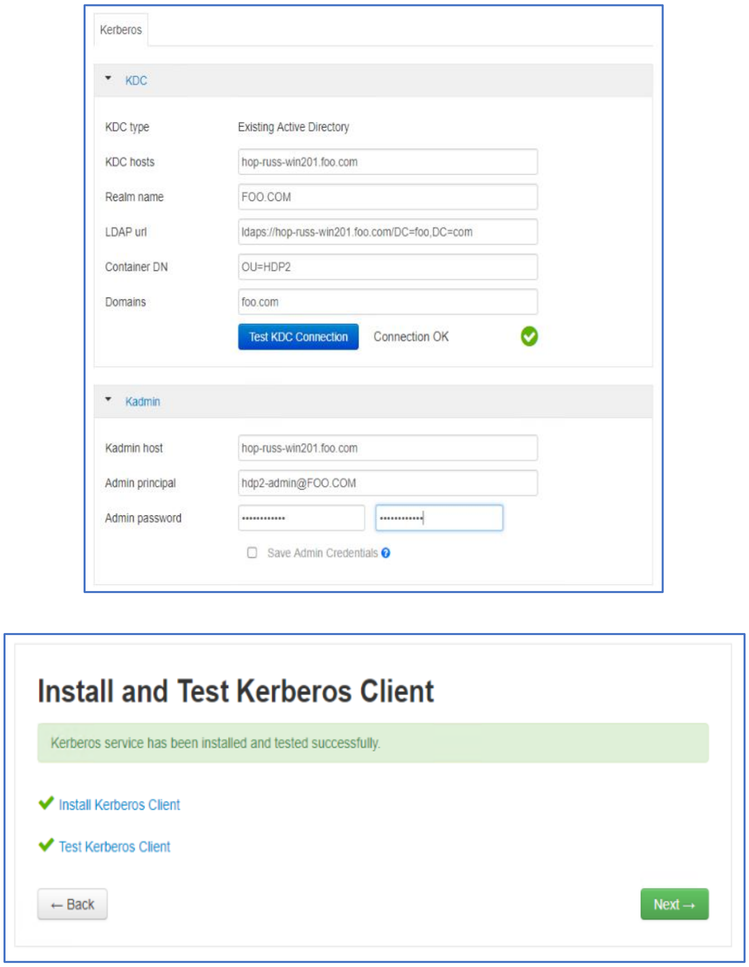


- Active Directory Users and Com
  - Saved Queries
  - foo.com
    - Builtin
    - CDH
    - CDH-511
    - Computers
    - Domain Controllers
    - ForeignSecurityPrincipal:
    - HDP
    - HDP2
    - LostAndFound
    - Managed Service Accour
    - Program Data
    - System
    - Users
    - NTDS Quotas
    - TPM Devices

Name	Type	Description
hdp2-admin	User	

Figure 4. Active Directory OU ready for Ambari Kerberization

2. Execute the Kerberos wizard and configure against the same Active Directory,



KDC

KDC type	Existing Active Directory
KDC hosts	hop-russ-win201.foo.com
Realm name	FOO.COM
LDAP url	ldaps://hop-russ-win201.foo.com/DC=foo,DC=com
Container DN	OU=HDP2
Domains	foo.com

Connection OK ✓

Kadmin

Kadmin host	hop-russ-win201.foo.com
Admin principal	hdp2-admin@FOO.COM
Admin password	••••••••••• •••••••••••

□ Save Admin Credentials

Install and Test Kerberos Client

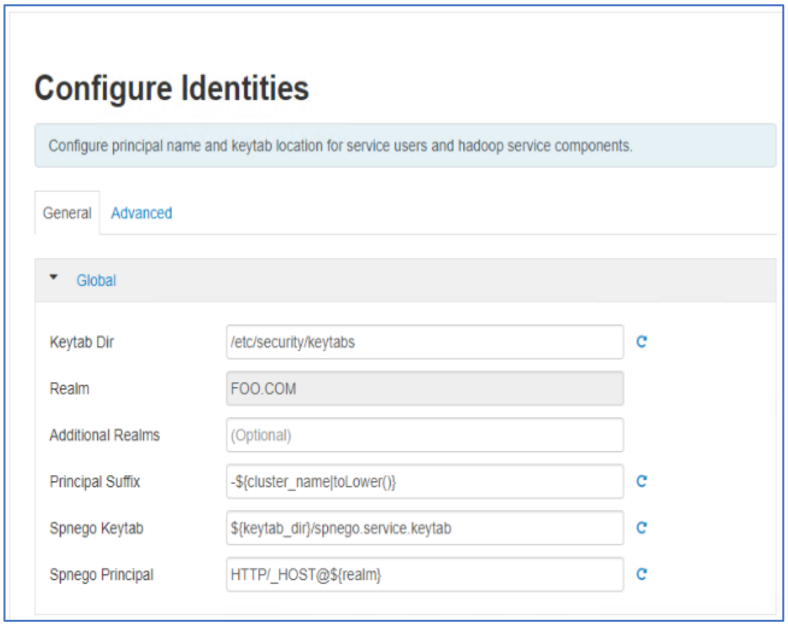
Kerberos service has been installed and tested successfully.

✓ Install Kerberos Client

✓ Test Kerberos Client

Figure 5. The Ambari Kerberos wizard

3. This step is where we deviate from the standard Kerberos Installation guide, leaving the Global and Ambari Principals fields as they are. By leaving the -${cluster_name|toLower()} as the Principal Suffix, Ambari will append the cluster name to any UPN created. This will create a unique UPN in Active Directory to facilitate multitenancy. This is the key difference required to facilitate multitenancy within AD with Isilon.



Configure Identities

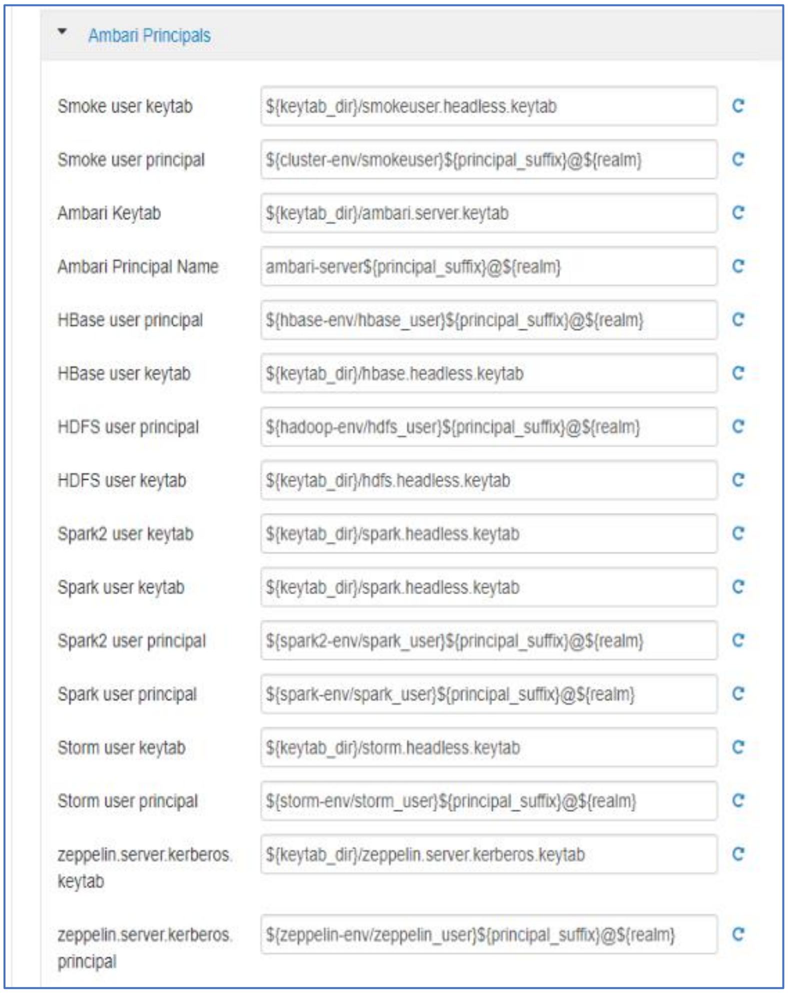
Configure principal name and keytab location for service users and hadoop service components.

General | Advanced

Global

Keytab Dir	/etc/security/keytabs
Realm	FOO.COM
Additional Realms	(Optional)
Principal Suffix	-${cluster_name|toLower()}
Spnego Keytab	${keytab_dir}/spnego.service.keytab
Spnego Principal	HTTP/_HOST@${realm}

Validate the Principal Suffix per the note below before proceeding.

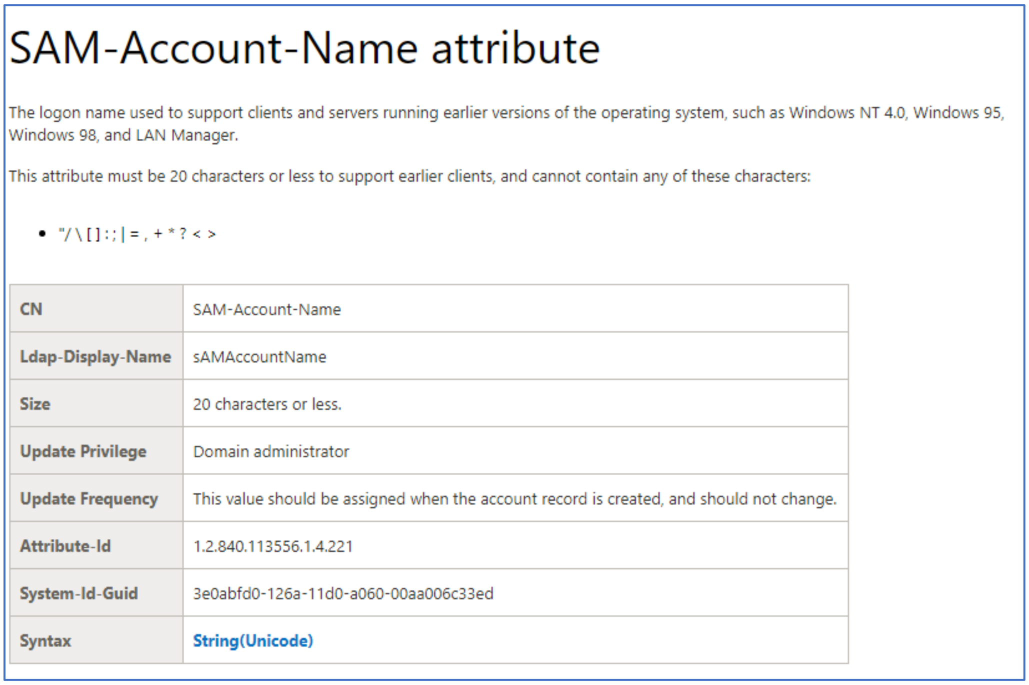


Ambari Principals

Smoke user keytab	${keytab_dir}/smokeuser.headless.keytab
Smoke user principal	${cluster-env/smokeuser}${principal_suffix}@${realm}
Ambari Keytab	${keytab_dir}/ambari.server.keytab
Ambari Principal Name	ambari-server${principal_suffix}@${realm}
HBase user principal	${hbase-env/hbase_user}${principal_suffix}@${realm}
HBase user keytab	${keytab_dir}/hbase.headless.keytab
HDFS user principal	${hadoop-env/hdfs_user}${principal_suffix}@${realm}
HDFS user keytab	${keytab_dir}/hdfs.headless.keytab
Spark2 user keytab	${keytab_dir}/spark.headless.keytab
Spark user keytab	${keytab_dir}/spark.headless.keytab
Spark2 user principal	${spark2-env/spark_user}${principal_suffix}@${realm}
Spark user principal	${spark-env/spark_user}${principal_suffix}@${realm}
Storm user keytab	${keytab_dir}/storm.headless.keytab
Storm user principal	${storm-env/storm_user}${principal_suffix}@${realm}
zeppelin.server.kerberos.keytab	${keytab_dir}/zeppelin.server.kerberos.keytab
zeppelin.server.kerberos.principal	${zeppelin-env/zeppelin_user}${principal_suffix}@${realm}

Figure 6. No modification to Ambari Principals are required

Note: The SAMAccount Name attribute in AD is limited to 20 characters, therefore we need to account for the length of the cluster name in the suffix added to the account names to create the UPNs and SAMAccount Name.



SAM-Account-Name attribute

The logon name used to support clients and servers running earlier versions of the operating system, such as Windows NT 4.0, Windows 95, Windows 98, and LAN Manager.

This attribute must be 20 characters or less to support earlier clients, and cannot contain any of these characters:

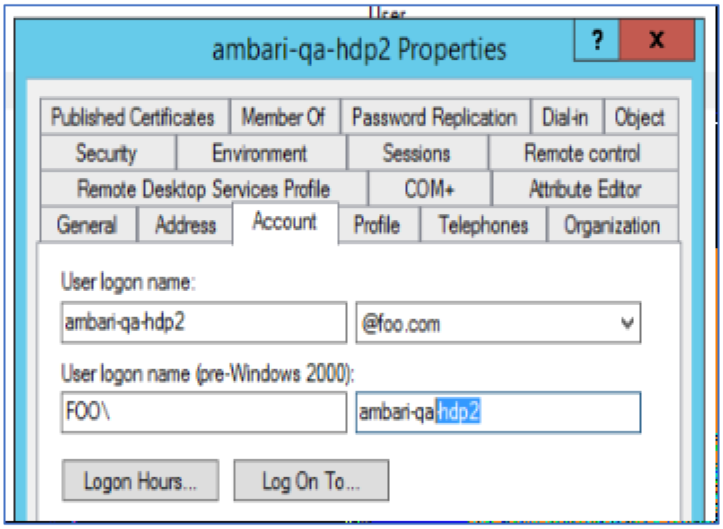
- "/ \ [ ] : ; | = , + * ? < >

CN	SAM-Account-Name
Ldap-Display-Name	sAMAccountName
Size	20 characters or less.
Update Privilege	Domain administrator
Update Frequency	This value should be assigned when the account record is created, and should not change.
Attribute-Id	1.2.840.113556.1.4.221
System-Id-Guid	3e0abfd0-126a-11d0-a060-00aa006c33ed
Syntax	String(Unicode)

Figure 7. SAMAccount Name AD attributes

For additional information on the SAMAccountName field, see the following article: https://msdn.microsoft.com/en-us/library/ms679635(v=vs.85).aspx.

Since the Ambari wizard will create several UPN's that will require matching SAMAccount Names we must make sure that the account plus the suffix is 20 characters or less. Based on this requirement you need to evaluate the cluster name suffix that will be generated and decide if it needs modification from the full clustername to an abbreviated version of the clustername.



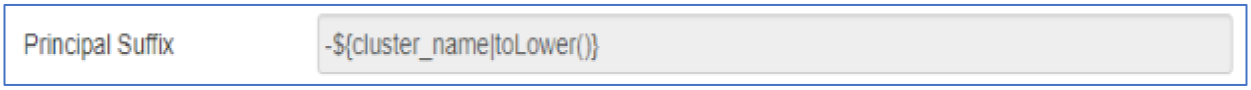
User logon name: ambari-qa-hdp2 @foo.com

User logon name (pre-Windows 2000): FOO\ ambari-qahdp2

Figure 8. The SAMAccount Name attribute in AD

Ambari creates the following UPN's; hdfs, ambari-qa, hbase, spark, storm, zeppelin and ambari-server, therefore if your Ambari cluster name is longer than 7 characters in length you will need to modify default variable to be used as you will not be able to create a corresponding SAMAccount Name equivalent to the UPN username.

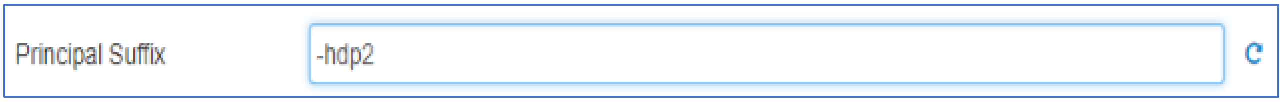
If the ambari clustername is less than 7 characters in length, we can use the default value with the full name.



Principal Suffix: -${cluster_name|toLower()}

Figure 9. Default clustername suffix added, the entire Ambari clustername

If our Ambari clustername is longer than 7 characters and will create UPN usernames of username-clustername in excess of 20 characters, which is valid for the username but we must modify the suffix to meet the 20 character requirement for SAMAccount Name.



Principal Suffix: -hdp2

Figure 10. A modified clustername suffix to meet the less than 20-character limit for SAMAccount Name

Ambari will still create internal mapping rules to map the service user account to the UPN, it will just use the modified suffix, we can now modify the SAMAccount Name filed in AD to match the UPN username and create mapping rules on Isilon.

If this is not done on the initial generation of principals in AD it can always be done later by removing and regenerating the principals with a modified suffix.

4. Make the required modifications to the 'Advanced Tab' to the service account SPNs as documented in the installation guide; hdfs, yarn, mapred etc.

   Note: These modifications are still required for Isilon integration to support valid identities for the Hadoop cluster service SPNs.

5. Complete the Wizard, continuing to follow the installation guide through the deployment of the Active Directory Principals. Before restarting services, follow the additional steps in the Kerberos installation guide.

6. Review the Ambari-generated principals in Active Directory; remove the Isilon cluster SPNs as created by Ambari and add them to the Isilon Computer Object as needed per the installation guide.

   6.1 Remove the Ambari-generated hdfs/isilonsmartconnectzone SPN in the Hadoop OU

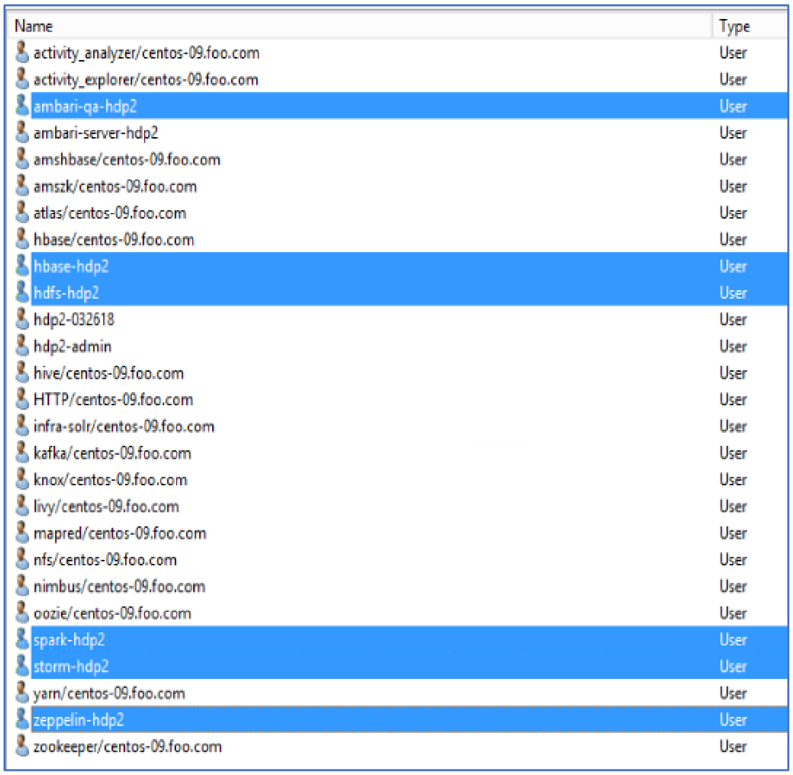
   6.2 Remove the Ambari-generated HTTP/isilonsmartconnectzone SPN SPN in the Hadoop OU

   6.3 Add the hdfs/isilonsmartconnectzone SPN to Isilon

   6.4 Add the HTTP/isilonsmartconnectzone SPN to Isilon

   6.5 Validate the hdfs/clustername SPN exists on Isilon

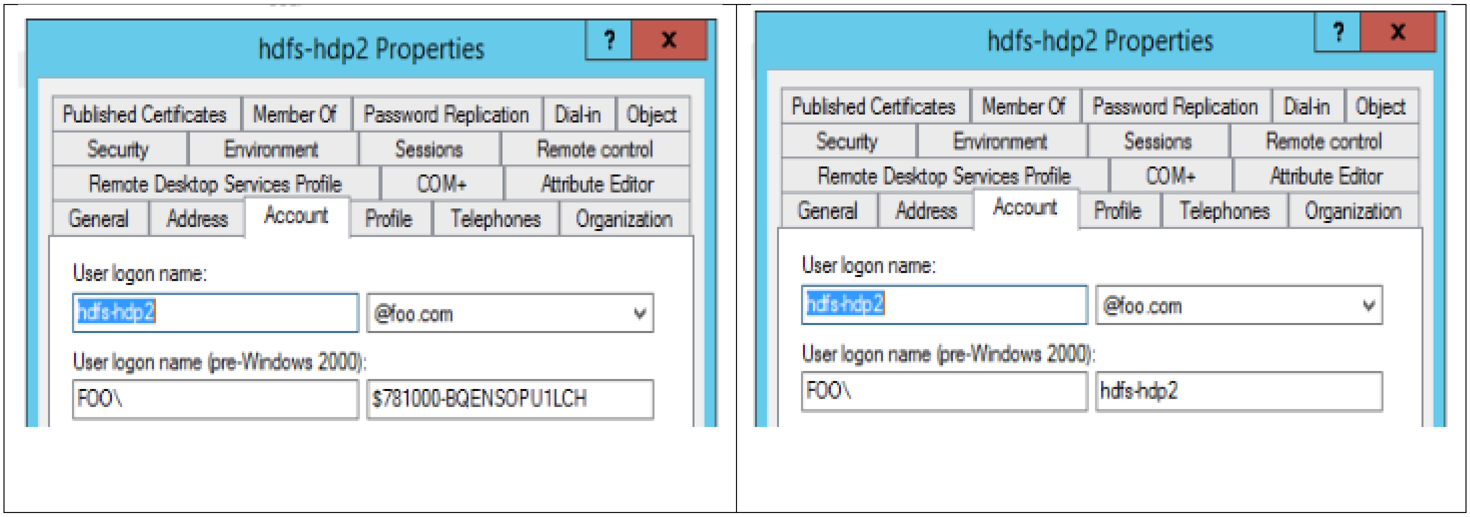
7. Review the Ambari smoke user UPNs, that include the Ambari clustername suffix.



Name	Type
activity_analyzer/centos-09.foo.com	User
activity_explorer/centos-09.foo.com	User
ambari-qa-hdp2	User
ambari-server-hdp2	User
amshbase/centos-09.foo.com	User
amszk/centos-09.foo.com	User
atlas/centos-09.foo.com	User
hbase/centos-09.foo.com	User
hbase-hdp2	User
hdfs-hdp2	User
hdp2-032618	User
hdp2-admin	User
hive/centos-09.foo.com	User
HTTP/centos-09.foo.com	User
infra-solr/centos-09.foo.com	User
kafka/centos-09.foo.com	User
knox/centos-09.foo.com	User
livy/centos-09.foo.com	User
mapred/centos-09.foo.com	User
nfs/centos-09.foo.com	User
nimbus/centos-09.foo.com	User
oozie/centos-09.foo.com	User
spark-hdp2	User
storm-hdp2	User
yarn/centos-09.foo.com	User
zeppelin-hdp2	User
zookeeper/centos-09.foo.com	User

Figure 11. Ambari generated smoke user UPN in Active Directory

8. Update the Ambari smoke user UPNs; modify the User logon name (pre-Windows 2000/SAMAccountName) to match the username portion of the UPN. Since we accounted for the clustername suffix in the wizard, and ensured the length is less than 20 characters, we can modify the SAMAccount name to be equal to the username portion from the UPN knowing it meets the AD requirement.



User logon name:
hdfs-hdp2 @foo.com

User logon name (pre-Windows 2000):
FOO\ $781000-BQENSOPU1LCH

User logon name:
hdfs-hdp2 @foo.com

User logon name (pre-Windows 2000):
FOO\ hdfs-hdp2

Figure 12. Modification required to the Ambari user account

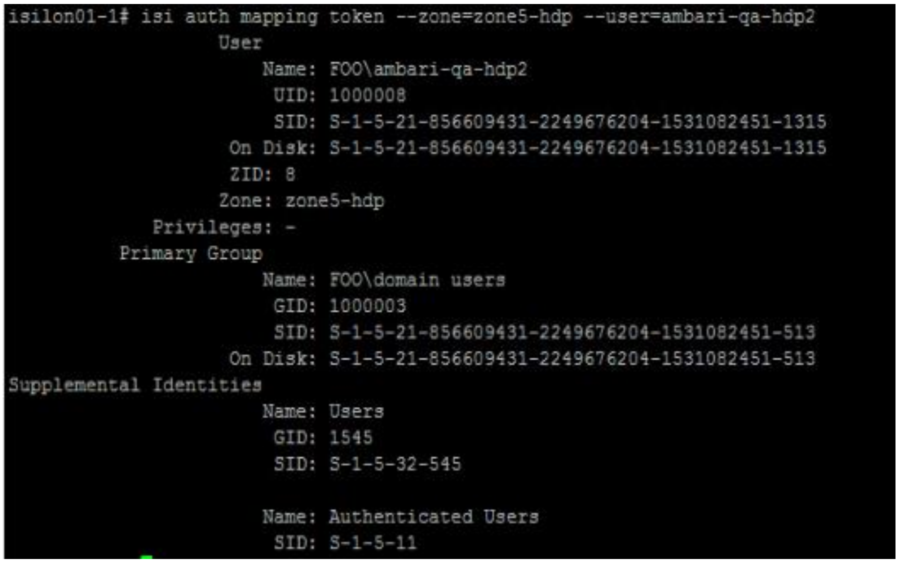
With a clustername suffix of ‘-hdp2’ the required AD modifications and Isilon mapping rules are found in the table below.

AMBARI ACCOUNT	AD UPN ACCOUNT USERNAME	SAMACCOUNTNAME (MODIFIED)	ISILON MAPPING RULE NEEDED	ISILON ACCOUNT
hdfs	hdfs-hdp2	hdfs-hdp2	<domainname>\hdfs-hdp2 => root []	hdfs
ambari-qa	ambari-qa-hdp2	ambari-qa-hdp2	<domainname>\\ambari-qa-hdp2 => ambari-qa []	ambari-qa
<service>	<service>-hdp2	<service>-hdp2	<domainname>\\<service>-hdp2=> ambari-qa []	<service>

Table 2. Required modification to AD User principals:

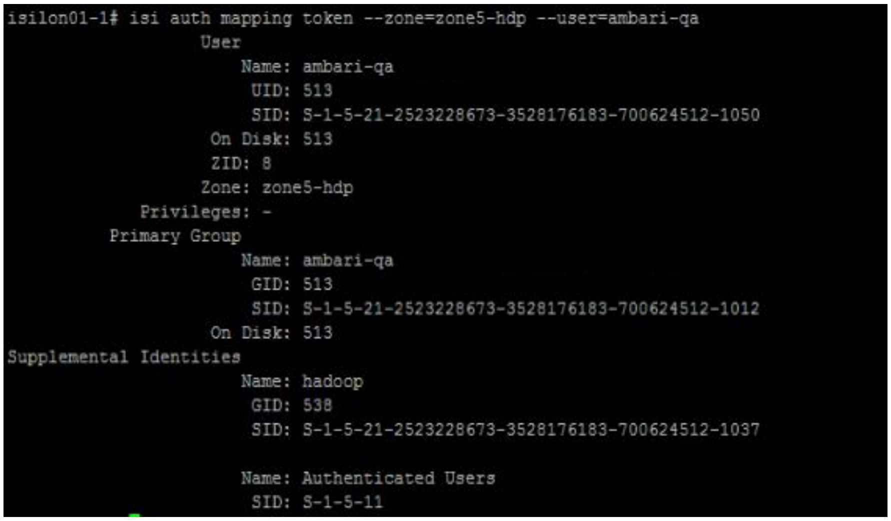
9. Review the current status of the Active Directory UPN ambari-qa-clustername account.

   Without the mapping rule, Isilon has visibility to the Ambari smoke user UPN account and the Local Isilon ambari-qa account (this was the account the cluster was installed with), but they are not mapped to each other.



```
isilon01-1# isi auth mapping token --zone=zone5-hdp --user=ambari-qa-hdp2
                   User
                       Name: FOO\ambari-qa-hdp2
                        UID: 1000008
                        SID: S-1-5-21-856609431-2249676204-1531082451-1315
                    On Disk: S-1-5-21-856609431-2249676204-1531082451-1315
                    ZID: 8
                   Zone: zone5-hdp
             Privileges: -
          Primary Group
                       Name: FOO\domain users
                        GID: 1000003
                        SID: S-1-5-21-856609431-2249676204-1531082451-513
                    On Disk: S-1-5-21-856609431-2249676204-1531082451-513
Supplemental Identities
                       Name: Users
                        GID: 1545
                        SID: S-1-5-32-545

                       Name: Authenticated Users
                        SID: S-1-5-11
```

Figure 13. Isilon sees the Active Directory ambari-qa-clustername user



```
isilon01-1# isi auth mapping token --zone=zone5-hdp --user=ambari-qa
                   User
                       Name: ambari-qa
                        UID: 513
                        SID: S-1-5-21-2523228673-3528176183-700624512-1050
                    On Disk: 513
                    ZID: 8
                   Zone: zone5-hdp
             Privileges: -
          Primary Group
                       Name: ambari-qa
                        GID: 513
                        SID: S-1-5-21-2523228673-3528176183-700624512-1012
                    On Disk: 513
Supplemental Identities
                       Name: hadoop
                        GID: 538
                        SID: S-1-5-21-2523228673-3528176183-700624512-1037

                       Name: Authenticated Users
                        SID: S-1-5-11
```

Figure 14. The local ambari-qa user account

In order to complete the identity management, add the mapping rules on Isilon OneFS to replace ambari-qa-cluster with the local ambari-qa user.

10. Add the required mapping rules to the Isilon Access Zone.

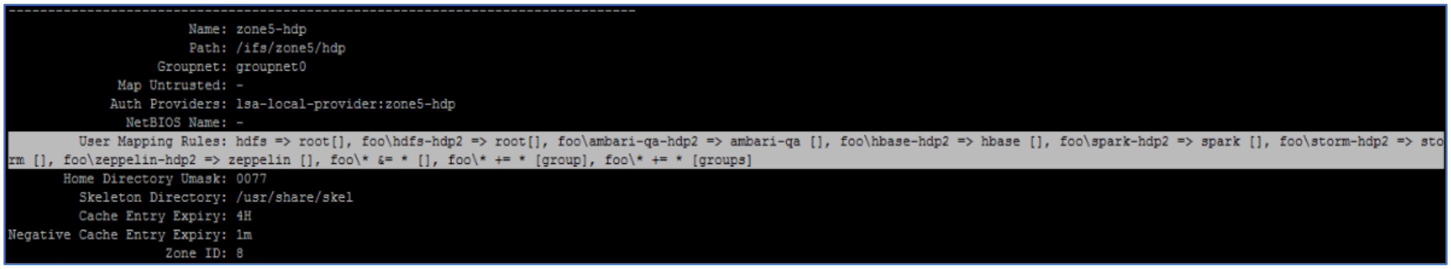
```
isi zone zones modify --zone=<zonename> --add-user-mapping-rules="hdfs => root[]"

isi zone zones modify --zone=<zonename>--add-user-mapping-rules="domain\hdfs-<clsname> => root[]"

isi zone zones modify --zone=<zonename>--add-user-mapping-rules="domain\ambari-qa-<clsname> => ambari-qa []"

isi zone zones list -v
```

Additional rules if needed:

```
isi zone zones modify --zone=<zonename>--add-user-mapping-rules="domain\hbase-<clsname> => hbase []"

isi zone zones modify --zone=<zonename>--add-user-mapping-rules="domain spark-<clsname> => spark []"

isi zone zones modify --zone=<zonename>--add-user-mapping-rules="domain\storm-<clsname> => storm []"

isi zone zones modify --zone=<zonename>--add-user-mapping-rules="domain\zeppelin-<clsname> => zeppelin []"
```

Optional rules to complete Isilon – Active Directory user mapping:

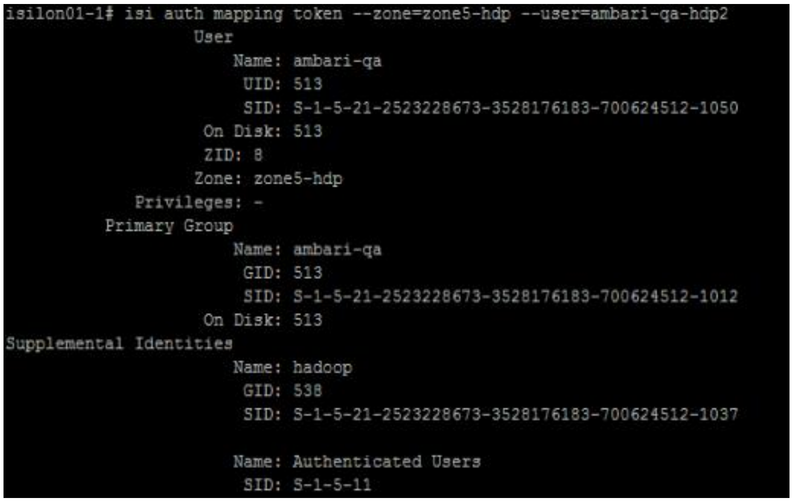
```
isi zone zones modify --zone=zone5-hdp --add-user-mapping-rules="domain\* &= * []"

isi zone zones modify --zone=zone5-hdp --add-user-mapping-rules="domain\* += * [group]"

isi zone zones modify --zone=zone5-hdp --add-user-mapping-rules="domain\* += * [groups]"
```



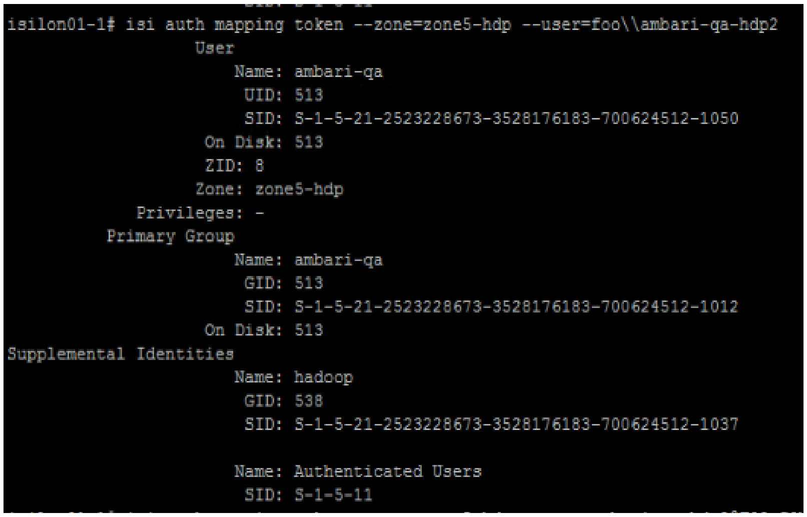
```
--------------------------------------------------------------------------------
                       Name: zone5-hdp
                       Path: /ifs/zone5/hdp
                   Groupnet: groupnet0
              Map Untrusted: -
             Auth Providers: lsa-local-provider:zone5-hdp
               NetBIOS Name: -
         User Mapping Rules: hdfs => root[], foo\hdfs-hdp2 => root[], foo\ambari-qa-hdp2 => ambari-qa [], foo\hbase-hdp2 => hbase [], foo\spark-hdp2 => spark [], foo\storm-hdp2 => sto
rm [], foo\zeppelin-hdp2 => zeppelin [], foo\* &= * [], foo\* += * [group], foo\* += * [groups]
       Home Directory Umask: 0077
         Skeleton Directory: /usr/share/skel
         Cache Entry Expiry: 4H
Negative Cache Entry Expiry: 1m
                    Zone ID: 8
```

Figure 15. Isilon mapping rules

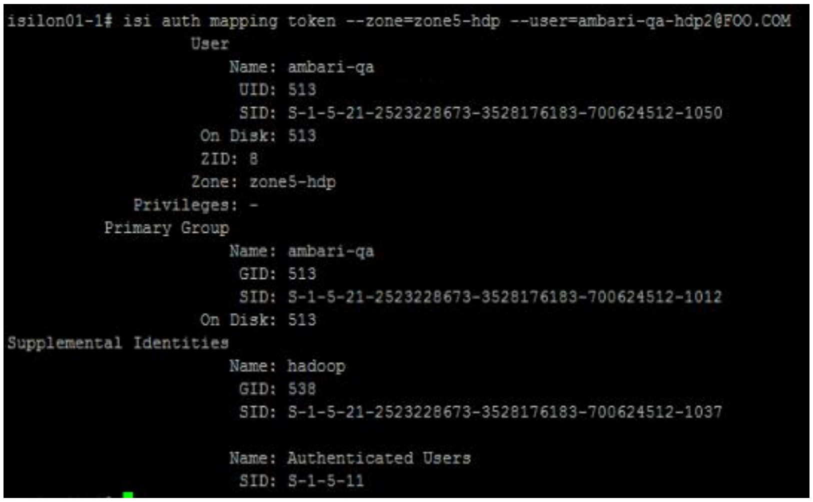
11. Validate the Isilon mapping rule to the Active Directory user account as shown in the following screens:



```
isilon01-1# isi auth mapping token --zone=zone5-hdp --user=ambari-qa-hdp2
                   User
                       Name: ambari-qa
                        UID: 513
                        SID: S-1-5-21-2523228673-3528176183-700624512-1050
                    On Disk: 513
                    ZID: 8
                   Zone: zone5-hdp
             Privileges: -
          Primary Group
                       Name: ambari-qa
                        GID: 513
                        SID: S-1-5-21-2523228673-3528176183-700624512-1012
                    On Disk: 513
Supplemental Identities
                       Name: hadoop
                        GID: 538
                        SID: S-1-5-21-2523228673-3528176183-700624512-1037

                       Name: Authenticated Users
                        SID: S-1-5-11
```



```
isilon01-1# isi auth mapping token --zone=zone5-hdp --user=foo\\ambari-qa-hdp2
                   User
                       Name: ambari-qa
                        UID: 513
                        SID: S-1-5-21-2523228673-3528176183-700624512-1050
                    On Disk: 513
                    ZID: 8
                   Zone: zone5-hdp
             Privileges: -
          Primary Group
                       Name: ambari-qa
                        GID: 513
                        SID: S-1-5-21-2523228673-3528176183-700624512-1012
                    On Disk: 513
Supplemental Identities
                       Name: hadoop
                        GID: 538
                        SID: S-1-5-21-2523228673-3528176183-700624512-1037

                       Name: Authenticated Users
                        SID: S-1-5-11
```

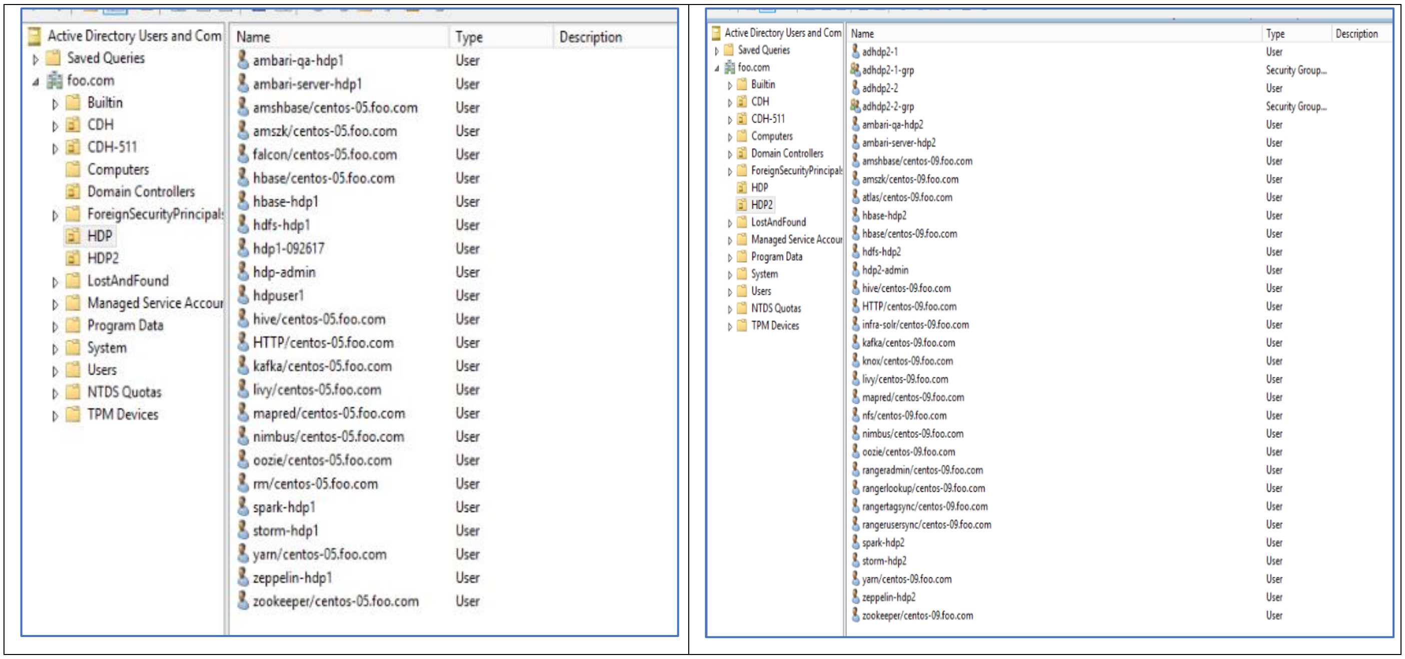


```
isilon01-1# isi auth mapping token --zone=zone5-hdp --user=ambari-qa-hdp2@FOO.COM
                   User
                       Name: ambari-qa
                        UID: 513
                        SID: S-1-5-21-2523228673-3528176183-700624512-1050
                    On Disk: 513
                    ZID: 8
                   Zone: zone5-hdp
             Privileges: -
          Primary Group
                       Name: ambari-qa
                        GID: 513
                        SID: S-1-5-21-2523228673-3528176183-700624512-1012
                    On Disk: 513
Supplemental Identities
                       Name: hadoop
                        GID: 538
                        SID: S-1-5-21-2523228673-3528176183-700624512-1037

                       Name: Authenticated Users
                        SID: S-1-5-11
```

Figure 16. Ambari smoke user UPN maps successfully back to the Isilon service account

With each Ambari UPN now being uniquely defined with the $clustername variable attached to it, we no longer have issues with overlapping SAMAccount Names which could cause installation issues. As such, multiple OUs and Ambari clusters can be Kerberized against the same AD without issues.

The following diagram illustrates two clusters deployed against the same AD with unique UPNs:



- Active Directory Users and Com
  - Saved Queries
  - foo.com
    - Builtin
    - CDH
    - CDH-511
    - Computers
    - Domain Controllers
    - ForeignSecurityPrincipal:
    - HDP
    - HDP2
    - LostAndFound
    - Managed Service Accour
    - Program Data
    - System
    - Users
    - NTDS Quotas
    - TPM Devices

Name	Type	Description
ambari-qa-hdp1	User	
ambari-server-hdp1	User	
amshbase/centos-05.foo.com	User	
amszk/centos-05.foo.com	User	
falcon/centos-05.foo.com	User	
hbase/centos-05.foo.com	User	
hbase-hdp1	User	
hdfs-hdp1	User	
hdp1-092617	User	
hdp-admin	User	
hdpuser1	User	
hive/centos-05.foo.com	User	
HTTP/centos-05.foo.com	User	
kafka/centos-05.foo.com	User	
livy/centos-05.foo.com	User	
mapred/centos-05.foo.com	User	
nimbus/centos-05.foo.com	User	
oozie/centos-05.foo.com	User	
rm/centos-05.foo.com	User	
spark-hdp1	User	
storm-hdp1	User	
yarn/centos-05.foo.com	User	
zeppelin-hdp1	User	
zookeeper/centos-05.foo.com	User	

- Active Directory Users and Com
  - Saved Queries
  - foo.com
    - Builtin
    - CDH
    - CDH-511
    - Computers
    - Domain Controllers
    - ForeignSecurityPrincipal:
    - HDP
    - HDP2
    - LostAndFound
    - Managed Service Accour
    - Program Data
    - System
    - Users
    - NTDS Quotas
    - TPM Devices

Name	Type	Description
adhdp2-1	User	
adhdp2-1-grp	Security Group...	
adhdp2-2	User	
adhdp2-2-grp	Security Group...	
ambari-qa-hdp2	User	
ambari-server-hdp2	User	
amshbase/centos-09.foo.com	User	
amszk/centos-09.foo.com	User	
atlas/centos-09.foo.com	User	
hbase-hdp2	User	
hbase/centos-09.foo.com	User	
hdfs-hdp2	User	
hdp2-admin	User	
hive/centos-09.foo.com	User	
HTTP/centos-09.foo.com	User	
infra-solr/centos-09.foo.com	User	
kafka/centos-09.foo.com	User	
knox/centos-09.foo.com	User	
livy/centos-09.foo.com	User	
mapred/centos-09.foo.com	User	
nfs/centos-09.foo.com	User	
nimbus/centos-09.foo.com	User	
oozie/centos-09.foo.com	User	
rangeradmin/centos-09.foo.com	User	
rangerlookup/centos-09.foo.com	User	
rangertagsync/centos-09.foo.com	User	
rangerusersync/centos-09.foo.com	User	
spark-hdp2	User	
storm-hdp2	User	
yarn/centos-09.foo.com	User	
zeppelin-hdp2	User	
zookeeper/centos-09.foo.com	User	

Figure 17. Two OUs in a single Active Directory supporting two Hadoop clusters

Advantages of this approach:

- Can be used post-deployment with standard user accounts, may need to recreate UPNs and SPN's
- Easy and simple to implement, does not require custom username deployment of Ambari or HDP
- Simplifies Kerberization; less modification to UPNs are needed

Drawbacks of this approach:

- Requires additional Isilon mapping rules

Ambari multitenant review

Having reviewed the deployment and integration of Ambari with Hortonworks HDP with Isilon Kerberos, we can see how—because of the Kerberos deployment methodology used by Ambari—an approach to multitenant deployments should be considered before deploying either Ambari-based Hortonworks clusters or the Isilon HDFS configuration.

The following highlights design decisions that should be considered:

- Dedicated KDCs are supported.
- Shared KDCs are supported.
- In a shared KDC deployment, a strategy should be determined prior to deployment (Option 1 or Option 2).
  - Option 1: UPNs with -clustername leverages additional mapping rules to support multitenancy
  - Option 2: custom usernames require additional configuration modifications and must be done at initial installation

- All standard Isilon Kerberos requirements and best practices should be assessed.

Recommended Approaches and Best Practices

Having reviewed the options and configuration requirements for deploying Kerberized clusters with OneFS, the following recommendations outline the suggested approaches to deploying multitenant Hadoop clusters against a single Isilon cluster.

Ultimately the choice of shared or dedicated KDCs is a specific environment decision and may ultimately be dictated by the existing Kerberos infrastructure in your environment, security policy, or how user identity management is implemented along with Kerberos. Again, it is important to recognize that this paper is focused on the deployment of multitenant Kerberos authentication architectures in a single Active Directory. Additional documentation should be consulted to support alternate deployments.

Dedicated KDCs do provide isolation of Principals and may provide easier administration to the Hadoop and Isilon clusters, while shared centralized KDCs will provide the benefit of central management and existing infrastructure.

Ambari Hortonworks HDP - Isilon OneFS deployment best practices

With Ambari creating SPNs, UPNs, and Isilon SPNs, additional configuration is required as discussed in this whitepaper.

Multiple Ambari Hortonworks HDP clusters can be Kerberized easily against shared or dedicated KDCs when the required configuration changes or modifications are accounted for prior to installation and during configuration.

- Use a per Hadoop cluster Access Zone
- Deploy local Hadoop service accounts with UID – GID parity on Isilon and Cloudera hosts with the standard name: hdfs, yarn, hbase
- Kerberize against KDC using standard Kerberos wizard options
- Leave the `cluster_name|toLower()` variable on Ambari UPNs; create all UPNs with -clustername attached (Option 2)
- If AD, modify Ambari UPN SAMAccount names to reflect the account name, example: hdfs-clustername, ambari-qa-clustername
- Add mapping rules in AD: hdfs-clustername => root, ambari-qa-clustername => ambari-qa
- Remove any Ambari-created Isilon SPNs in the Active Directory OU or the KDC
- Validate and fix any required Isilon SPNs in the KDC or on the Isilon computer object in Active Directory

In all Kerberos implementations, follow the recommended vendor documentation and all Isilon Hadoop Best Practices.

General Isilon Hadoop Kerberos best practices

- Enable forward and reverse DNS lookups on all the hosts.
  - All the compute hosts must have forward DNS lookup resolved correctly for all the hosts.
  - Isilon SmartConnect zone name lookups must resolve correctly.
  - Reverse PTR records for all IP addresses in the SmartConnect pool must exist.
  - Isilon OneFS must be able to resolve all the hosts, KDCs, and Active Directory servers as needed.
- Kerberos implementations are highly dependent accurate time; all Hadoop nodes and OneFS must be configured to use a reliable accurate time source such as NTP. This should be a strata 3 or higher time source. Oftentimes an AD or DNS controller will also server out NTP requests on an enterprise network because of the high availability nature of these infrastructure services.

Conclusion

The approach of using the Ambari appended the '-$clustername' variable to the Ambari user principals in the Kerberos wizard with matching Isilon mapping rules provides the ability to create unique users in Active Directory and will not create conflicts for multitenant implementations where Isilon requires a valid SAMAccount Name. This approach is the recommended method of implementing Kerberos authentication with Ambari and Isilon OneFS at this time.

Resources

- Using Hadoop with Isilon - Isilon Info Hub
- Isilon and Hadoop Cluster Install Guides
- EMC Isilon Best Practices Guide for Hadoop Data Storage
- Isilon OneFS Considerations for Active Directory-based Kerberos with Hadoop