Kerberos Authentication

When running Aspire inside of Cloudera as part of the Hadoop Ecosystem, it is almost always required to interact with Kerberized Hadoop Components such as HBase and HDFS. This page provides a guide on how to prepare the credentials and the configuration needed for Aspire to be able to talk to Kerberized Hadoop Components.

Aspire credentials

It is recommended that aspire has its own credentials on Kerberos or Active Directory.

For creating an account on a MIT Kerberos Server, in the kadmin.local or kadmin shell you must run the following command:

```
$ kadmin
kadmin: addprinc aspire@REALM.COM
```

replace REALM.COM with your own realm.

Then create a keytab for the aspire user:

```
kadmin: xst -k aspire.keytab aspire
```

Test your newly created account

First destroy any kerberos ticket on the cache:

```
$ kdestroy
$ klist
klist: No credentials cache found (filename: /tmp/krb5cc_1000)
```

And then authenticate using the aspire account and keytab

```
$ kinit -kt aspire.keytab aspire
$ klist
Ticket cache: FILE:/tmp/krb5cc_1000
Default principal: aspire@REALM.COM

Valid starting       Expires              Service principal
04/27/2018 22:46:14  04/28/2018 22:46:14  krbtgt/REALM.COM@REALM.COM
        renew until 05/04/2018 22:46:14
```

Set the necessary permissions on HDFS and/or HBase

> ⚠ You might need to execute the following commands with an existing account with sufficient permissions

HDFS user directory

First if you want to be able to write to HDFS from Aspire, you may want to create a user directory for aspire in HDFS. First make sure you have correctly authenticated with Kerberos using the kinit command from above. Then create the /user/aspire directory by executing the following commands:

```
$ hadoop fs -mkdir /user/aspire
$ hadoop fs -chown aspire /user/aspire
```

If you want Aspire to be able to read from an specific HDFS directory, then make sure the aspire user can read it by looking at the permissions from the directoy:

```
$ hadoop fs -ls /doc
Found 1 items
drwxrwxrwx - hdfs supergroup 0 2017-12-06 19:53 /doc/sourceId
```

HBase

To test connection with HBase after authenticating with the kinit command, let's open the hbase shell:

```
$ hbase shell
2018-04-27 23:06:44,384 INFO [main] Configuration.deprecation: hadoop.native.lib is deprecated. Instead, use io.
native.lib.available
HBase Shell; enter 'help<RETURN>' for list of supported commands.
Type "exit<RETURN>" to leave the HBase Shell
Version 1.2.0-cdh5.12.1, rUnknown, Thu Aug 24 09:37:07 PDT 2017

hbase(main):001:0>
```

And execute the list command to test the aspire user permissions:

```
hbase(main):001:0> list
```

If you find any troubles with HBase permissions for your aspire user check Cloudera HBase Authorization for step by step instructions on how to set the appropiate permissions.

If you are using HBase for crawls metadata, take into account that you might need either **Admin** permissions for Aspire to be able to create new namespaces for each new content source, or the namespaces already be created and assign **Create, Read** and **Write** permissions to those namespaces.