Cloudera DataFlow for Data Hub 7.2.11

Configuring Flow Management clusters to hot load custom NARs

Date published: 2019-12-16
Date modified: 2021-09-09

Legal Notice

Contents

Configuring Flow Management clusters to hot load custom NARs.....4

Configuring Flow Management clusters to hot load custom NARs

Hot loading custom NARs means that you can store the NARs for your NiFi custom processors in an external folder, and load them from that directory for use on the NiFi canvas. This makes it easy to edit custom processors and add new ones without having to restart your Flow Management cluster.

Before you begin

- Collect your Kerberos information.
  1. In Management Console, click Environments.
  2. From the Actions drop-down, click Manage Access.
  3. Create a Machine User with the EnvironmentUser Resource Role and set its Workload password.

     Note:

     It is best practice to have a dedicated Machine User in this role. This Machine User is in charge of accessing the object store resources for NiFi, having a dedicated user ensures that there is least privilege access and separation of duties.
  4. Note down the user you create there and password, for your safety valves.
- You have created an ID Broker mapping.
  1. Access IDBroker mappings.
     a. To access IDBroker mappings in your environment, click Actions > Manage Access.
     b. Choose the IDBroker Mappings tab where you can provide mappings for users or groups and click Edit.
  2. Add your CDP Machine User and the corresponding AWS role that provides write access to your folder in your S3 bucket to the Current Mappings section by clicking the blue + sign.

     Note: You can get the AWS IAM role ARN from the Roles Summary page in AWS and can copy it into the IDBroker role field. The selected AWS IAM role must have a trust policy allowing IDBroker to assume this role.
  3. Click Save and Sync.

About this task

- You can configure NiFi to hot load custom NARs in all supported cloud providers.
- You can only access resources in the object store of the cloud provider in which you environment is running. But but you can have a development and a production environment, for example, and configure both environments to access the same location to hot load custom NARs
- It is possible to define multiple locations if desired by repeating the configuration properties with a different “[**custom-name**]” for each location. But all locations should be in the object store of the underlying cloud provider of where the NiFi cluster is located.
- NiFi polls for new NARs in the configured location(s) every 5 minutes.

Procedure

1. In Management Console, select Data Hub Clusters, then go to the Flow Management cluster, then NiFi Service.
2. Click Configuration.
3. Search for Safety Valve and then Find: NiFi Node Advanced Configuration Snippet (Safety Valve) for staging/nifi.properties.xml
4. Create the following new properties:

```
Name: nifi.nar.library.provider.[**custom-name**].implementation
```

```
Example: nifi.nar.library.provider.narstorage1.implementation
Value: org.apache.nifi.nar.hadoop.HDFSNarProvider

Path to the configuration files
Name: nifi.nar.library.provider.[**custom-name**].resources
Example: nifi.nar.library.provider.narstorage1.resources
Value, by default: /etc/hadoop/conf.cloudera.core_settings/core-site.xml

Storage location for the NARs (the bucket name in S3 for example)
Name: nifi.nar.library.provider.[**custom-name**].storage.location
Example: nifi.nar.library.provider.narstorage1.storage.location
Value: s3a://my-custom-nars-bucket

Directory library with path to the NARs relative to the storage location
Name: nifi.nar.library.provider.[**custom-name**].source.directory
Example: nifi.nar.library.provider.narstorage1.source.directory
Value: /myNARs

Workload username and password of the machine user to use to access the r
esources

Name: nifi.nar.library.provider.[**custom-name**].kerberos.principal
Example: nifi.nar.library.provider.narstorage1.kerberos.principal
Value: srv_my_s3_nar_user

Name: nifi.nar.library.provider.[**custom-name**].kerberos.password
Example: nifi.nar.library.provider.narstorage1.kerberos.password
Value: myPassword123
```

Note:

Cloudera recommends that the workload user name and password that you configure above is strictly dedicated to accessing custom NARs, and that the only permissions granted to this user are to access the custom NAR object store location.

Alternately, you can also use the keytab file. However, this file must be distributed on all NiFi nodes. This becomes problematic when scaling up a NiFi cluster in DataHub. In addition, the file permissions must be properly changed to ensure the file is readable by the `nifi` user. However, if desired, the below configuration can be used:

```
Name: nifi.nar.library.provider.[**custom-name**].kerberos.keytab
Example: nifi.nar.library.provider.narstorage1.kerberos.keytab
Value: /path/to/my/file.keytab
```

5. Restart the NiFi service in the Cloudera Manager.

Results

Refresh the NiFi UI and then drag and drop a new processor to NiFi canvas and search for your custom processor.