Cloudera DataFlow

Test Sessions

Date published: 2021-04-06
Date modified: 2024-06-03

CLOUDERA

Legal Notice

Contents

Test sessions

Starting a Test Session provisions NiFi resources, acting like a development sandbox for a particular draft. It allows you to work with live data to validate your data flow logic while updating your draft. You can suspend a test session any time and change the configuration of the NiFi cluster, then resume testing with the updated configuration.

Starting a test session

To validate your draft flow, start a test session, provisioning a NiFi cluster.

Before you begin

- Make sure that you have DFDeveloper permission to perform this task. For information on account and resource roles, see Cloudera DataFlow Authorization.
- You have created a draft in Flow Designer.

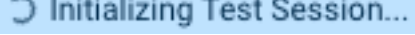## Procedure

1. Select Flow Options Test Session .

   - Add Custom NAR Configuration if your flow requires custom processors that you have built.
   - Select Inbound Connections Allow NiFi to receive data if your your draft uses listen processors with inbound connection endpoint support.

     The deployment wizard auto synchronizes port information with the configuration settings you made for your listen processor(s).
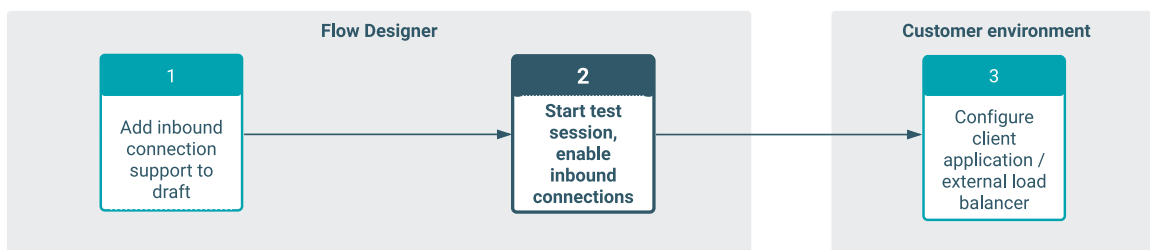
   > **Note:** You can also change NiFi version but, unless you have a specific reason to select an earlier one, Cloudera recommends using the latest version.

2. Click Start Test Session.

   Test Session status changes to Initializing Test Session...

3.

   Wait for the status to change to Active Test Session.

4. Click Flow Options Services to enable Services.

5.

   Select a service you want to enable, then click ⏻ Enable.

   Repeat this step for all Services.

   > **Tip:** You can also select Enable Service and Referencing Components. It will not only enable the controller service, but also any component that references it. This way, you do not need to enable the component separately to run the test session.

6. Click Back To Flow Designer to return to the flow design Canvas.

7.
   Start all processors that are not running (you see a ✖ Stopped symbol next to the processor name) by selecting

   the processor on the Canvas then clicking ⊙ Start.

Results

The flow starts executing. On the Flow Design Canvas you can observe statistics on your processors change as they consume data and execute their respective tasks. You can also observe one or more blue Notification Pill, providing information about the current task. Potential issues that need to be addressed are called out in an orange

🔔 Notification Pill on the individual processors.

Connecting an external application to a test session

Once a test session with inbound connection is running, you can go on and connect an external application to start sending data.
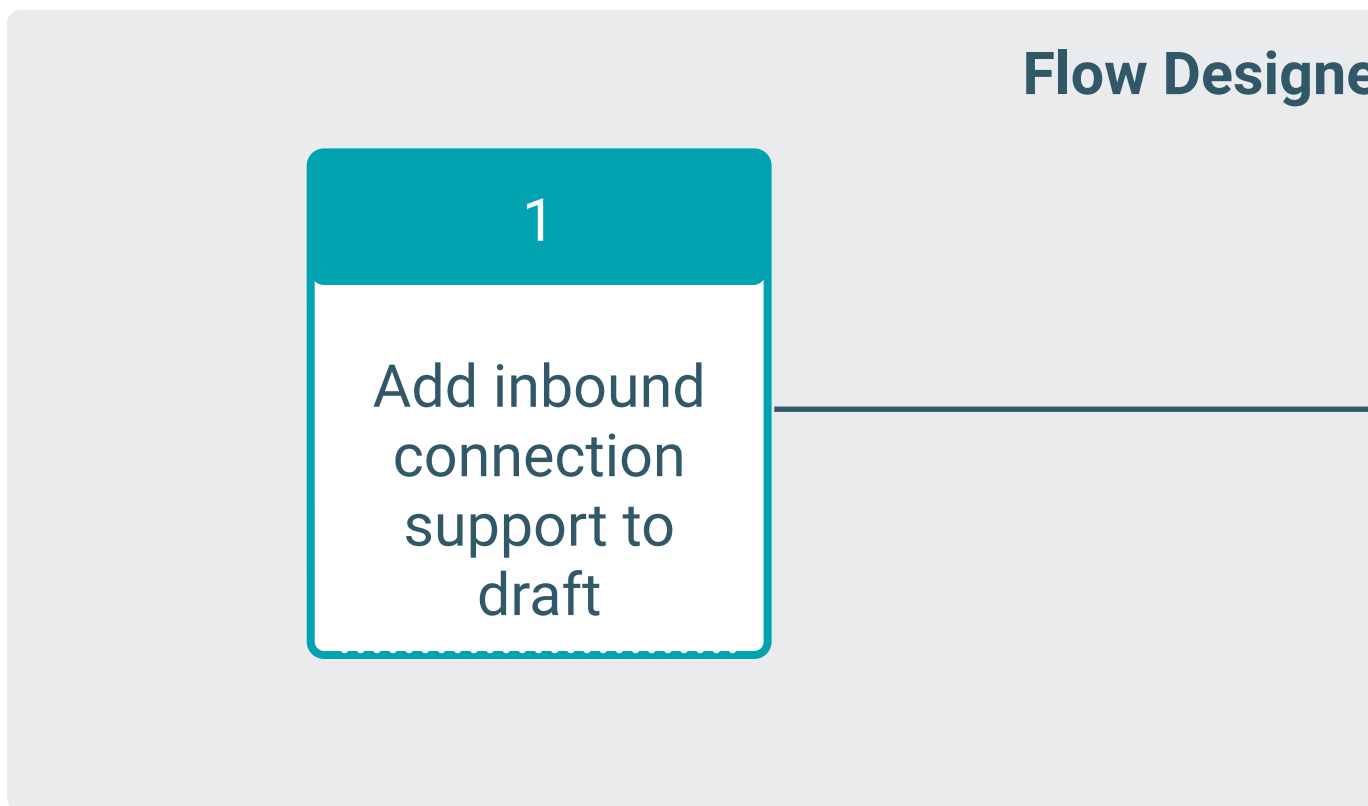
Before you begin

- A test session with inbound connection configured is running.
- A network connection through which the client can reach the NiFi endpoint is available.
- You have been assigned at least the DFFlowDeveloper role for the environment to which you want to configure the inbound connection.

Procedure

1. Select the draft where you want to send data and go to Flow Options Inbound Connection Details .

2. Make a note of the ENDPOINT HOSTNAME and LISTENING PORTS.

3. Click Download Client Certificate.

   The X.509 client certificate downloads to your computer in PEM format.

4. Click Download Private Key to obtain the RSA Private Key.

   The unencrypted RSA Private Key encoded with PKCS8 downloads to your computer in PEM format.

5. Depending on your client, you may have to convert the certificate and the private key to a different format.

   For example, to convert PEM to PKCS12 format, use the following command:

   ```
   openssl pkcs12 -export -in [***DOWNLOADED PEM CERT FILE***] -in
   key [***DOWNLOADED PEM PRIVATE KEY***] -out certificate.p12
   ```

   To further convert the PKCS12 file to JKS format for a Java client, run the following command:

   ```
   keytool -importkeystore -srckeystore [***CERTIFICATE NAME***].p12 -srcsto
   retype pkcs12 -destkeystore [***DESTINATION KEYSTORE***].jks
   ```

6. Add the certificate file and the private key files to the keystore of your application.

7. Configure your application to stream data to the Endpoint Hostname, port, and protocol of the NiFi instance
   behind the Test Session.

Stopping a test session

You can stop a test session to reconfigure your NiFi cluster. You can continue editing your draft flow in the
meantime.

Before you begin

- Make sure that you have DFDeveloper permission to perform this task. For information on account and resource
  roles, see Cloudera DataFlow Authorization.
- You have an active test session.

Procedure

1. You have two options to end an active test session:

   - Click the toggle next to the Active Test Session notification.
   - Click Flow Options then click 🗎 Stop under Test Session.

2. Click End.

   Test session status first changes to ⟳ Ending Test Session... Ending Test Session... then, after the NiFi

   cluster shuts down, it changes to Test Session Not Active ◖ Test Session Not Active.

What to do next

Once the test session stops, you can update the configuration of your NiFi testbed cluster. To do so, select Flow

Options Test Session ✎ Edit Configuration .

Restarting a test session

Learn about restarting an inactive test session.

Before you begin

- Make sure that you have DFDeveloper permission to perform this task. For information on account and resource roles, see Cloudera DataFlow Authorization.
- You have an inactive test session.

Procedure

1. You can optionally update the configuration of your NiFi development sandbox before restarting it. Click Flow Options then click Edit Configuration under Test Session.

2. You have two options to restart an inactive test session:

   - Click the toggle next to the Test Session Not Active ⬤ Test Session Not Active notification.
   - Click Flow Options then click 📋 Restart under Test Session.

Verifying flow file content

Before you begin

- Make sure that you have DFDeveloper permission to perform this task. For information on account and resource roles, see Cloudera DataFlow Authorization.
- You have an active test session.

Procedure

1. On the Flow Design Canvas, click the connection where you want to inspect the flow files.

2.

In the Configuration pane select ⋮ ☰ List Queue .
In the List Queue view, All Queued Flow Files for [***RELATIONSHIP NAME***] Relationship opens.

3. Select a flow file by clicking it to view its Attributes in the right-hand pane.

4.

If there is an attribute that is of special interest to you, hover the cursor over it and select ⚲ Pin this attribute
The selected attribute is pinned to the top of the list of attributes. Pinning persists, if you select a different flow file, or return to the Canvas and decide to view flow files later, you will find the pinned attribute(s) at the top of the list.

»

19e761c1-6a14-4b25-a30b-77c030338541
📋 19e761c1-6a14-4b25-a30b-77c030338541

More Details ∨

⬇   📖

Attributes (29)   📌 Pinned (1) ▾   📋 Copy

invokehttp.status.code   📌
200

date
Mon, 06 Mar 2023 10:43:10 GMT

server
mw2284.codfw.wmnet

invokehttp.tx.id
65ebe530-3c4b-4b28-99c7-aed19f70688b

vary
Accept-Encoding,Treat-as-Untrusted,X-Forwarded-Proto,Cookie,Authorization

x-frame-options
SAMEORIGIN

x-client-ip
34.213.89.169

invokehttp.remote.dn
CN=*.wikipedia.org

invokehttp.request.duration
226

uuid
19e761c1-6a14-4b25-a30b-77c030338541

invokehttp.request.url
https://en.wikipedia.org/w/api.php?action=query&list=recentchanges&format=jso...

5.
To view flow file content, click Open in Data Viewer on top of the Attributes pane.

> **Tip:**
> - Click More details ⌄ to view file metadata.
> - To format the file printout, select an available grammar from the View as drop-down.

The contents of the flow file is displayed in a new tab in your browser.

Queue management

You can remove all flow files from a selected connection or all flow files on process group level.

Before you begin

- Make sure that you have DFDeveloper permission to perform this task. For information on account and resource roles, see Cloudera DataFlow Authorization.
- You have an active test session.

Procedure

- To view the queue of a connection, right click on the connection and select List Queue.
  The top 100 flow files are listed on the **List Queue** page.

Emptying a queue

Learn how to delete enqueued flow files from a connection.

Procedure

1. Right click on the connection where you want to delete the enqueued flow files.

2.
Select ⊖ Empty Queue.

3.
Click 🗑 Empty.

All FlowFiles waiting on the selected connection at the time of the request are removed.

Emptying all queues

Learn how to delete enqueued flow files from all connections on Process Group level.

Procedure

1. Click on an empty area of the Canvas.
   The **Configuration** pane on the right displays configuration options for the process group you are currently in.

2. Click ⊖ Empty All Queues.

3. Click 🗑 Empty.

   All FlowFiles waiting on all connections within the process group at the time of the request are removed.