

Apache NiFi 3

Managing a DataFlow

Date of Publish: 2018-08-13

<http://docs.hortonworks.com>

Contents

Command and Control of the DataFlow.....	3
Starting a Component.....	3
Stopping a Component.....	3
Enabling/Disabling a Component.....	3
Remote Process Group Transmission.....	4
Individual Port Transmission.....	4
Other Management Features.....	7


Command and Control of the DataFlow

When a component is added to the NiFi canvas, it is in the Stopped state. In order to cause the component to be triggered, the component must be started. Once started, the component can be stopped at any time. From a Stopped state, the component can be configured, started, or disabled.

Starting a Component

In order to start a component, the following conditions must be met:

- The component's configuration must be valid.
- All defined Relationships for the component must be connected to another component or auto-terminated.
- The component must be stopped.
- The component must be enabled.
- The component must have no active tasks.


Components can be started by selecting all of the components to start and then clicking the Start button () in the Operate Palette or by right-clicking a single component and choosing Start from the context menu.

If starting a Process Group, all components within that Process Group (including child Process Groups) will be started, with the exception of those components that are invalid or disabled.

Once started, the status indicator of a Processor will change to a Play symbol ().

Stopping a Component

A component can be stopped any time that it is running. A component is stopped by right-clicking on the component

and clicking Stop from the context menu, or by selecting the component and clicking the Stop button () in the Operate Palette.

If a Process Group is stopped, all of the components within the Process Group (including child Process Groups) will be stopped.

Once stopped, the status indicator of a component will change to the Stop symbol ().

Stopping a component does not interrupt its currently running tasks. Rather, it stops scheduling new tasks to be performed. The number of active tasks is shown in the top-right corner of the Processor.



Enabling/Disabling a Component

When a component is enabled, it is able to be started. Users may choose to disable components when they are part of a dataflow that is still being assembled, for example. Typically, if a component is not intended to be run, the component is disabled, rather than being left in the Stopped state. This helps to distinguish between components that are intentionally not running and those that may have been stopped temporarily (for instance, to change the component's configuration) and inadvertently were never restarted.

When it is desirable to re-enable a component, it can be enabled by selecting the component and clicking the Enable



button () in the Operate Palette. This is available only when the selected component or components are disabled. Alternatively, a component can be enabled by checking the checkbox next to the "Enabled" option in the Settings tab of the Processor configuration dialog or the configuration dialog for a Port.

Once enabled, the component's status indicator will change to either Invalid () or Stopped (), depending on whether or not the component is valid.






A component is then disabled by selecting the component and clicking the Disable button () in the Operate Palette, or by clearing the checkbox next to the "Enabled" option in the Settings tab of the Processor configuration dialog or the configuration dialog for a Port.

Only Ports and Processors can be enabled and disabled.

Remote Process Group Transmission

Remote Process Groups provide a mechanism for sending data to or retrieving data from a remote instance of NiFi. When a Remote Process Group (RPG) is added to the canvas, it is added with the Transmission Disabled, as indicated

by the icon () in the top-left corner. When Transmission is Disabled, it can be enabled by right-clicking on the RPG and clicking the "Enable Transmission" menu item. This will cause all ports for which there is a Connection to begin transmitting data. This will cause the status indicator to then change to the Transmission Enabled icon ().

If there are problems communicating with the Remote Process Group, a Warning indicator () may instead be present in the top-left corner. Hovering over this Warning indicator with the mouse will provide more information about the problem.

Individual Port Transmission

There are times when the DFM may want to either enable or disable transmission for only a specific Port within the Remote Process Group. This can be accomplished by right-clicking on the Remote Process Group and choosing the "Remote ports" menu item. This provides a configuration dialog from which each Port can be configured:

Remote Process Group Ports

Name

NiFi Flow

URL

http://nifi-01:8080/nifi

Input ports	
<input checked="" type="checkbox"/>	<p>Tweets No description specified.</p> <p>Concurrent tasks ? 1</p> <p>Compressed No</p>
<input type="checkbox"/>	<p>Provenance Data No description specified.</p> <p>Concurrent tasks ? 1</p> <p>Compressed No</p>

Output ports

The left-hand side lists all of the Input Ports that the remote instance of NiFi allows data to be sent to. The right-hand side lists all of the Output Ports from which this instance is able to pull data. If the remote instance is using secure communications (the URL of the NiFi instance begins with https://, rather than http://), any Ports that the remote instance has not made available to this instance will not be shown.

Note: If a Port that is expected to be shown is not shown in this dialog, ensure that the instance has proper permissions and that the Remote Process Group's flow is current. This can be checked by closing the Port Configuration Dialog and looking at the bottom-right corner of the Remote Process Group. The date at which the flow was last refreshed is shown. If the flow appears to be outdated, it can be updated by right-clicking on


the Remote Process Group and selecting "Refresh flow." (See Anatomy of a Remote Process Group for more information).

Each Port is shown with the Port name, followed by its description, currently configured number of Concurrent tasks, and whether or not data sent to this port will be compressed. To the left of this information is a switch to turn the Port on or off. Those Ports that have no Connections attached to them are grayed out:

Remote Process Group Ports

Name		URL							
NiFi Flow	Currently Transmitting	http://nifi-01:8080/nifi							
<div style="display: flex; justify-content: space-between;"> <div style="width: 70%;"> <div style="background-color: #4a7c8c; color: white; padding: 5px; margin-bottom: 5px;">Input ports</div> <table border="1" style="width: 100%; border-collapse: collapse;"> <tbody> <tr> <td style="width: 15%; text-align: center;"> <input checked="" type="checkbox"/> </td> <td style="width: 45%;"> <p>Tweets No description specified.</p> <p>Concurrent tasks ? 1</p> </td> <td style="width: 40%; text-align: center;"> <p>Compressed No</p> </td> </tr> <tr> <td style="text-align: center;"> <input type="checkbox"/> </td> <td> <p>Provenance Data No description specified.</p> <p>Concurrent tasks ? 1</p> </td> <td style="text-align: center;"> <p>Compressed No</p> </td> </tr> </tbody> </table> </div> <div style="width: 25%; border-left: 1px solid #ccc; padding-left: 10px;"> <div style="background-color: #4a7c8c; color: white; padding: 5px; margin-bottom: 5px;">Output ports</div> <table border="1" style="width: 100%; border-collapse: collapse;"> <tbody> <tr> <td style="height: 20px;"></td> </tr> </tbody> </table> </div> </div>			<input checked="" type="checkbox"/>	<p>Tweets No description specified.</p> <p>Concurrent tasks ? 1</p>	<p>Compressed No</p>	<input type="checkbox"/>	<p>Provenance Data No description specified.</p> <p>Concurrent tasks ? 1</p>	<p>Compressed No</p>	
<input checked="" type="checkbox"/>	<p>Tweets No description specified.</p> <p>Concurrent tasks ? 1</p>	<p>Compressed No</p>							
<input type="checkbox"/>	<p>Provenance Data No description specified.</p> <p>Concurrent tasks ? 1</p>	<p>Compressed No</p>							
	Not Connected								

The on/off switch provides a mechanism to enable and disable transmission for each Port in the Remote Process Group independently. Those Ports that are connected but are not currently transmitting can be configured by clicking

the pencil icon () below the on/off switch. Clicking this icon will allow the DFM to change the number of Concurrent tasks and whether or not compression should be used when transmitting data to or from this Port.

Other Management Features

In addition to the Summary Page, Data Provenance Page, Template Management Page, and Bulletin Board Page, there are other tools in the Global Menu that are useful to the DFM. Select Flow Configuration History to view all the changes that have been made to the dataflow. The history can aid in troubleshooting, such as if a recent change to the dataflow has caused a problem and needs to be fixed. The DFM can see what changes have been made and adjust the flow as needed to fix the problem. While NiFi does not have an "undo" feature, the DFM can make new changes to the dataflow that will fix the problem.

Two other tools in the Global Menu are Controller Settings and Users. The Controller Settings page provides the ability to change the name of the NiFi instance, add comments describing the NiFi instance, and set the maximum number of threads that are available to the application.