

Apache NiFi

Advanced Flow Design

Organizing Flows

A basic flow may be laid out as a simple chain of processors. More complicated flows can benefit from moving some portion of the flow into a Process Group.

What is a Process Group?

- A collection of processors, connections, and input/output ports
- A convenient way to separate flow logic into more understandable parts
- The default canvas is the Root Process Group and all other Process Groups are children of the Root group
- In short: A Process Group is a flow

Process Groups

Process Group Name

Bulletin Notification

Unpack All the Files!

1 1

0 0 4 0 0 0

Active
Tasks

Component
Count

Queued	67 / 4.76 MB	
In	1 / 281.21 MB	(5 min)
Read/Write	281.27 MB / 281.08 MB	(5 min)
Out	110 / 276.32 MB	(5 min)

Unpack zip and tar files based on Mime Type




Comments

5-minute Stats

Benefits

- Reduce clutter on canvas
- Make flows more understandable
- Allow for easier flow administration and navigation
- Allows for metric information at the sub flow level

Creating a Process Group

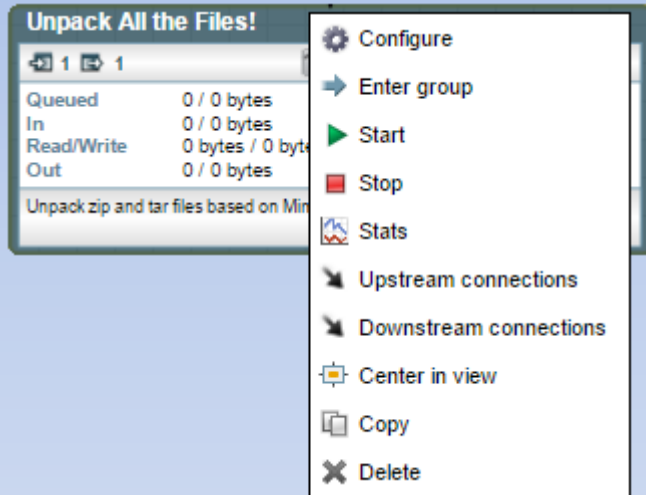
- Drag process group icon () to canvas
- Give the process group a name and click 'Okay'
- Double-click to enter process group and build the internal flow
- Drag  and  to create input/output ports as necessary for receiving and sending data

Using a Process Group

- Return to the parent group by right clicking the canvas and choosing 'Leave group', or by using the breadcrumbs above the canvas
- Create connections to and from the process group using the same procedure as connecting processors

Process Group

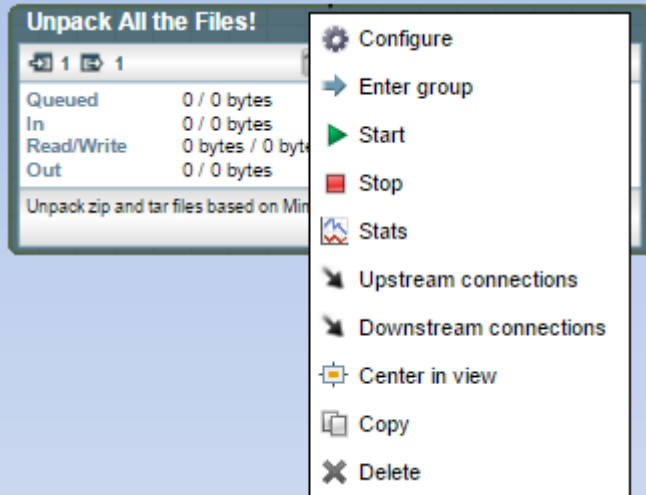
Context Menu Options



- Configure – Update name and comments
- Enter group – View the inner flow
- Start – Start all components
- Stop – Stop all components

Process Group

Context Menu Options

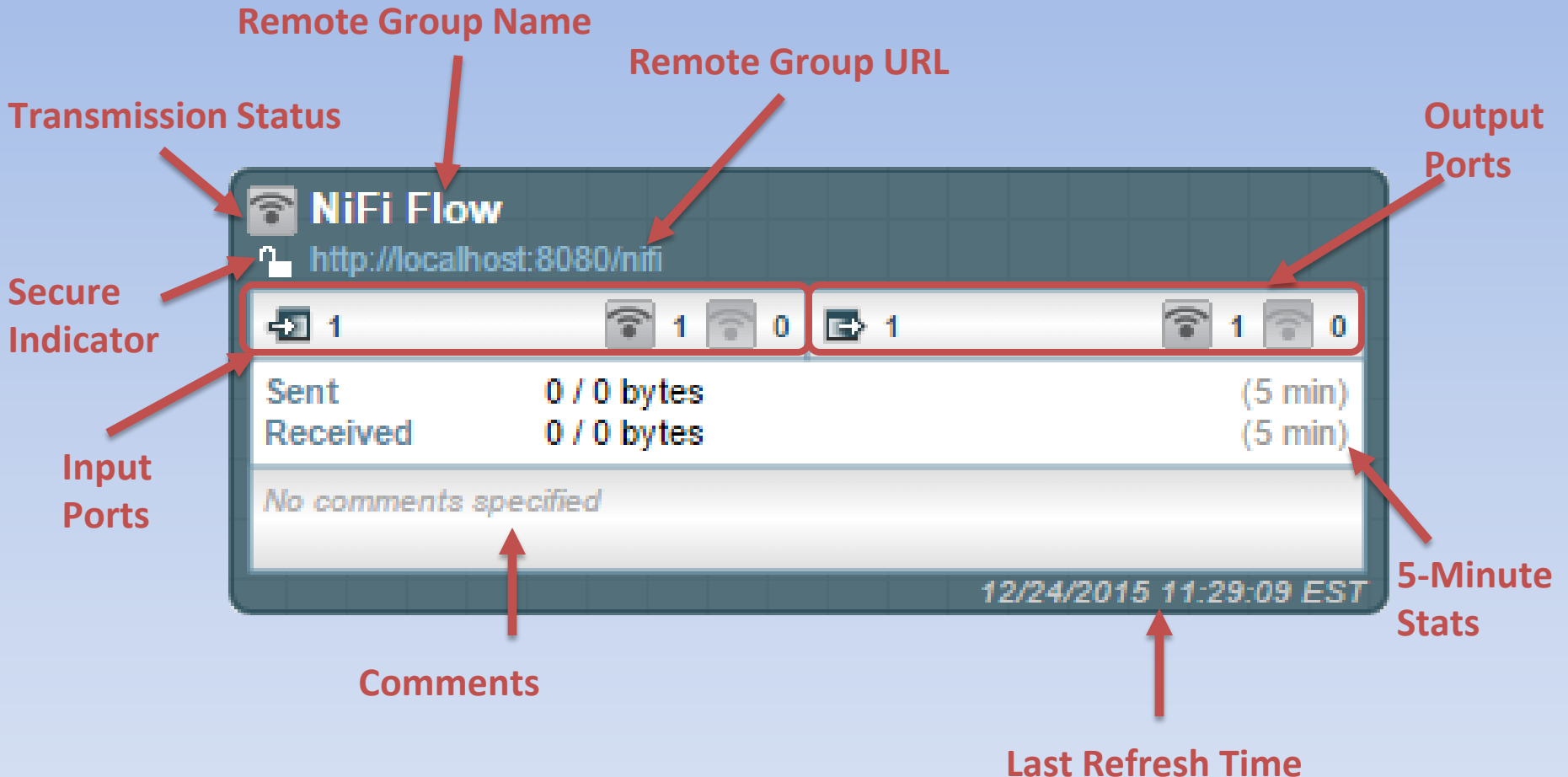


- Stats – View combined stats
- Upstream Connections – View all connections coming into the group
- Downstream Connections – View all connections coming out of the group
- Center in view – Center the current view around the group

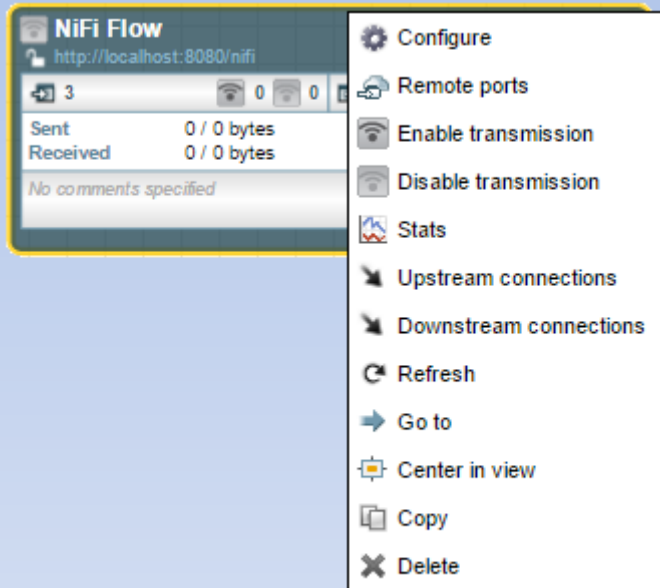
Remote Process Groups

- Appear and behave similar to Process Groups.
- When dropped on the canvas, the user is prompted for the URL of a remote NiFi instance rather than a name
- If the remote NiFi is clustered, the URL should be the URL of the remote instance's NiFi Cluster Manager (NCM)

Remote Process Group

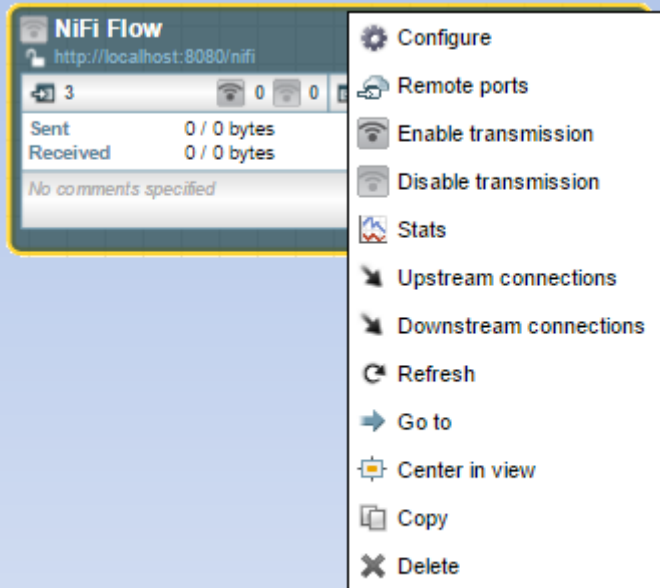


Remote Process Group Context Menu



- Configure – Set timeout and yield options
- Remote ports – View all remote ports for the group
- Enable transmission
- Disable transmission
- Stats – View Remote Flow stats

Remote Process Group Context Menu



- Upstream Connections – View all connections coming into the group
- Downstream connections – View all connections coming out of the group
- Refresh – Refresh the stats and input/output port information
- Go to – Opens a new tab for the URL of the remote system

Port Configuration

Remote Process Group Ports

Name
NiFi Flow

URL
http://localhost:8080/nifi

Input ports

Process 1
No description specified.
OFF
Concurrent tasks ? 1 Compressed No

Process 2
No description specified.
ON
Concurrent tasks ? 1 Compressed No

Process 3
No description specified.
OFF
Concurrent tasks ? 1 Compressed No

Output ports

output
No description specified.
OFF
Concurrent tasks ? 3 Compressed No

Connected, Not Transmitting

Connected, Transmitting

Not Connected

Edit Port Configuration

Close

Funnel

- Funnel allow aggregation of multiple connections/queues
- Allow configuration of back pressure and prioritization on aggregate queue
- Can help reduce clutter in flow and make it easier to follow

Templates

- Combination of Processors, Funnel, Input/Output ports, Process Groups, and Remote Process Groups
- Useful for large and complex data flows by allowing sub-flows to be saved as single units
- Can be exported and imported to allow for sharing between systems with minimal reconfiguration

Creating a Template

Select the components that will be part of the template

Click the 'Create template' button



Enter a name and description for the template and click 'Create'

Finding Templates

Once a Template has been created, it can be found in NiFi's *conf/templates* directory and can be downloaded directly from the 'Templates' menu in the UI

As noted previously, these templates can be used throughout the dataflow and can be utilized on other NiFi systems as well

Importing a Template

Click the 'Templates' button () in the Flow Management portion of the UI

In the new window, click 'Browse' and select the template to import

Upon a successful import, the template will appear in the list of templates

Using a Template

From the 'Draggable Components' section of the UI, drag the 'Template' button () onto the canvas



Select the desired template from the drop-down menu and click 'Add'

Further Resources

- RequiTest Website:
<http://requitest.com/>
- Apache NiFi Website:
<http://nifi.apache.org/>
- Apache NiFi Users Mailing List:
http://mail-archives.apache.org/mod_mbox/nifi-users/
- Apache NiFi Developers Mailing List:
http://mail-archives.apache.org/mod_mbox/nifi-dev/