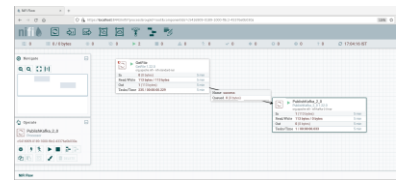


**By Sai Kumar**



## Nifi Kafka Pipeline

### STEP 1

Creating a topic named NiFi message.

```
[root@elephant ~]# kafka-topics --create --bootstrap-server elephant.kafka.com:9092 --topic nifimessage
Created topic nifimessage.
```

### STEP 2

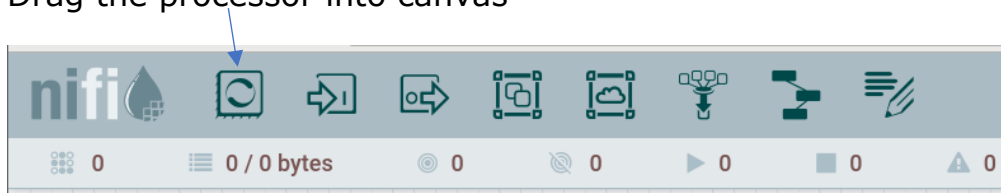
Create a directory named kafkanifi in root directory

```
[root@elephant ~]# mkdir /root/kafkanifi/
[root@elephant ~]# chmod 777 /root/kafkanifi
```

### Step 3

Configuring GetFile Processor

Drag the processor into canvas



**Add Processor**

Source

all groups

Displaying 348 of 348

Filter

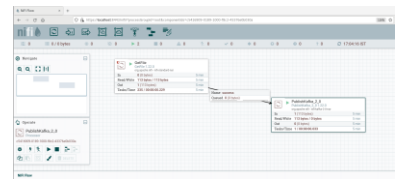
Type	Version	Tags
AttributeRollingWindow	1.22.0	rolling, data science, Attribute E...
AttributesToCSV	1.22.0	flowfile, csv, attributes
AttributesToJSON	1.22.0	flowfile, json, attributes
Base64EncodeContent	1.22.0	encode, base64
CalculateRecordStats	1.22.0	stats, record, metrics
CaptureChangeMySQL	1.22.0	cdc, jdbc, mysql, event, transact...
CompareFuzzyHash	1.22.0	fuzzy-hashing, hashing, cyber-s...
CompressContent	1.22.0	compress, snappy framed, brotl...
ConnectWebsocket	1.22.0	subscribe, consume, listen, We...
ConsumeAMQP	1.22.0	receive, amqp, rabbit, get, cons...
ConsumeAvroKafkaTopic	1.22.0	cloud, avro, kafka, stream, con...

**AttributeRollingWindow 1.22.0** org.apache.nifi - nifi-stateful-analysis-nar

Track a Rolling Window based on evaluating an Expression Language expression on each FlowFile and add that value to the processor's state. Each FlowFile will be emitted with the count of FlowFiles and total aggregate value of values processed in the current time window.

**CANCEL** **ADD**

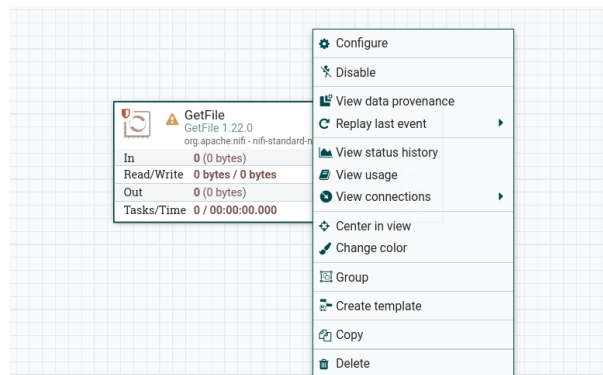
**By Sai Kumar**



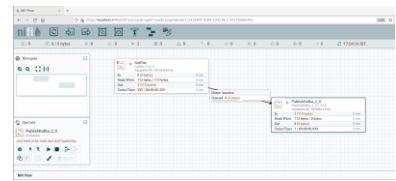
Search for **GetFile** and click **ADD**



Right click on GetFile and Select Configure



**By Sai Kumar**



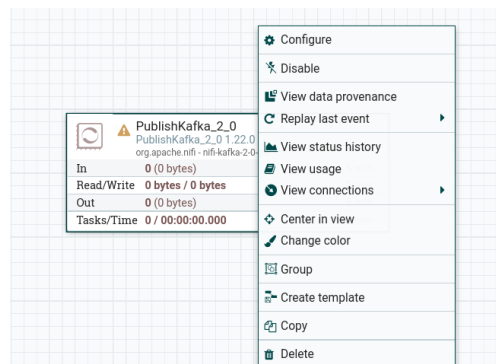
Select **PROPERTIES** and mention the OS directory with path, under Input Directory.

A screenshot of the 'Configure Processor' dialog for 'GetFile 1.22.0'. The 'PROPERTIES' tab is selected. The 'Input Directory' property is highlighted, and a value of '/root/kafkanifi' is entered in the text field. The 'Set empty string' checkbox is unchecked. The 'OK' button is visible at the bottom right of the dialog.

Click **OK** and click on **APPLY**

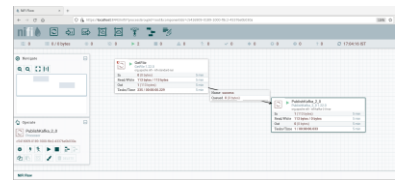
## Step 4

Now again drag Processor and search "PublishKafka\_2\_0", and select ADD



Right click and select **configure**

**By Sai Kumar**



**Configure Processor** | PublishKafka\_2\_0 1.15.3

**Invalid**

**SETTINGS** | SCHEDULING | PROPERTIES | COMMENTS

Name: PublishKafka\_2\_0 ☒ Enabled

Id: 638aa628-0189-1000-ae43-688e8dceaf38

Type: PublishKafka\_2\_0 1.15.3

Automatically Terminate Relationships **failure** ☒ **success** ☒

Any FlowFile that cannot be sent to Kafka will be routed to this Relationship

FlowFiles for which all content was sent to Kafka.

select **PROPERTIES**

Provide

kafka brokers: **elephant.training.com:9092**

Topic name: **nifimessage**

Use transactions: **false**

**Configure Processor** | PublishKafka\_2\_0 1.22.0

**Invalid**

**SETTINGS** | SCHEDULING | **PROPERTIES** | RELATIONSHIPS | COMMENTS

Required field ☒ ☐

Property	Value
Kafka Brokers	<input checked="" type="checkbox"/> elephant.training.com:9092
Security Protocol	<input checked="" type="checkbox"/> PLAINTEXT
SASL Mechanism	<input checked="" type="checkbox"/> GSSAPI
Kerberos Service Name	<input checked="" type="checkbox"/> No value set
Kerberos Credentials Service	<input checked="" type="checkbox"/> No value set
Kerberos Principal	<input checked="" type="checkbox"/> No value set
Kerberos Keytab	<input checked="" type="checkbox"/> No value set
SSL Context Service	<input checked="" type="checkbox"/> No value set
Topic Name	<input checked="" type="checkbox"/> nifimessage
Delivery Guarantee	<input checked="" type="checkbox"/> Guarantee Replicated Delivery
Failure Strategy	<input checked="" type="checkbox"/> Route to Failure
Use Transactions	<input checked="" type="checkbox"/> false

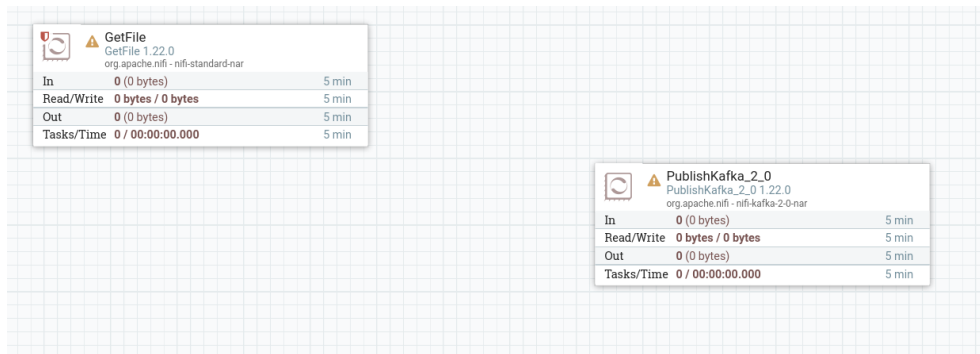
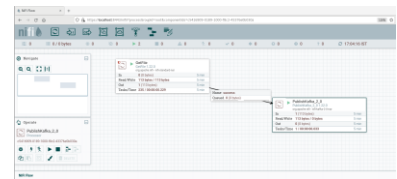
**CANCEL** **APPLY**

Click on **APPLY**

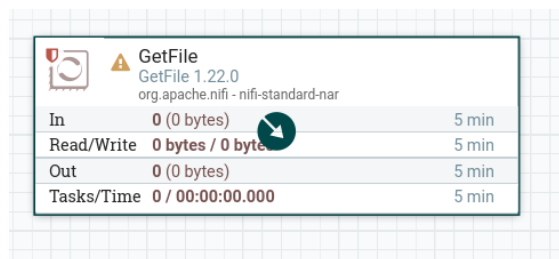
**By Sai Kumar**

## Step 5

### Creating relationship



Move the cursor to center of GetFile to get relationship Arrow and drag it to Publish kafka to create relationship.



Create Connection

DETAILS SETTINGS

From Processor  
GetFile  
GetFile

Within Group  
NIFI Flow

For Relationships  
☒ success

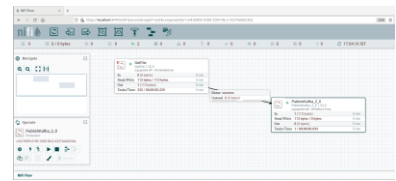
To Processor  
PublishKafka\_2\_0  
PublishKafka\_2\_0

Within Group  
NIFI Flow

CANCEL ADD

Click on **ADD**

**By Sai Kumar**



Right click on Publish kafka and select Configure

**Configure Processor** | PublishKafka\_2\_0 1.22.0

Stopped

**SETTINGS** | SCHEDULING | PROPERTIES | RELATIONSHIPS | COMMENTS

Automatically Terminate / Retry Relationships ⓘ

**failure**

☒ terminate ☐ retry

Any FlowFile that cannot be sent to Kafka will be routed to this Relationship

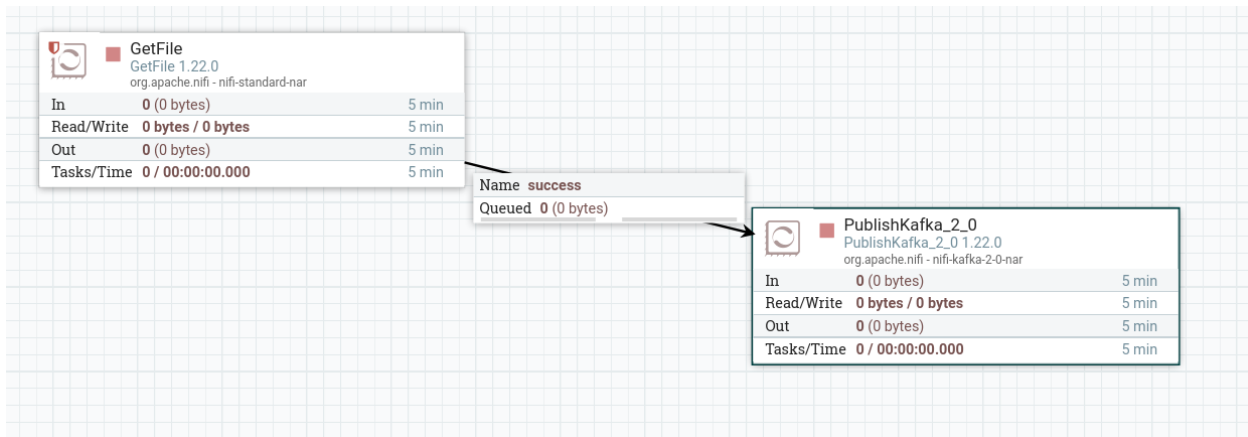
**success**

☒ terminate ☐ retry

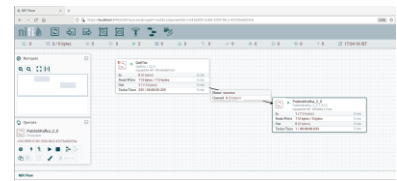
FlowFiles for which all content was sent to Kafka.

CANCEL APPLY

Tick on terminate under failure and success, and click on **APPLY**



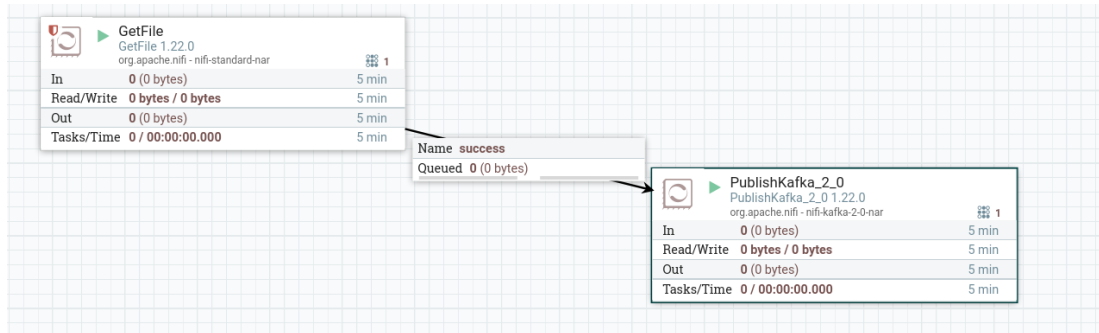
**By Sai Kumar**



## STEP 6

Start the consumer and leave the terminal.

Start GetFile and PublishKafka, by right clicking on both the processors one by one and selecting Start. After start it looks like below.

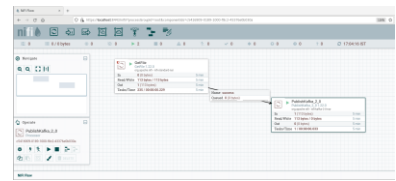


## STEP 7

Now copy the provided people.txt into /root/kafkanifi

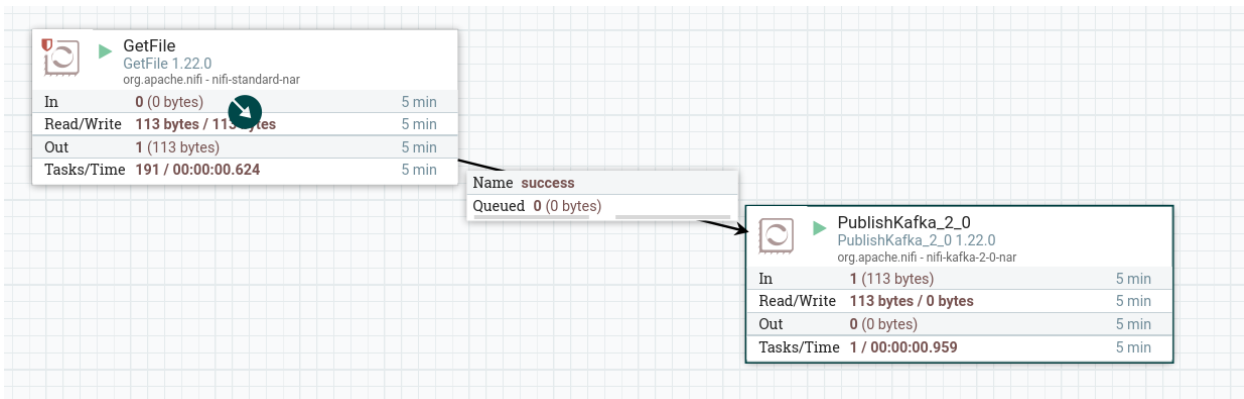
```
|# cp people.txt /root/kafkanifi/
```

**By Sai Kumar**



## STEP 8

Check the data., showing bytes in **Read/Write**.



Check the consumer

```
[root@saibigdata ~]# kafka-console-consumer --bootstrap-server saibigdata.training.com:9092 --topic nifimessage
02134,Hopper,Grace,52
94020,Turing,Alan,32
94020,Lovelace,Ada,28
87501,Babbage,Charles,49
02134,Wirth,Niklaus,48
```

**\*\*\* Happy Learning \*\*\***