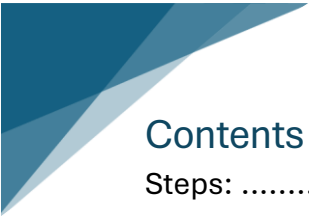


NIFI

Rana Hassan Hosny

GIZA SYSTEMS | RANA.HOSNY@GIZASYSTEMS.ONMICROSOFT.COM



Contents


Steps:2

Final result:6

Steps:

1- Make those configurations for ListSFTP:


- Hostname: SFTP
- Port: 22
- Username: foo
- Password: pass
- Remote Path: ./upload
- Search Recursively: true



ListSFTP
ListSFTP 2.0.0
org.apache.nifi - nifi-standard-nar

In	0 (0 bytes)	5 min
Read/Write	0 bytes / 0 bytes	5 min
Out	0 (0 bytes)	5 min
Tasks/Time	27 / 00:00:08.763	5 min

2- Fetch content using FetchSFTP:



FetchSFTP
FetchSFTP 2.0.0
org.apache.nifi - nifi-standard-nar

In	0 (0 bytes)	5 min
Read/Write	0 bytes / 0 bytes	5 min
Out	0 (0 bytes)	5 min
Tasks/Time	0 / 00:00:00.000	5 min










Processor Details

Settings



Scheduling

Properties

Required field

Property	Value
Hostname	 sftp
Port	 22
Username	 foo
Password	 Sensitive value set
Private Key Path	 No value set
Private Key Passphrase	 No value set
Remote File	 \${path}/\${filename}
Completion Strategy	 None
Move Destination Directory	 No value set

3- Split the reading per line using SplitText:

	 SplitText SplitText 2.0.0 org.apache.nifi - nifi-standard-nar	
In	0 (0 bytes)	5 min
Read/Write	0 bytes / 0 bytes	5 min
Out	0 (0 bytes)	5 min
Tasks/Time	0 / 00:00:00.000	5 min

Processor Details






Settings

Scheduling


Properties

Re

Required field

Property	Value
Line Split Count	 1
Maximum Fragment Size	 No value set
Header Line Count	 0
Header Line Marker Characters	 No value set
Remove Trailing Newlines	 true

4- Convert to json using ReplaceText:

	ReplaceText ReplaceText 2.0.0 org.apache.nifi - nifi-standard-nar	
In	0 (0 bytes)	5 min
Read/Write	0 bytes / 0 bytes	5 min
Out	0 (0 bytes)	5 min
Tasks/Time	0 / 00:00:00.000	5 min



Processor Details

Settings

Scheduling

Properties

Rela

Required field

Vt

Property	Value
Replacement Strategy	Regex Replace
Search Value	(.+?) ,([^\,]+)\$
Replacement Value	{"timestamp": "\$1", "value": "\$2"}
Character Set	UTF-8
Maximum Buffer Size	1 MB
Evaluation Mode	Line-by-Line
Line-by-Line Evaluation Mode	All

5- Extract the device name and add it into the record using UpdateRecord:

- Create two services JsonTreeReader and JsonRecordSetWriter and enable them.
- Use them in the configuration:
- Add property /device and this parm: `${path:replaceAll('.*(device\\d+).*', '$1')}`

	UpdateRecord UpdateRecord 2.0.0 org.apache.nifi - nifi-standard-nar
In	0 (0 bytes) 5 min
Read/Write	0 bytes / 0 bytes 5 min
Out	0 (0 bytes) 5 min
Tasks/Time	0 / 00:00:00.000 5 min

Processor Details

Settings

Scheduling

Properties

Rela

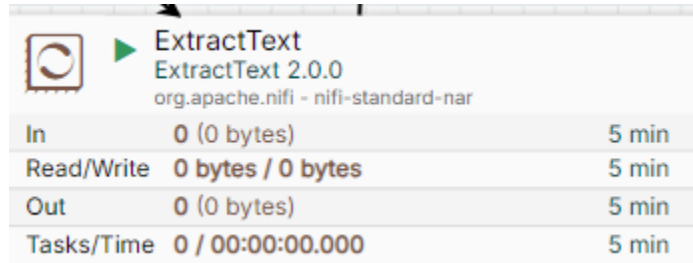
Required field

V

Property	Value
Record Reader	JsonTreeReader
Record Writer	JsonRecordSetWriter
Replacement Value Strategy	Literal Value
/device	<code>\${path:replaceAll('.*(device\\d+).*', '\$1')}</code>

6- Extract the timestamp as attribute using ExtractText:

- Create property timestamp with those parameters: `(\w{3} \w{3} \d{2} \d{4} \s* \d{2}:\d{2}:\d{2})`

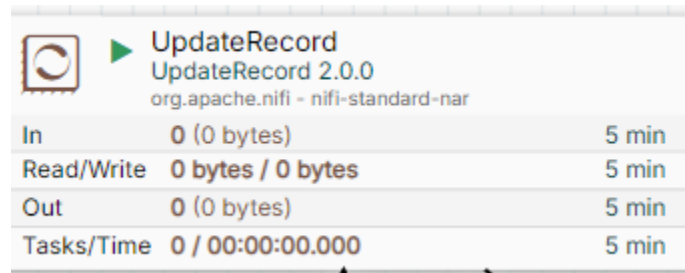


The screenshot shows the 'ExtractText' processor interface. It has a green play button icon and the text 'ExtractText 2.0.0' and 'org.apache.nifi - nifi-standard-nar'. Below this is a table with the following data:

In	0 (0 bytes)	5 min
Read/Write	0 bytes / 0 bytes	5 min
Out	0 (0 bytes)	5 min
Tasks/Time	0 / 00:00:00.000	5 min

7- Convert the timestamp into using UpdateRecord:

- Create /timestamp property with those parameters:
`${timestamp:toDate("EEE MMM dd yyyy HH:mm:ss", "GMT"):format("yyyy-MM-dd'T'HH:mm:ssXXX")}`

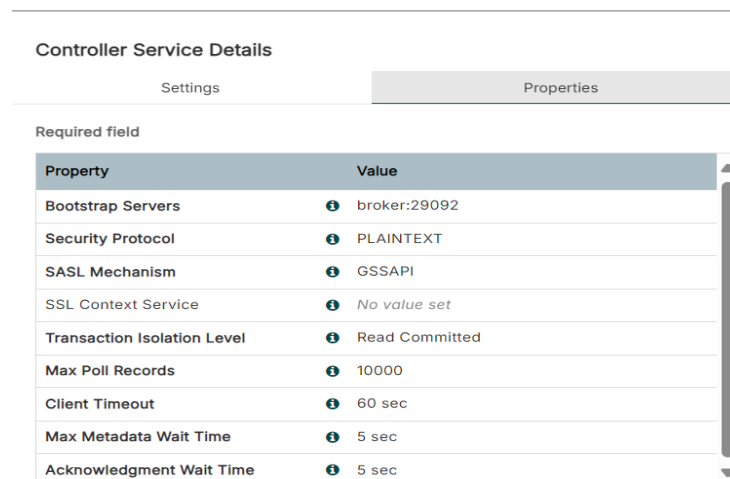


The screenshot shows the 'UpdateRecord' processor interface. It has a green play button icon and the text 'UpdateRecord 2.0.0' and 'org.apache.nifi - nifi-standard-nar'. Below this is a table with the following data:

In	0 (0 bytes)	5 min
Read/Write	0 bytes / 0 bytes	5 min
Out	0 (0 bytes)	5 min
Tasks/Time	0 / 00:00:00.000	5 min

8- Create Publishkafka to publish in test topic :


- You need to create Kafka3ConnectionService at first with those parmeters :



The screenshot shows the 'Controller Service Details' page for 'Kafka3ConnectionService'. It has two tabs: 'Settings' and 'Properties'. The 'Properties' tab is selected. Below the tabs is a table with the following data:

Property	Value
Bootstrap Servers	broker:29092
Security Protocol	PLAINTEXT
SASL Mechanism	GSSAPI
SSL Context Service	No value set
Transaction Isolation Level	Read Committed
Max Poll Records	10000
Client Timeout	60 sec
Max Metadata Wait Time	5 sec
Acknowledgment Wait Time	5 sec

- Then create the PublishKafka:

	PublishKafka PublishKafka 2.0.0 org.apache.nifi - nifi-kafka-nar	
In	0 (0 bytes)	5 min
Read/Write	0 bytes / 0 bytes	5 min
Out	0 (0 bytes)	5 min
Tasks/Time	0 / 00:00:00.000	5 min

Processor Details











Settings

Scheduling

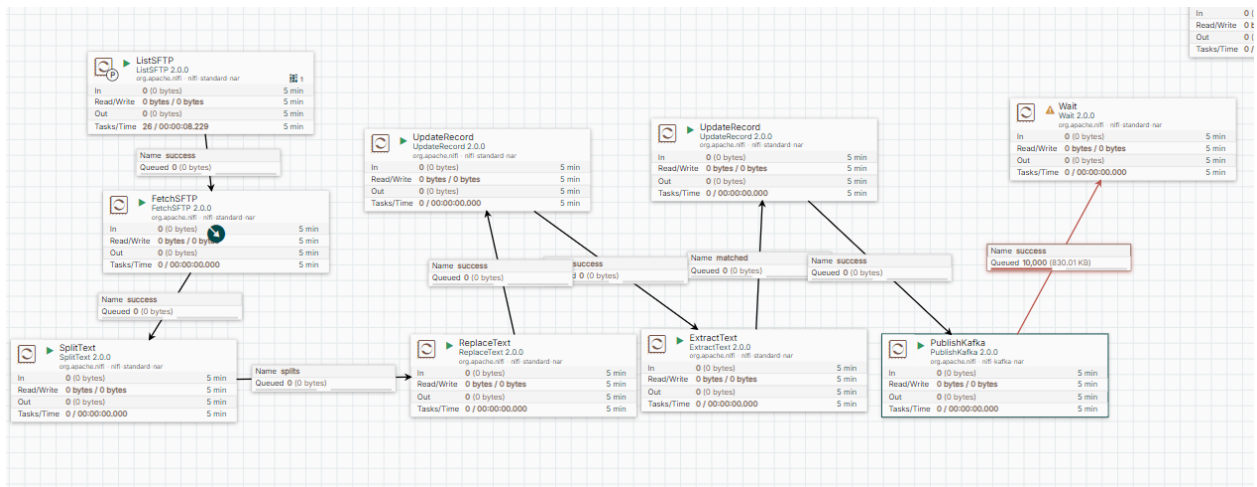
Properties

Rel

Required field

Property	Value
Kafka Connection Service	 Kafka3ConnectionService 
Topic Name	 test
Failure Strategy	 Route to Failure
Delivery Guarantee	 Guarantee Replicated Delivery
Compression Type	 none
Max Request Size	 1 MB
Transactions Enabled	 true
Transactional ID Prefix	 No value set
Partitioner Class	 DefaultPartitioner

Final result:





Cluster overview

Brokers

Topics

Connect

ksqDB

Consumers

Replicators

Cluster settings

Health+

test

ConfigurationMessagesSchema

Production

Consumption

Total messages

Retention time

Filter by timestamp, offset, key or value

All partitions

Latest

Max 50 results

20 messages shownAuto-refresh on

Timestamp	Offset	Partition	Key	Value
1734955695373	10513	0	""	["timestamp":"2024-10-27T12:03:10Z","value":"PR38-DA62YDK","device":"device3"]
1734955695373	10514	0	""	["timestamp":"2024-10-27T12:03:10Z","value":"PR38-DA62Y5UR","device":"device3"]
1734955695373	10515	0	""	["timestamp":"2024-10-27T12:03:10Z","value":"PR38-DA62Y57P","device":"device3"]
1734955695370	10500	0	""	["timestamp":"2024-10-27T12:03:07Z","value":"PR38-DA62YHJN","device":"device3"]
1734955695370	10501	0	""	["timestamp":"2024-10-27T12:03:07Z","value":"PR38-DA62YHVG","device":"device3"]

IRON

3,10-

Search

ENG

2:45 PM

12/23/2024