

Файл Nifi-scheme.xml и DAG файл airflow_dag.py к работе приложены

Разворачивание Docker

```
/usr/bin/bash --login -i
bash-4.4$ cd C:\docker_ssau
bash-4.4$ git clone https://github.com/ssau-data-engineering/Prerequisites.git
Cloning into 'Prerequisites'...
remote: Enumerating objects: 69, done.
remote: Counting objects: 100% (69/69), done.
remote: Compressing objects: 100% (44/44), done.
remote: Total 69 (delta 28), reused 56 (delta 19), pack-reused 0
Receiving objects: 100% (69/69), 90.39 KiB | 675.00 KiB/s, done.
Resolving deltas: 100% (28/28), done.
bash-4.4$

bash-4.4$ cd Prerequisites\
>
bash-4.4$ ls
'Python Jupyter Remote Docker.code-profile'  airflow  docker-compose.elasticsearch.yaml  docker-compose.nifi.yaml  images  nifi
README.md  docker-compose.airflow.yaml  docker-compose.miflow.yaml  docker-compose.postgresql.yaml  miflow  requirements.txt
bash-4.4$
```

1. docker network create data-engineering-labs-network

```
PS C:\docker_ssau\Prerequisites> docker network create data-engineering-labs-network
08d07ade1bd93ed351d399a51769de55e8481de03193f95cb21ae3903d5d8a3b
PS C:\docker_ssau\Prerequisites> _
```

docker compose -f docker-compose.airflow.yaml up airflow-init

```
PS C:\docker_ssau\Prerequisites> docker compose -f docker-compose.airflow.yaml up airflow-init
[+] Running 21/21
redis 7 layers [#####] 0B/0B Pulled 33.0s
  1f7ce2fa46ab Pull complete 24.8s
  3c6368585bf1 Pull complete 0.8s
  3911d271d7d8 Pull complete 4.9s
  ac88aa9d4021 Pull complete 16.6s
  127cd75a68a2 Pull complete 6.1s
  4f4fb700ef54 Pull complete 7.8s
  f3993c1104fc Pull complete 9.7s
postgres 12 layers [#####] 0B/0B Pulled 79.8s
  d8e29f57d52c Pull complete 11.0s
  c8a84bacdc9c Pull complete 18.7s
  4c9c7f55c6be Pull complete 19.7s
  4f2487cd6295 Pull complete 27.7s
  5a5f050db629 Pull complete 23.9s
  691624362977 Pull complete 24.9s
  f464a58894c5 Pull complete 25.5s
  a600810358c9 Pull complete 71.7s
  285c975ab2ee Pull complete 27.0s
  58308e6e68e7 Pull complete 27.9s
  6277814ae501 Pull complete 28.4s
  607fd62806ab Pull complete 76.8s
```

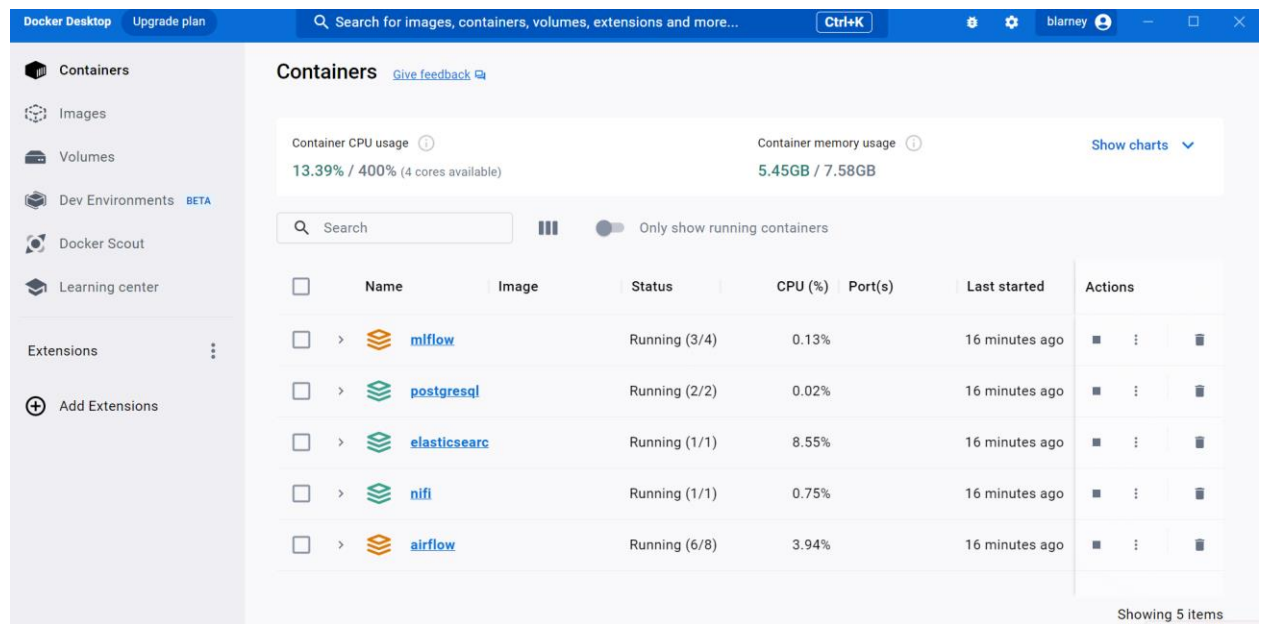
docker compose -f docker-compose.airflow.yaml up --build -d

```

PS C:\docker_ssau\Prerequisites> docker compose -f docker-compose.airflow.yaml up --build -d
[+] Running 3/15
 - docker-proxy 14 layers [#####] 6.598MB/51.41MB Pulling 29.4s
   ✓ 96526aa774ef Pull complete 14.4s
   ✓ 979f7a62ca1b Pull complete 15.4s
   ✓ 4f4fb700ef54 Pull complete 0.8s
 - fb84691a0857 Downloading [=====>] 3.55MB/... 26.2s
 - 18b1b5d57ca7 Downloading [=====>] 2.14MB/... 26.2s
 - ce1e8d1970d0 Downloading [==>] 907.6kB/... 26.2s
 - 4ffc5f9802ef Waiting 26.2s
 - d3856c109dcb Waiting 26.2s
 - 49f0f622e61c Waiting 26.2s
 - 81640b6a1021 Waiting 26.2s
 - 8c7513b91529 Waiting 26.2s
 - 6aa916c39934 Waiting 26.2s
 - 47a389dfac47 Waiting 26.2s
 - 7832b6a0e052 Waiting 26.2s

```

Запускаем остальные контейнеры

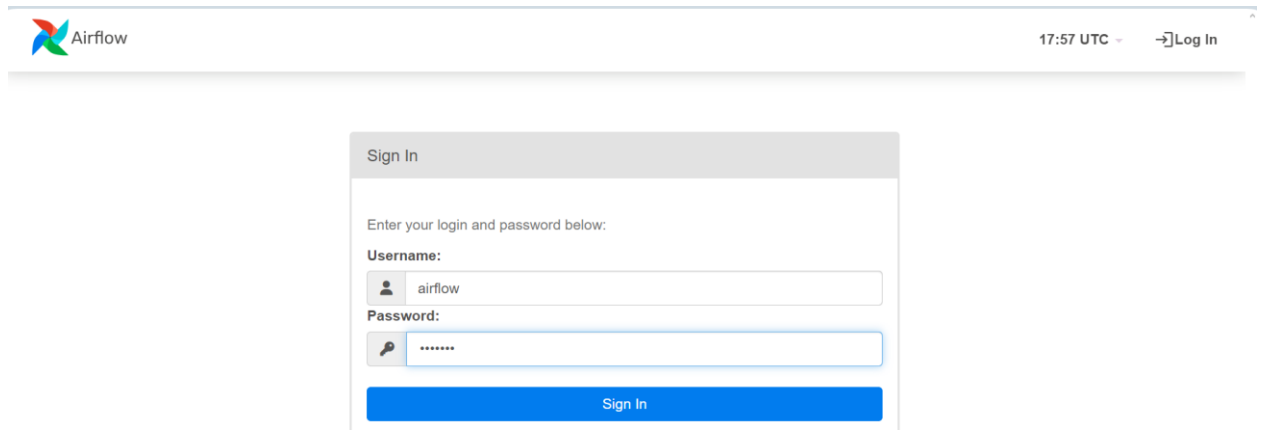


AIRFLOW

Создаём DAG и сохраняем его по пути C:\docker_ssau\Prerequisites\airflow\dags

```
File Edit Selection View Go Run Terminal Help • airflow_dag.py - Visual Studio Code
EXPLORER
NO FOLDER OPENED
You have not yet opened a folder.
Open Folder
Opening a folder will close all currently open editors. To keep them open, add a folder instead.
OUTLINE
TIMELINE
C:\> docker_ssau > Prerequisites > airflow > dags > airflow_dag.py > ...
1 from airflow import DAG
2 from airflow.operators.python import PythonOperator
3 from elasticsearch import Elasticsearch
4 from datetime import datetime, timedelta
5 import pandas as pd
6 import numpy as np
7 import uuid
8
9 default_args = {
10     'owner': 'vlados',
11     'start_date': datetime(2023, 12, 2),
12 }
13
14 dag = DAG(
15     'airflow_sender',
16     default_args=default_args,
17     description='This DAG will send some data to Elasticsearch',
18     catchup=False,
19 )
20
21
22 def data_download():
23     dag_dataframe = pd.DataFrame()
24     for i in range(26):
25         chunk = pd.read_csv(f"/opt/airflow/data/chunk{i}.csv")
26         dag_dataframe = pd.concat([dag_dataframe, chunk])
27     return dag_dataframe
28
29
30
```

Заходим по ссылке <http://localhost:8080/> на Airflow логин airflow пароль airflow



Вот мы и на главной странице

DAGs

All 54		Active 0	Paused 54	Running 0	Failed 0	Filter DAGs by tag	Search DAGs	<input checked="" type="checkbox"/> Auto-refresh	
DAG	Owner	Runs	Schedule	Last Run	Next Run	Recent Tasks			
airflow_sender	vlados		1 day, 0:00:00		2023-12-02, 00:00:00				
dataset_consumes_1 <div>consumes dataset-scheduled</div>	airflow		Dataset		On s3://dag1/output_1.txt				
dataset_consumes_1_and_2 <div>consumes dataset-scheduled</div>	airflow		Dataset		0 of 2 datasets updated				
dataset_consumes_1_never_scheduled <div>consumes dataset-scheduled</div>	airflow		Dataset		0 of 2 datasets updated				
dataset_consumes_unknown_never_scheduled <div>dataset-scheduled</div>	airflow		Dataset		0 of 2 datasets updated				

Загружаем доки

Этот компьютер > Acer (C:) > docker_ssau > Prerequisites > airflow > data

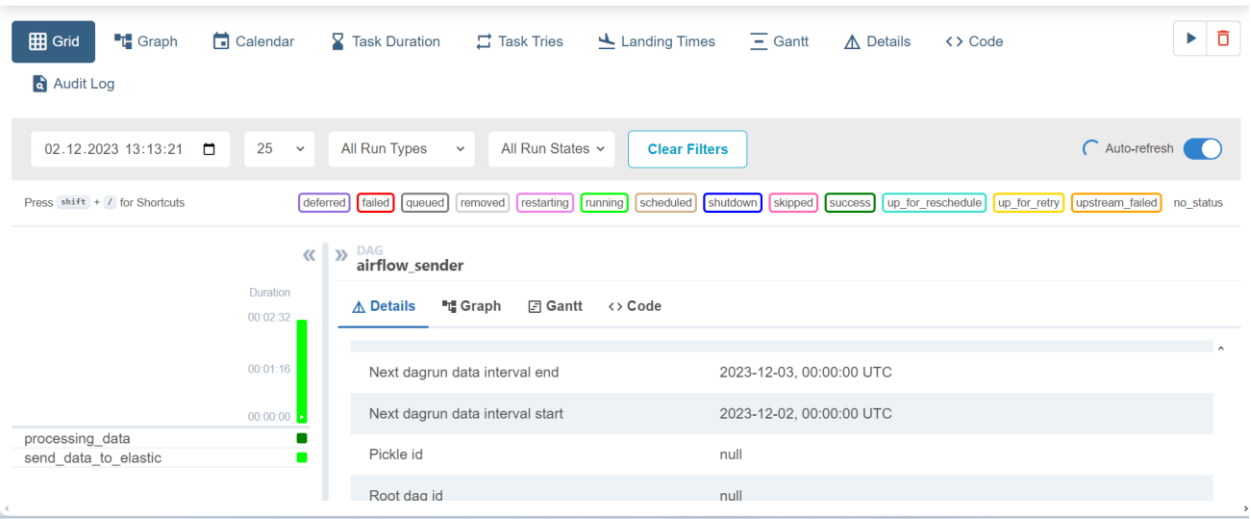
Поиск в: data

Сортировать Просмотреть

Сведения

Имя	Дата изменения	Тип	Размер
chunk0.csv	01.12.2023 23:52	Файл Microsoft Excel...	1 952 КБ
chunk1.csv	01.12.2023 23:52	Файл Microsoft Excel...	2 012 КБ
chunk2.csv	01.12.2023 23:52	Файл Microsoft Excel...	1 965 КБ
chunk3.csv	01.12.2023 23:52	Файл Microsoft Excel...	2 001 КБ
chunk4.csv	01.12.2023 23:52	Файл Microsoft Excel...	1 972 КБ
chunk5.csv	01.12.2023 23:52	Файл Microsoft Excel...	1 986 КБ
chunk6.csv	01.12.2023 23:52	Файл Microsoft Excel...	1 984 КБ
chunk7.csv	01.12.2023 23:52	Файл Microsoft Excel...	2 016 КБ
chunk8.csv	01.12.2023 23:52	Файл Microsoft Excel...	1 965 КБ
chunk9.csv	01.12.2023 23:52	Файл Microsoft Excel...	2 012 КБ
chunk10.csv	01.12.2023 23:52	Файл Microsoft Excel...	1 990 КБ
chunk11.csv	01.12.2023 23:52	Файл Microsoft Excel...	2 006 КБ
chunk12.csv	01.12.2023 23:52	Файл Microsoft Excel...	1 948 КБ
chunk13.csv	01.12.2023 23:52	Файл Microsoft Excel...	1 980 КБ
chunk14.csv	01.12.2023 23:52	Файл Microsoft Excel...	1 996 КБ

Заходим в DAG и запускаем



Этот компьютер

Acer (C:)

docker_ssau

Prerequisites

airflow

data

Поиск в: data

Сортировать

Просмотреть

...

Имя

Дата изменения

Тип

Размер

chunk11.csv

01.12.2023 23:52

Файл Microsoft Excel...

2 006 КБ

chunk12.csv

01.12.2023 23:52

Файл Microsoft Excel...

1 948 КБ

chunk13.csv

01.12.2023 23:52

Файл Microsoft Excel...

1 980 КБ

chunk14.csv

01.12.2023 23:52

Файл Microsoft Excel...

1 996 КБ

chunk15.csv

01.12.2023 23:52

Файл Microsoft Excel...

2 000 КБ

chunk16.csv

01.12.2023 23:52

Файл Microsoft Excel...

2 006 КБ

chunk17.csv

01.12.2023 23:52

Файл Microsoft Excel...

2 006 КБ

chunk18.csv

01.12.2023 23:52

Файл Microsoft Excel...

1 969 КБ

chunk19.csv

01.12.2023 23:52

Файл Microsoft Excel...

1 973 КБ

chunk20.csv

01.12.2023 23:52

Файл Microsoft Excel...

2 000 КБ

chunk21.csv

01.12.2023 23:52

Файл Microsoft Excel...

1 984 КБ

chunk22.csv

01.12.2023 23:52

Файл Microsoft Excel...

1 959 КБ

chunk23.csv

01.12.2023 23:52

Файл Microsoft Excel...

2 006 КБ

chunk24.csv

01.12.2023 23:52

Файл Microsoft Excel...

2 000 КБ

chunk25.csv

01.12.2023 23:52

Файл Microsoft Excel...

1 994 КБ

data.csv

02.12.2023 17:13

Файл Microsoft Excel...

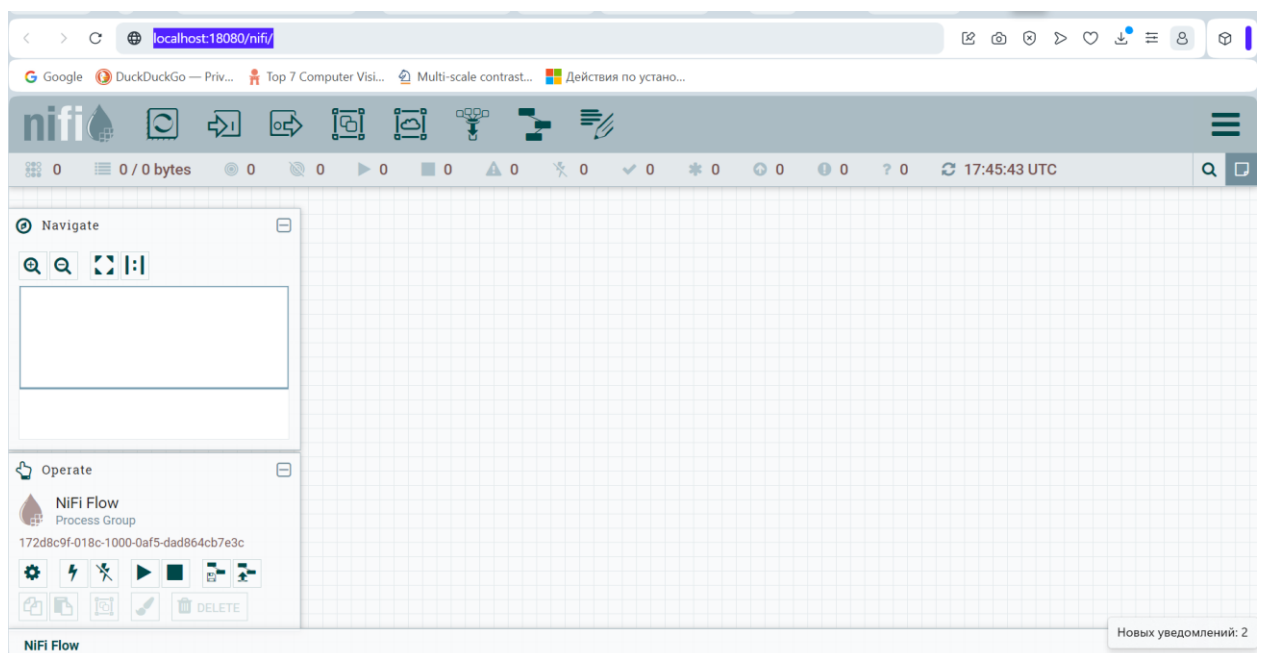
30 561 КБ

ан 1 элемент: 29.8 МБ

НИФИ

1. Заходим на Nifi

<http://localhost:18080/nifi/>



1. Настраиваем GetFile

Выбираем тип процессора. Делать это придётся при создании каждого процессора, но покажу единожды

Add Processor


Source

Displaying 1 of 349

GetFi

all groups ▾

amazon attributes
avro aws azure
cloud consume
csv fetch get
google ingest
json listen logs
message
microsoft pubsub
put record
restricted source
storage text
update

Type ▴	Version	Tags
 GetFile	1.23.2	ingress, input, restricted, get, fil...

GetFile 1.23.2 org.apache.nifi - nifi-standard-nar

Creates FlowFiles from files in a directory. NiFi will ignore files it doesn't have at least read permissions for.

CANCEL

ADD

Configure Processor | GetFile 1.23.2

 Stopped

SETTINGS

SCHEDULING













PROPERTIES

RELATIONSHIPS

COMMENTS

Required field



Property	Value	
Input Directory	 /opt/nifi/nifi-current/data	
File Filter	 [^\\.]*	
Path Filter	 No value set	
Batch Size	 100	
Keep Source File	 false	
Recurse Subdirectories	 true	
Polling Interval	 0 sec	
Ignore Hidden Files	 true	
Minimum File Age	 0 sec	
Maximum File Age	 No value set	
Minimum File Size	 0 B	
Maximum File Size	 No value set	

CANCEL

APPLY

Тут ничего не трогаем

SETTINGS | SCHEDULING | PROPERTIES | RELATIONSHIPS | COMMENTS

Automatically Terminate / Retry Relationships ⓘ

success

☐ terminate ☐ retry

All files are routed to success

2. Заполняем splitRecord

Для этого в Properties заполняем выбираем Create new service

Configure Processor | SplitRecord 1.23.2

Invalid

SETTINGS | SCHEDULING | PROPERTIES | RELATIONSHIPS | COMMENTS

Required field ✔ +

Property	Value
Record Reader	
Record Writer	
Records Per Split	

Reference parameter...
Reference parameter...
Create new service...

CANCELOK

CANCELAPPLY

Тут выбираем Controller Service

Add Controller Service

Requires Controller Service

RecordSetWriterFactory 1.23.2 from org.apache.nifi - nifi-standard-services-api-nar

Compatible Controller Services

CSVRecordSetWriter 1.23.2	✓
AvroRecordSetWriter 1.23.2	?
CSVRecordSetWriter 1.23.2	?
FreeFormTextRecordSetWriter 1.23.2	?
JsonRecordSetWriter 1.23.2	?
ParquetRecordSetWriter 1.23.2	?
RecordSetWriterLookup 1.23.2	?
ScriptedRecordSetWriter 1.23.2	?

Итак, вкладка Properties выглядит вот так

Configure Processor | SplitRecord 1.23.2

⚠ Invalid

SETTINGS	SCHEDULING	PROPERTIES	RELATIONSHIPS	COMMENTS
----------	------------	------------	---------------	----------

Required field ✓ +

Property	Value
Record Reader	CSVReader →
Record Writer	CSVRecordSetWriter →
Records Per Split	100000

CANCEL

APPLY

Задаём Relationship

 Invalid

SETTINGS

SCHEDULING

PROPERTIES

RELATIONSHIPS

COMMENTS

Automatically Terminate / Retry Relationships **failure**☒ terminate ☐ retry

If a FlowFile cannot be transformed from the configured input format to the configured output format, the unchanged FlowFile will be routed to this relationship.

original☒ terminate ☐ retry

Upon successfully splitting an input FlowFile, the original FlowFile will be sent to this relationship.

splits☐ terminate ☐ retry

The individual 'segments' of the original FlowFile will be routed to this relationship.

CANCEL

APPLY

2 QueryRecord (первый)

Поскольку мы прежде уже выбрали это значение Value в SplitRecord, нам предлагается выбрать его снова. Так и поступим

Invalid

SETTINGS | SCHEDULING | PROPERTIES | RELATIONSHIPS | COMMENTS

Required field

Property	Value
Record Reader	
Record Writer	CSVRecordSetWriter
Include Zero Record FlowFiles	
Cache Schema	
Default Decimal Precision	10
Default Decimal Scale	0

CANCEL APPLY

Добавим пользовательское значение correct и имеем вид

Invalid

SETTINGS | SCHEDULING | PROPERTIES | RELATIONSHIPS | COMMENTS

Required field

Property	Value
Record Reader	CSVReader
Record Writer	CSVRecordSetWriter
Include Zero Record FlowFiles	true
Cache Schema	true
Default Decimal Precision	10
Default Decimal Scale	0
correct	SELECT * FROM FLOWFILE WHERE designation SELEC...

CANCEL APPLY

correct = SELECT * FROM FLOWFILE WHERE designation <> " AND region_1
<>"

 Invalid

SETTINGS

SCHEDULING

PROPERTIES

RELATIONSHIPS

COMMENTS

Automatically Terminate / Retry Relationships 

correct

☐ terminate ☐ retry

User-defined relationship that specifies where data that matches the specified SQL query should be routed

failure

☒ terminate ☐ retry

If a FlowFile fails processing for any reason (for example, the SQL statement contains columns not present in input data), the original FlowFile it will be routed to this relationship

original

☒ terminate ☐ retry

The original FlowFile is routed to this relationship

CANCEL

APPLY

4 Теперь настраиваем QueryRecord (второй)

Добавим пару пользовательских значений

Configure Processor | QueryRecord 1.23.2

Invalid

SETTINGS

SCHEDULING

PROPERTIES

RELATIONSHIPS

COMMENTS

Required field

Property

Record Reader

Record Writer

Include Zero Record FlowFiles

Cache Schema

Default Decimal Precision

Default Decimal Scale

→

→

Add Property

Property Name

not_null_price

Sensitive Value ?

Yes

No

CANCEL

OK

CANCEL

APPLY

Итого у нас получилось

Configure Processor | QueryRecord 1.23.2

Invalid

SETTINGS

SCHEDULING

PROPERTIES

RELATIONSHIPS

COMMENTS

Required field

Property

Record Reader

Record Writer

Include Zero Record FlowFiles

Cache Schema

Default Decimal Precision

Default Decimal Scale

not_null_price

null_price

?

?

?

?

?

?

?

?

CSVReader

CSVRecordSetWriter

true

true

10

0

SELECT * FROM FLOWFILE WHERE price is not null

SELECT * FROM FLOWFILE WHERE price is null

→

→

🗑

🗑

CANCEL

APPLY

Перезаходим в настройку процессора и понимаем, что в Relationship появилось два новых значения, но трогать мы их не будем

Configure Processor

QueryRecord 1.23.2

Invalid

SETTINGS

SCHEDULING

PROPERTIES

RELATIONSHIPS

COMMENTS

failure

☒ terminate

☐ retry

If a FlowFile fails processing for any reason (for example, the SQL statement contains columns not present in input data), the original FlowFile it will be routed to this relationship

not_null_price

☐ terminate

☐ retry

User-defined relationship that specifies where data that matches the specified SQL query should be routed

null_price

☐ terminate

☐ retry

User-defined relationship that specifies where data that matches the specified SQL query should be routed

original

☒ terminate

☐ retry

The original FlowFile is routed to this relationship

CANCEL

APPLY

5 Настраиваем UpdateRecord

Properties

Configure Processor

UpdateRecord 1.23.2

Invalid

SETTINGS

SCHEDULING

PROPERTIES

RELATIONSHIPS

COMMENTS

Required field

☒

☐

+

Property	Value
Record Reader	<div><div>?</div>CSVReader</div> <div>→</div>
Record Writer	<div><div>?</div>CSVRecordSetWriter</div> <div>→</div>
Replacement Value Strategy	<div><div>?</div>Literal Value</div> <div></div>
/price	<div><div>?</div>0.0</div> <div><div>🗑</div></div>

Relationships (На этот раз дополнительная строчка по имени пользовательского Property не создалась)

Configure Processor | UpdateRecord 1.23.2

Invalid

SETTINGS

SCHEDULING

PROPERTIES

RELATIONSHIPS

COMMENTS

Automatically Terminate / Retry Relationships

failure

☒ terminate

☐ retry

If a FlowFile cannot be transformed from the configured input format to the configured output format, the unchanged FlowFile will be routed to this relationship

success

☐ terminate

☐ retry

FlowFiles that are successfully transformed will be routed to this relationship

CANCEL

APPLY

6 MergeContent

В Properties настраиваем Delimiter Strategy

Invalid

SETTINGS

SCHEDULING

PROPERTIES

RELATIONSHIPS

COMMENTS

Required field

✔

+

Property		Value	
Attribute Strategy	?	Keep Only Common Attributes	^
Correlation Attribute Name	?	No value set	
Minimum Number of Entries	?	1	
Maximum Number of Entries	?	1000	
Minimum Group Size	?	0 B	
Maximum Group Size	?	No value set	
Max Bin Age	?	No value set	
Maximum number of Bins	?	5	
Delimiter Strategy	?	Text	
Header	?	id,country,description,designation,points,price,provinc...	
Footer	?	No value set	
Demarcator	?	No value set	▼

CANCEL

APPLY

В Header значение:
id,country,description,designation,points,price,province,region_1,region_2,taster_name,taster_
twitter_handle,title,variety,winery

 Invalid

SETTINGS

SCHEDULING

PROPERTIES

RELATIONSHIPS

COMMENTS

Automatically Terminate / Retry Relationships 

failure

☒ terminate ☐ retry

If the bundle cannot be created, all FlowFiles that would have been used to created the bundle will be transferred to failure

merged

☐ terminate ☐ retry

The FlowFile containing the merged content

original

☒ terminate ☐ retry

The FlowFiles that were used to create the bundle

CANCEL

APPLY

7 UpdateAttribute

Configure Processor | UpdateAttribute 1.23.2

 Invalid

- SETTINGS

SCHEDULING







PROPERTIES

RELATIONSHIPS

COMMENTS

Required field



Property		Value	
Delete Attributes Expression		No value set	
Store State		Do not store state	
Stateful Variables Initial Value		No value set	
Cache Value Lookup Cache Size		100	
filename		<code>\${fragment.identifier}_\${random()}.csv</code>	

 ADVANCED

CANCELAPPLY

filename -= `${fragment.identifier}_${random()}.csv`

Configure Processor | UpdateAttribute 1.23.2

 Invalid

- SETTINGS

SCHEDULING

PROPERTIES

RELATIONSHIPS


COMMENTS

Automatically Terminate / Retry Relationships 

success

☒ terminate ☐ retry

All successful FlowFiles are routed to this relationship

 ADVANCED

CANCELAPPLY

8 PutElasticsearchRecord

Configure Processor

PutElasticsearchRecord 1.23.2

Invalid

SETTINGS

SCHEDULING

PROPERTIES

RELATIONSHIPS

COMMENTS

Required field

+

Property		Value	
Index Operation	?	index	
Index	?	nifi	
Type	?	_doc	
@timestamp Value	?	No value set	
Client Service	?	ElasticSearchClientServiceImpl	→
Record Reader	?	CSVReader	→
Batch Size	?	100	
ID Record Path	?	No value set	
Index Operation Record Path	?	No value set	
Index Record Path	?	No value set	
Type Record Path	?	No value set	
@timestamp Record Path	?	No value set	

CANCEL

APPLY

Client Service = ElasticSearchClientServiceImpl

Configure Processor | PutElasticsearchRecord 1.23.2

 Invalid

SETTINGS

SCHEDULING

PROPERTIES

RELATIONSHIPS

COMMENTS

Automatically Terminate / Retry Relationships 

errors

☒ terminate ☐ retry

If a "Result Record Writer" is set, any Record(s) corresponding to Elasticsearch document(s) that resulted in an "error" (within Elasticsearch) will be routed here.

failure

☒ terminate ☐ retry

All flowfiles that fail for reasons unrelated to server availability go to this relationship.

retry

☒ terminate ☐ retry

All flowfiles that fail due to server/cluster availability go to this relationship.

success

☒ terminate ☐ retry

CANCEL

APPLY

9 PutFile

Configure Processor | PutFile 1.23.2

 Invalid

SETTINGS

SCHEDULING









PROPERTIES

RELATIONSHIPS

COMMENTS

Required field



Property		Value	
Directory		/opt/nifi/nifi-current/data	
Conflict Resolution Strategy		fail	
Create Missing Directories		true	
Maximum File Count		No value set	
Last Modified Time		No value set	
Permissions		No value set	
Owner		No value set	
Group		No value set	

CANCEL

APPLY

Configure Processor | PutFile 1.23.2

 Invalid

SETTINGS

SCHEDULING

PROPERTIES

RELATIONSHIPS

COMMENTS

Automatically Terminate / Retry Relationships 

failure

☒ terminate ☐ retry

Files that could not be written to the output directory for some reason are transferred to this relationship

success

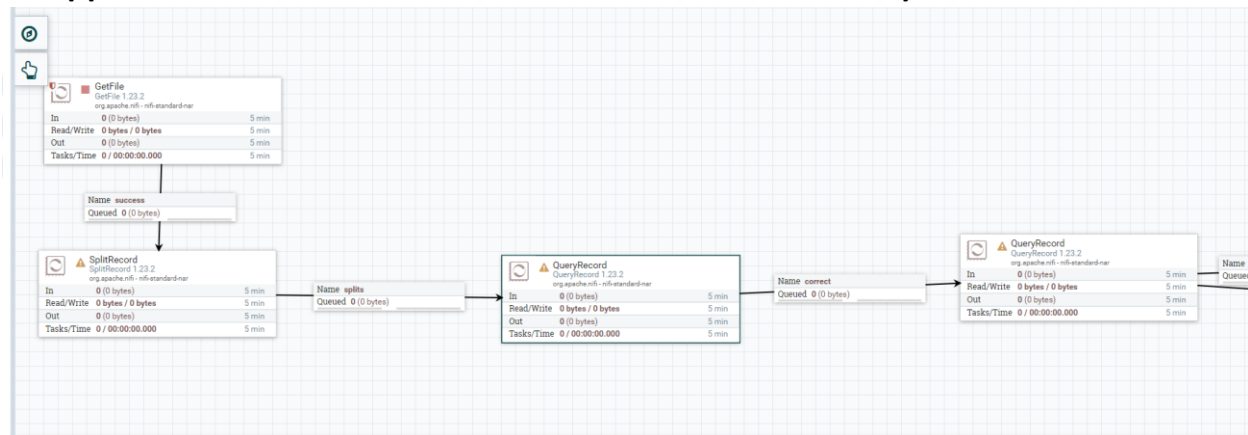
☒ terminate ☐ retry

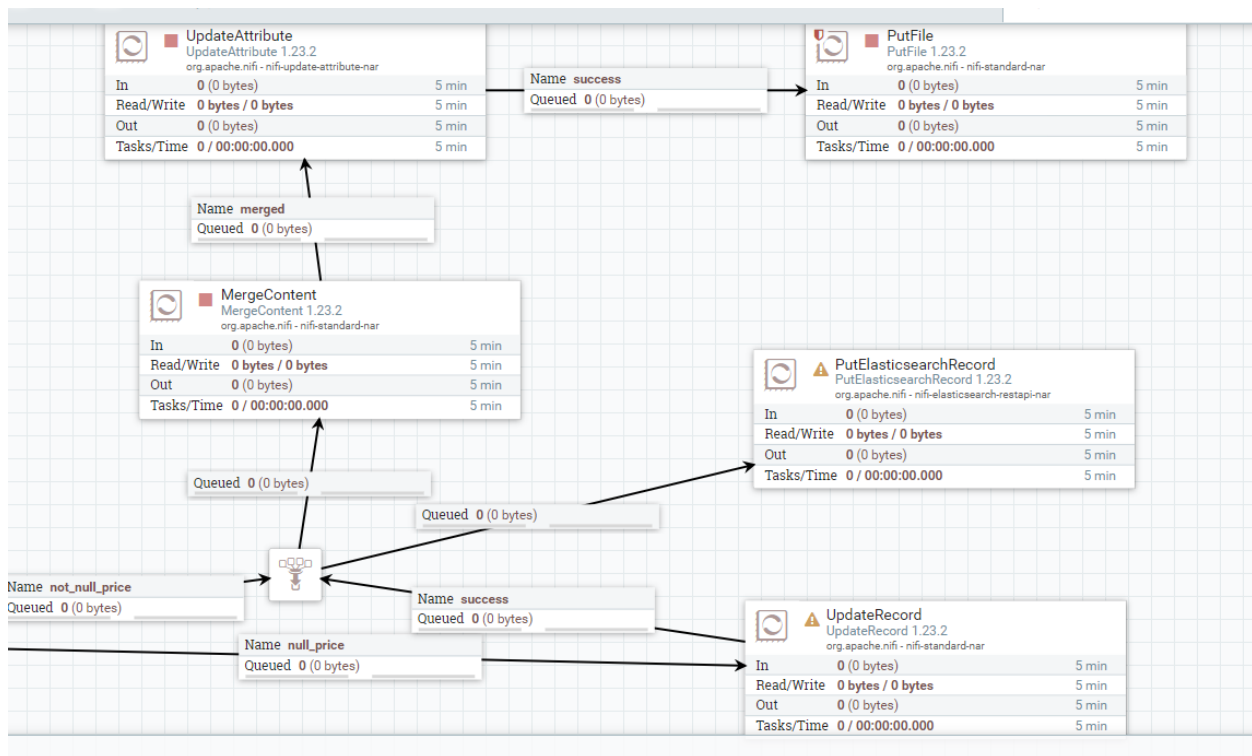
Files that have been successfully written to the output directory are transferred to this relationship

CANCEL

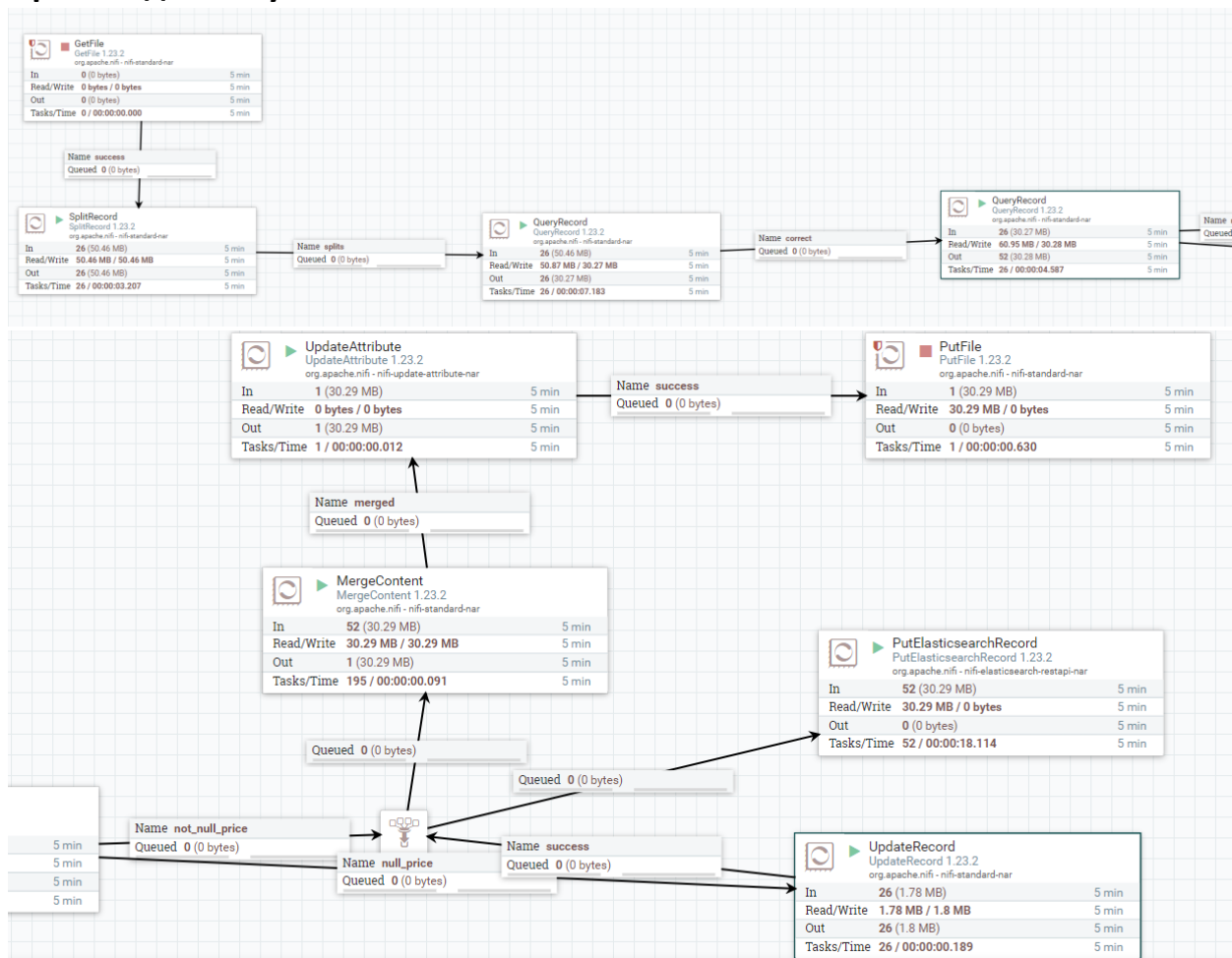
APPLY

**ИТОГО ПОЛУЧИЛАСЬ СЛЕДУЮЩАЯ СХЕМА (НА ДВУХ СКРИНШОТАХ
ВВИДУ ОСОБЕННОСТЕЙ МАСШТАБИРОВАНИЯ AIRFLOW)**





Произведён запуск



localhost:15601

localhost:15601/app/home#/?

Google DuckDuckGo — Priv... Top 7 Computer Visi... Multi-scale contrast... Действия по устано...



Welcome to Elastic



Start by adding integrations

Add data to your cluster from any source, then analyze and



Search Elastic



Home

Welcome home



Enterprise Search

Create search experiences with a refined set of APIs and tools.



Observability

Consolidate your logs, metrics, application traces, and system availability with purpose-built UIs.



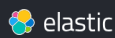
Security

Prevent, collect, detect, and respond to threats for unified protection across your infrastructure.



Analytics

Explore, visualize, and analyze your data using a powerful suite of analytical tools and applications.



Search Elastic



Stack Management



Management

Ingest

Ingest Pipelines

Data

Index Management

Index Lifecycle Policies

Snapshot and Restore

Rollup Jobs

Transforms

Remote Clusters

Alerts and Insights

Rules and Connectors



Welcome to Stack Management 7.17.9

Manage your indices, index patterns, saved objects, Kibana settings, and more.

A complete list of apps is in the menu on the left.

Management

Ingest [ⓘ]
Ingest Pipelines

Data [ⓘ]
[Index Management](#)
Index Lifecycle Policies
Snapshot and Restore
Rollup Jobs
Transforms
Remote Clusters

Alerts and Insights [ⓘ]
Rules and Connectors

Index Management

[Indices](#)[Data Streams](#)[Index Templates](#)[Component Templates](#)

Update your Elasticsearch indices individually or in bulk. [Learn more.](#) [↗](#)

☐ Include rollup indices

☐ Include hidden indices

☐ Name

Health

Status

Primaries

Replicas

Docs count

Storage size

Data stream

☐ nifi

● yellow

open

1

1

75036

51.5mb

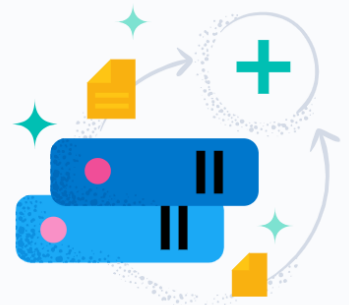
Rows per page: 10 [↕](#)

[Reload indices](#)

< 1 >

You have data in Elasticsearch. Now, create an index pattern.

Kibana requires an index pattern to identify which data streams, indices, and index aliases you want to explore. An index pattern can point to a specific index, for example, your log data from yesterday, or all indices that contain your log data.

[+ Create index pattern](#)

Want to learn more? [Read documentation](#) [↗](#)

Create index pattern

Name

Use an asterisk (*) to match multiple characters. Spaces and the characters `/`, `?`, `"`, `<`, `>`, `|` are not allowed.

Timestamp field

No matching data stream, index, or index alias has a timestamp field.

[Show advanced settings](#)

[× Close](#)[Create index pattern](#)

✓ Your index pattern matches 1 source.

nifi

Index

Rows per page: 10 [↕](#)

