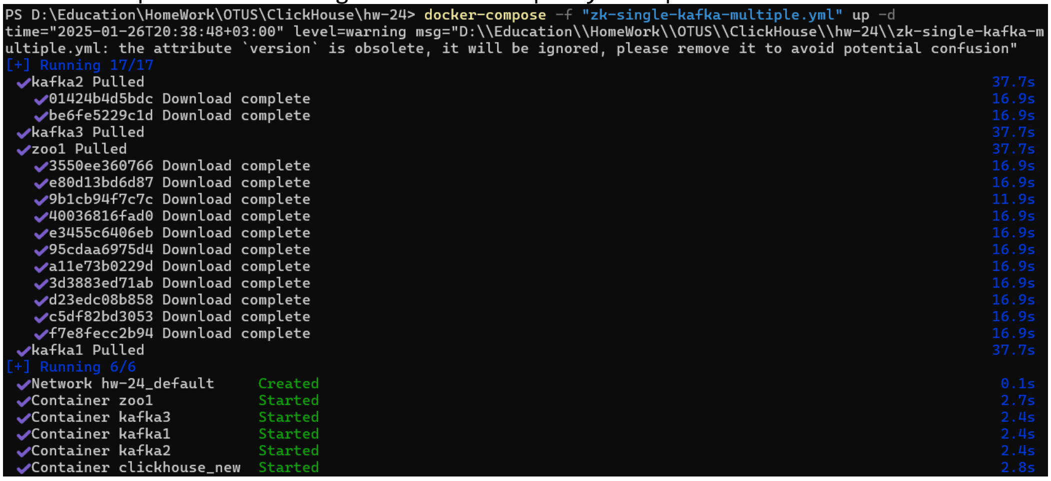
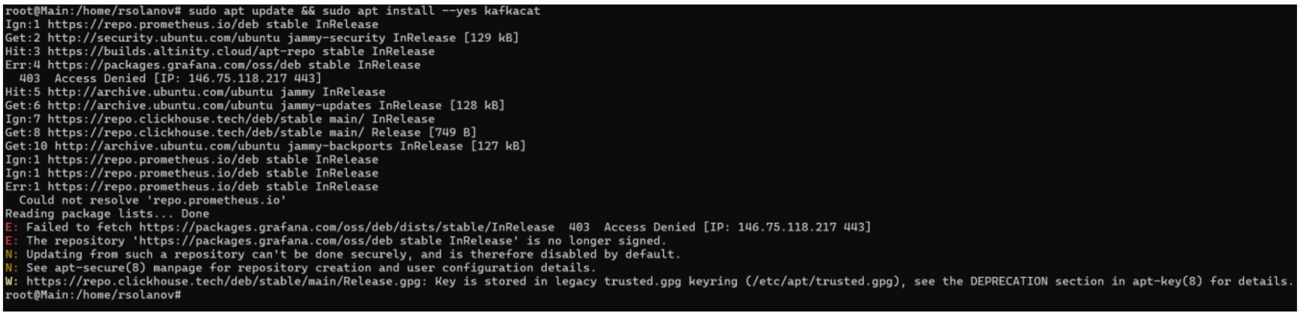
- Установка Kafka с помощью Docker
  В docker-compose файл добавлен ClickHouse
  `docker-compose -f "zk-single-kafka-multiple.yml" up -d`



```
PS D:\Education\HomeWork\OTUS\ClickHouse\hw-24> docker-compose -f "zk-single-kafka-multiple.yml" up -d
time="2025-01-26T20:38:48+03:00" level=warning msg="D:\\Education\\HomeWork\\OTUS\\ClickHouse\\hw-24\\zk-single-kafka-m
ultiple.yml: the attribute `version` is obsolete, it will be ignored, please remove it to avoid potential confusion"
[+] Running 17/17
 ✔kafka2 Pulled                                                                                           37.7s
   ✔01424b4d5bdc Download complete                                                                        16.9s
   ✔be6fe5229c1d Download complete                                                                        16.9s
 ✔kafka3 Pulled                                                                                           37.7s
 ✔zoo1 Pulled                                                                                             37.7s
   ✔3550ee360766 Download complete                                                                        16.9s
   ✔e80d13bd6d87 Download complete                                                                        16.9s
   ✔9b1cb94f7c7c Download complete                                                                        11.9s
   ✔40036816fad0 Download complete                                                                        16.9s
   ✔e3455c6406eb Download complete                                                                        16.9s
   ✔95cdaa6975d4 Download complete                                                                        16.9s
   ✔a11e73b0229d Download complete                                                                        16.9s
   ✔3d3883ed71ab Download complete                                                                        16.9s
   ✔d23edc08b858 Download complete                                                                        16.9s
   ✔c5df82bd3053 Download complete                                                                        16.9s
   ✔f7e8fecc2b94 Download complete                                                                        16.9s
 ✔kafka1 Pulled                                                                                           37.7s
[+] Running 6/6
 ✔Network hw-24_default     Created                                                                        0.1s
 ✔Container zoo1            Started                                                                        2.7s
 ✔Container kafka3          Started                                                                        2.4s
 ✔Container kafka1          Started                                                                        2.4s
 ✔Container kafka2          Started                                                                        2.4s
 ✔Container clickhouse_new  Started                                                                        2.8s
```

- Установка kafkacat

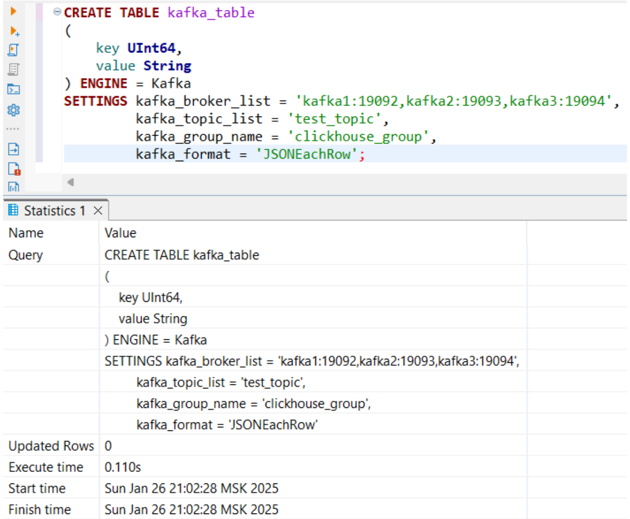
`sudo apt update && sudo apt install --yes kafkacat`



```
root@Main:/home/rsolanov# sudo apt update && sudo apt install --yes kafkacat
Ign:1 https://repo.prometheus.io/deb stable InRelease
Get:2 http://security.ubuntu.com/ubuntu jammy-security InRelease [129 kB]
Hit:3 https://builds.altinity.cloud/apt-repo stable InRelease
Err:4 https://packages.grafana.com/oss/deb stable InRelease
  403  Access Denied [IP: 146.75.118.217 443]
Hit:5 http://archive.ubuntu.com/ubuntu jammy InRelease
Get:6 http://archive.ubuntu.com/ubuntu jammy-updates InRelease [128 kB]
Ign:7 https://repo.clickhouse.tech/deb/stable main/ InRelease
Get:8 https://repo.clickhouse.tech/deb/stable main/ Release [749 B]
Get:10 http://archive.ubuntu.com/ubuntu jammy-backports InRelease [127 kB]
Ign:1 https://repo.prometheus.io/deb stable InRelease
Ign:1 https://repo.prometheus.io/deb stable InRelease
Err:1 https://repo.prometheus.io/deb stable InRelease
  Could not resolve 'repo.prometheus.io'
Reading package lists... Done
E: Failed to fetch https://packages.grafana.com/oss/deb/dists/stable/InRelease  403  Access Denied [IP: 146.75.118.217 443]
E: The repository 'https://packages.grafana.com/oss/deb stable InRelease' is no longer signed.
N: Updating from such a repository can't be done securely, and is therefore disabled by default.
N: See apt-secure(8) manpage for repository creation and user configuration details.
W: https://repo.clickhouse.tech/deb/stable/main/Release.gpg: Key is stored in legacy trusted.gpg keyring (/etc/apt/trusted.gpg), see the DEPRECATION section in apt-key(8) for details.
root@Main:/home/rsolanov#
```

- Создание таблицы с движком Kafka

```
CREATE TABLE kafka_table
(
    key UInt64,
    value String
) ENGINE = Kafka
SETTINGS kafka_broker_list = 'kafka1:19092,kafka2:19093,kafka3:19094',
         kafka_topic_list = 'test_topic',
         kafka_group_name = 'clickhouse_group',
         kafka_format = 'JSONEachRow';
```



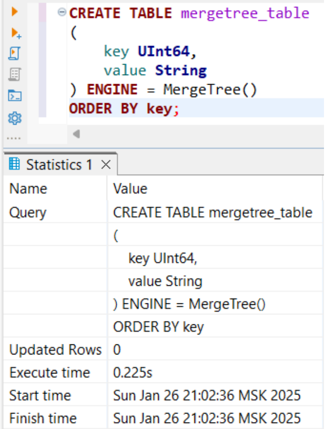
```
CREATE TABLE kafka_table
(
    key UInt64,
    value String
) ENGINE = Kafka
SETTINGS kafka_broker_list = 'kafka1:19092,kafka2:19093,kafka3:19094',
         kafka_topic_list = 'test_topic',
         kafka_group_name = 'clickhouse_group',
         kafka_format = 'JSONEachRow';
```

Name	Value
Query	CREATE TABLE kafka_table (key UInt64, value String) ENGINE = Kafka SETTINGS kafka_broker_list = 'kafka1:19092,kafka2:19093,kafka3:19094', kafka_topic_list = 'test_topic', kafka_group_name = 'clickhouse_group', kafka_format = 'JSONEachRow'
Updated Rows	0
Execute time	0.110s
Start time	Sun Jan 26 21:02:28 MSK 2025
Finish time	Sun Jan 26 21:02:28 MSK 2025

- Создание таблицы с движком MergeTree

```
CREATE TABLE mergetree_table
```

```
(
    key UInt64,
    value String
) ENGINE = MergeTree()
ORDER BY key;
```



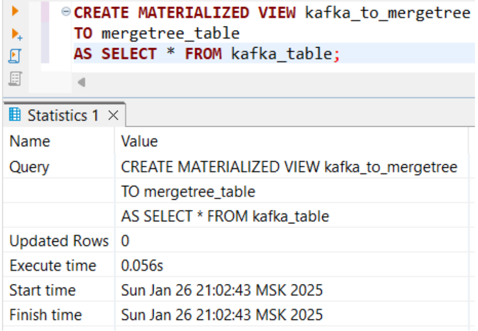
```
CREATE TABLE mergetree_table
(
    key UInt64,
    value String
) ENGINE = MergeTree()
ORDER BY key;
```

Statistics 1

Name	Value
Query	CREATE TABLE mergetree_table (key UInt64, value String) ENGINE = MergeTree() ORDER BY key
Updated Rows	0
Execute time	0.225s
Start time	Sun Jan 26 21:02:36 MSK 2025
Finish time	Sun Jan 26 21:02:36 MSK 2025

- Создание материализованного представления

```
CREATE MATERIALIZED VIEW kafka_to_mergetree
TO mergetree_table
AS SELECT * FROM kafka_table;
```



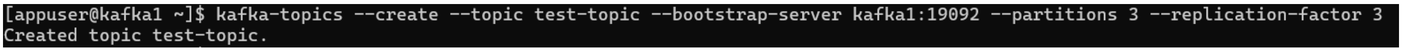
```
CREATE MATERIALIZED VIEW kafka_to_mergetree
TO mergetree_table
AS SELECT * FROM kafka_table;
```

Statistics 1

Name	Value
Query	CREATE MATERIALIZED VIEW kafka_to_mergetree TO mergetree_table AS SELECT * FROM kafka_table
Updated Rows	0
Execute time	0.056s
Start time	Sun Jan 26 21:02:43 MSK 2025
Finish time	Sun Jan 26 21:02:43 MSK 2025

- Создание топика Kafka

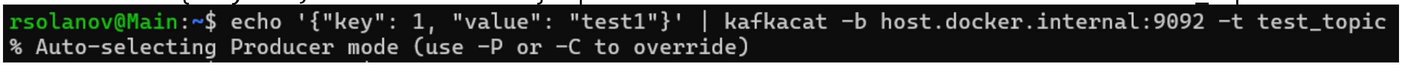
```
docker exec -it kafka1 /bin/bash
cd /usr/bin
kafka-topics --create --topic test-topic --bootstrap-server kafka1:19092 --partitions 3 --
replication-factor 3
```



```
[appuser@kafka1 ~]$ kafka-topics --create --topic test-topic --bootstrap-server kafka1:19092 --partitions 3 --replication-factor 3
Created topic test-topic.
```

- Отправка сообщений в Kafka

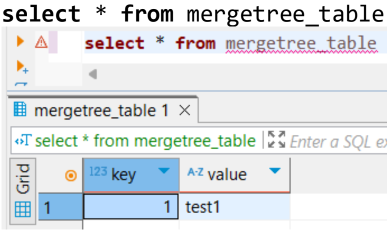
```
echo '{"key": 1, "value": "test1"}' | kafkacat -b host.docker.internal:9092 -t test_topic
```



```
rsolanov@Main:~$ echo '{"key": 1, "value": "test1"}' | kafkacat -b host.docker.internal:9092 -t test_topic
% Auto-selecting Producer mode (use -P or -C to override)
```

- Проверка получения сообщений от kafka и запись в ClickHouse

```
select * from mergetree_table
```



```
select * from mergetree_table
```

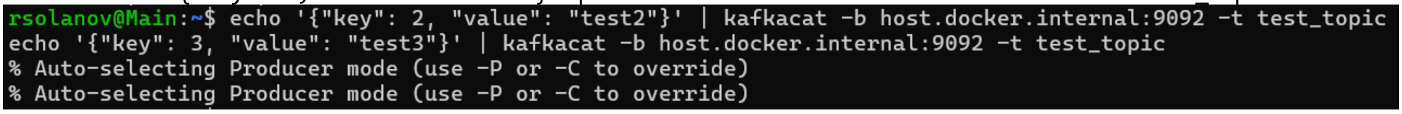
mergetree_table 1

select * from mergetree_table

	key	value
1	1	test1

- Отправка в Kafka новых сообщений

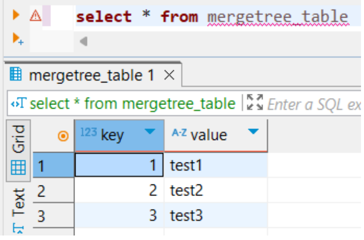
```
echo '{"key": 2, "value": "test2"}' | kafkacat -b host.docker.internal:9092 -t test_topic
echo '{"key": 3, "value": "test3"}' | kafkacat -b host.docker.internal:9092 -t test_topic
```



```
rsolanov@Main:~$ echo '{"key": 2, "value": "test2"}' | kafkacat -b host.docker.internal:9092 -t test_topic
echo '{"key": 3, "value": "test3"}' | kafkacat -b host.docker.internal:9092 -t test_topic
% Auto-selecting Producer mode (use -P or -C to override)
% Auto-selecting Producer mode (use -P or -C to override)
```

- Проверка получения сообщений в ClickHouse

```
select * from mergetree_table
```



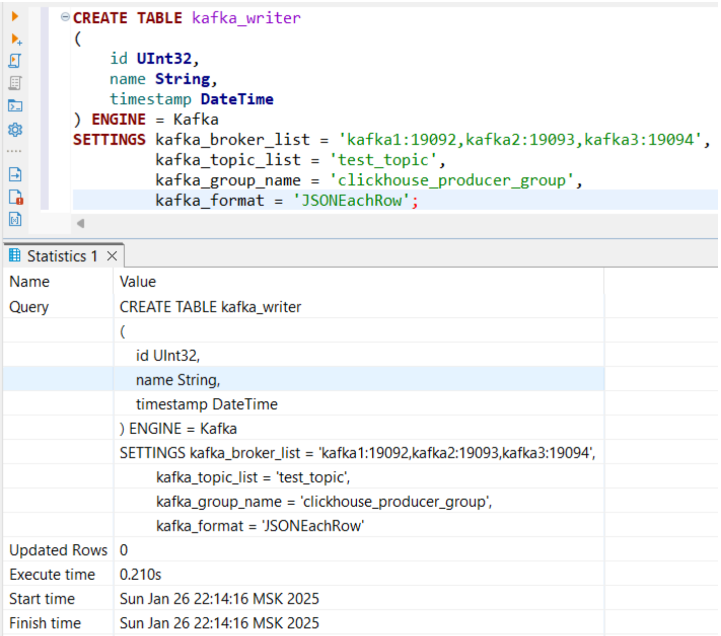
```
select * from mergetree_table
```

	key	value
1	1	test1
2	2	test2
3	3	test3

Записать данные в кафку с помощью ClickHouse Kafka Engine

- Создание таблицы для записи в Kafka

```
CREATE TABLE kafka_writer
(
    id UInt32,
    name String,
    timestamp DateTime
) ENGINE = Kafka
SETTINGS kafka_broker_list = 'kafka1:19092,kafka2:19093,kafka3:19094',
         kafka_topic_list = 'test_topic',
         kafka_group_name = 'clickhouse_producer_group',
         kafka_format = 'JSONEachRow';
```



```
CREATE TABLE kafka_writer
(
    id UInt32,
    name String,
    timestamp DateTime
) ENGINE = Kafka
SETTINGS kafka_broker_list = 'kafka1:19092,kafka2:19093,kafka3:19094',
         kafka_topic_list = 'test_topic',
         kafka_group_name = 'clickhouse_producer_group',
         kafka_format = 'JSONEachRow';
```

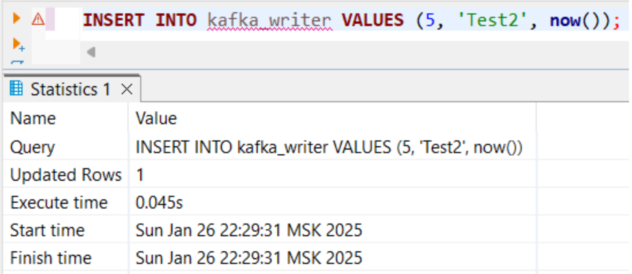
Name	Value
Query	CREATE TABLE kafka_writer (id UInt32, name String, timestamp DateTime) ENGINE = Kafka SETTINGS kafka_broker_list = 'kafka1:19092,kafka2:19093,kafka3:19094', kafka_topic_list = 'test_topic', kafka_group_name = 'clickhouse_producer_group', kafka_format = 'JSONEachRow'
Updated Rows	0
Execute time	0.210s
Start time	Sun Jan 26 22:14:16 MSK 2025
Finish time	Sun Jan 26 22:14:16 MSK 2025

- Пересоздание топика

```
kafka-topics --bootstrap-server kafka1:19092 --delete --topic test_topic
kafka-topics --bootstrap-server kafka1:19092 --create --topic test_topic --partitions 3 --
replication-factor 3
```

- Вставка данных в таблицу

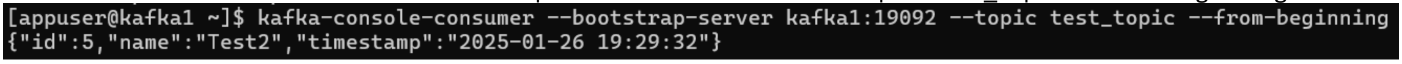
```
INSERT INTO kafka_writer VALUES (5, 'Test2', now());
```



```
INSERT INTO kafka_writer VALUES (5, 'Test2', now());
```

Name	Value
Query	INSERT INTO kafka_writer VALUES (5, 'Test2', now())
Updated Rows	1
Execute time	0.045s
Start time	Sun Jan 26 22:29:31 MSK 2025
Finish time	Sun Jan 26 22:29:31 MSK 2025

- Чтение данных из топика
  kafka-console-consumer --bootstrap-server kafka1:19092 --topic test_topic --from-beginning



```
[appuser@kafka1 ~]$ kafka-console-consumer --bootstrap-server kafka1:19092 --topic test_topic --from-beginning
{"id":5,"name":"Test2","timestamp":"2025-01-26 19:29:32"}
```

Построить тот же пайплайн обработки данных НЕ через Kafka Engine

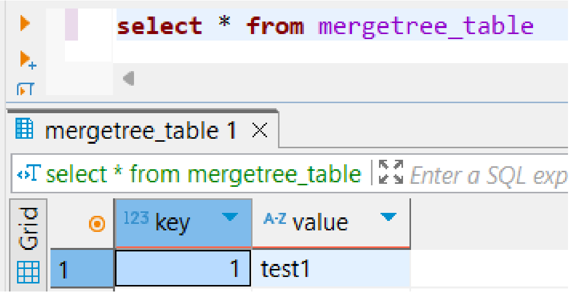
- Пересоздаем топик, очищаем таблицу «mergetree_table» и добавляем сообщение
  echo '{"key": 1, "value": "test1"}' | kafkacat -b host.docker.internal:9092 -t test_topic

- Запускаем скрипт Python

```python
from confluent_kafka import Consumer
import clickhouse_connect
import json


def main():
    consumer = Consumer({
        'bootstrap.servers': 'host.docker.internal:9092',
        'group.id': 'python_consumer_group',
        'auto.offset.reset': 'earliest'
    })
    consumer.subscribe(['test_topic'])

    clickhouse_client = clickhouse_connect.get_client(host='host.docker.internal', port=8123, username='default', password='default')

    while True:
        msg = consumer.poll(1.0)
        if msg is None:
            continue
        if msg.error():
            print(f"Ошибка потребителя: {msg.error()}")
            continue

        data = json.loads(msg.value().decode('utf-8'))
        print(json.dumps(data, indent=2))
        clickhouse_client.command(f"INSERT INTO mergetree_table FORMAT JSONEachRow {json.dumps(data)}")

    consumer.close()

if __name__ == '__main__':
    main()
```

- Проверка данных таблицы

  select * **from** mergetree_table



```sql
select * from mergetree_table
```

	key	value
1	1	test1