

CDC streaming pipeline from PostgreSQL ? Kafka ? Flink ? ClickHouse.

- [CDC streaming pipeline from PostgreSQL → Kafka → Flink → ClickHouse.](#)
- [Jar create](#)

CDC streaming pipeline from PostgreSQL ? Kafka ? Flink ? ClickHouse.

```
sudo wget https://repo1.maven.org/maven2/com/clickhouse/flink/flink-connector-clickhouse-2.0.0/0.1.1/flink-connector-clickhouse-2.0.0-0.1.1-all.jar
```

```
wget https://repo1.maven.org/maven2/org/apache/flink/flink-connector-kafka/3.1.0-1.18/flink-connector-kafka-3.1.0-1.18.jar
```

```
sudo /opt/flink/bin/flink run flink-clickhouse-job-1.0-SNAPSHOT.jar
```

```
/opt/flink/bin/sql-client.sh
```

```
ADD JAR 'file:///opt/flink/lib/flink-connector-clickhouse-2.0.0-0.1.1-all.jar';
```

```
ADD JAR 'file:///opt/flink/lib/flink-connector-kafka-3.1.0-1.18.jar';
```

```
ADD JAR 'file:///opt/flink/lib/clickhouse-jdbc-0.6.0.jar';
```

```
ADD JAR 'file:///opt/flink/lib/flink-connector-jdbc-3.1.2-1.18.jar';
```

? Kafka Source

```
CREATE TABLE kafka_users (  
  id BIGINT,  
  name STRING,  
  email STRING,  
  created_at TIMESTAMP(3),  
  WATERMARK FOR created_at AS created_at - INTERVAL '5' SECOND  
) WITH (
```

```
'connector'='kafka',  
'topic'='pg.public.users',  
'properties.bootstrap.servers'='192.168.1.226:9092',  
'properties.group.id'='flink-users-group',  
'scan.startup.mode'='earliest-offset',  
'format'='json',  
'json.fail-on-missing-field'='false',  
'json.ignore-parse-errors'='true'  
);
```

? ClickHouse Sink via JDBC

```
CREATE TABLE clickhouse_users (  
  id BIGINT,  
  name STRING,  
  email STRING,  
  created_at TIMESTAMP(3)  
) WITH (  
  'connector' = 'jdbc',  
  'url' = 'jdbc:clickhouse://192.168.1.228:8123/userdb',  
  'table-name' = 'users',  
  'username' = 'default',  
  'password' = ''  
);
```

? Streaming insert

```
INSERT INTO clickhouse_users  
SELECT id, name, email, created_at  
FROM kafka_users;
```

Jar create

`mvn clean package -DskipTests`

`mvn clean install`