

FIREBIRD 🐦 CONF 2025

Игорь Вальченко

Использование Change Data Capture в HQbird и РЕД База Данных

Содержание

- Что такое Change Data Capture?
- Посмотрим в недалекое прошлое
- Что есть на текущий момент?
- Как устроено в HQbird?
- Как устроено в РЕД База Данных?
- Топология использования CDC

What to take Change Data Capture?

Что такое Change Data Capture?

1. Отслеживание изменений в базе данных через лог транзакций
2. Отслеживание изменений в базе данных с помощью триггеров
3. Отслеживание изменений по временным меткам

Посмотрим в недалекое прошлое

Логическая асинхронная репликация с помощью триггеров

- IBReplicator
- FBReplicator
- CopyCat LiveMirror
- DBRE
- Самописные инструменты репликации

Логическая асинхронная репликация с помощью триггеров

- Отказоустойчивые решения
- ETL — Extract, Transform, Load

Логическая асинхронная репликация с помощью триггеров

Минусы:

- Дополнительные триггеры
- Дополнительные таблицы
- Дополнительные манипуляции
- Ручная репликация DDL
- Небольшая просадка производительности

Новые возможности Firebird

Нативная репликация

РЕД База Данных 2.5

HQBird 2.5-3.0

Firebird 4

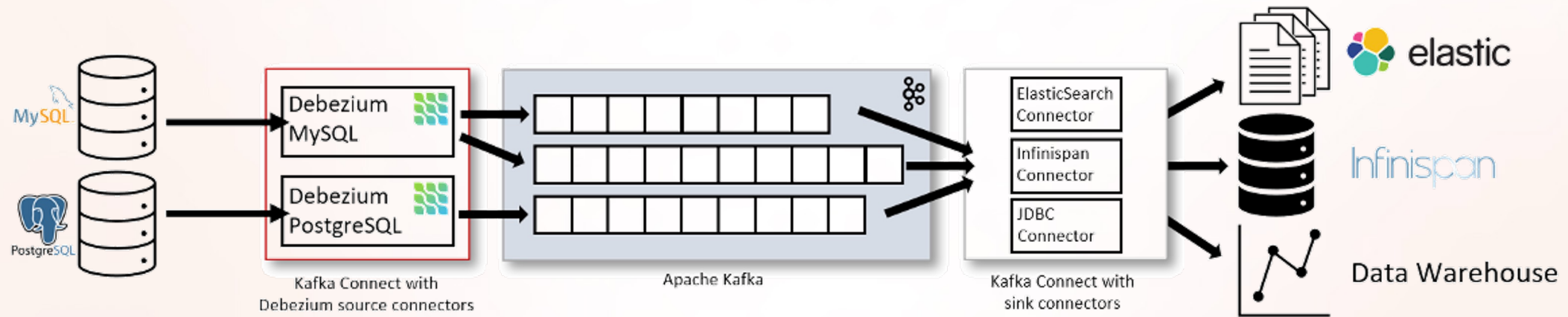
Что есть на текущий момент?

Использование логов транзакций

- PostgreSQL — WAL (write-ahead log)
- MySQL - binlog

Другие варианты

- Kafka Connector
- Debezium
- Apache Kafka



Debezium

Преимущества:

- Open-source продукт
- Поддержка Red Hat
- Широкое применение

Отечественные форки Debezium

- SIDEC (SberInfra Data Base Connect)
- ВТБ-Дебезиум

Для Firebird

- HQBird - сервис fb_streaming
- РЕД База Данных - плагин репликации данных в Apache Kafka

Почему используем Apache Kafka?

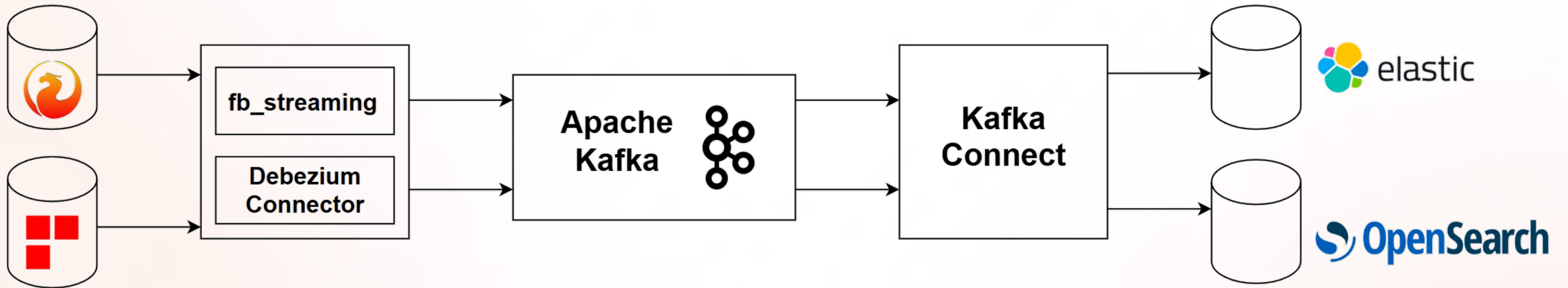
- Простота
- Надежность
- Горизонтальная масштабируемость
- Kafka Connect

Почему лог репликации лучше триггерной репликации?

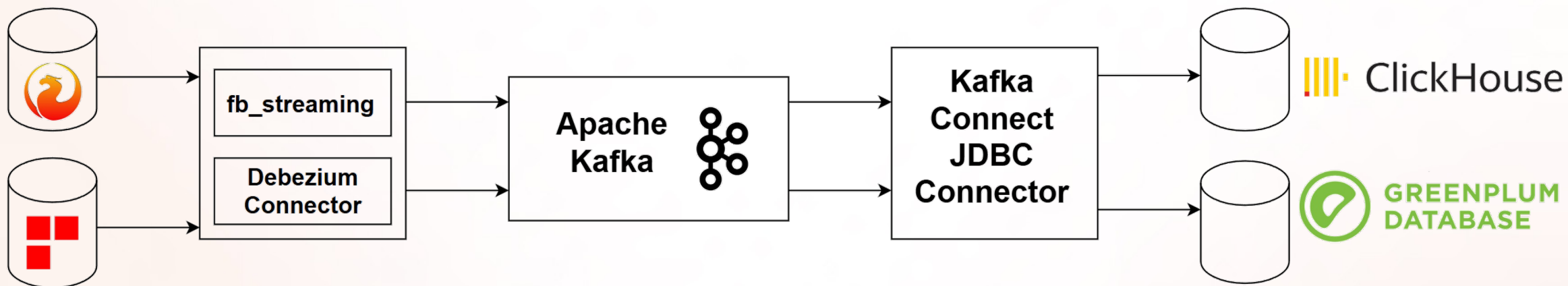
Преимущества:

- Не требуются триггеры для фиксации изменений данных
- Не требуется таблица для хранения изменений
- Нагрузка на СУБД меньше

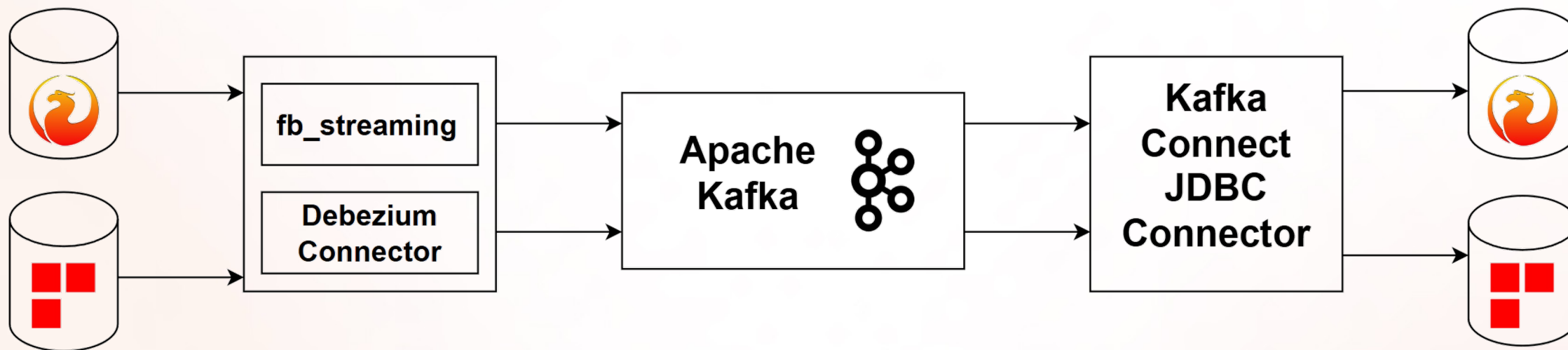
Firebird/РБД → Full Text Search



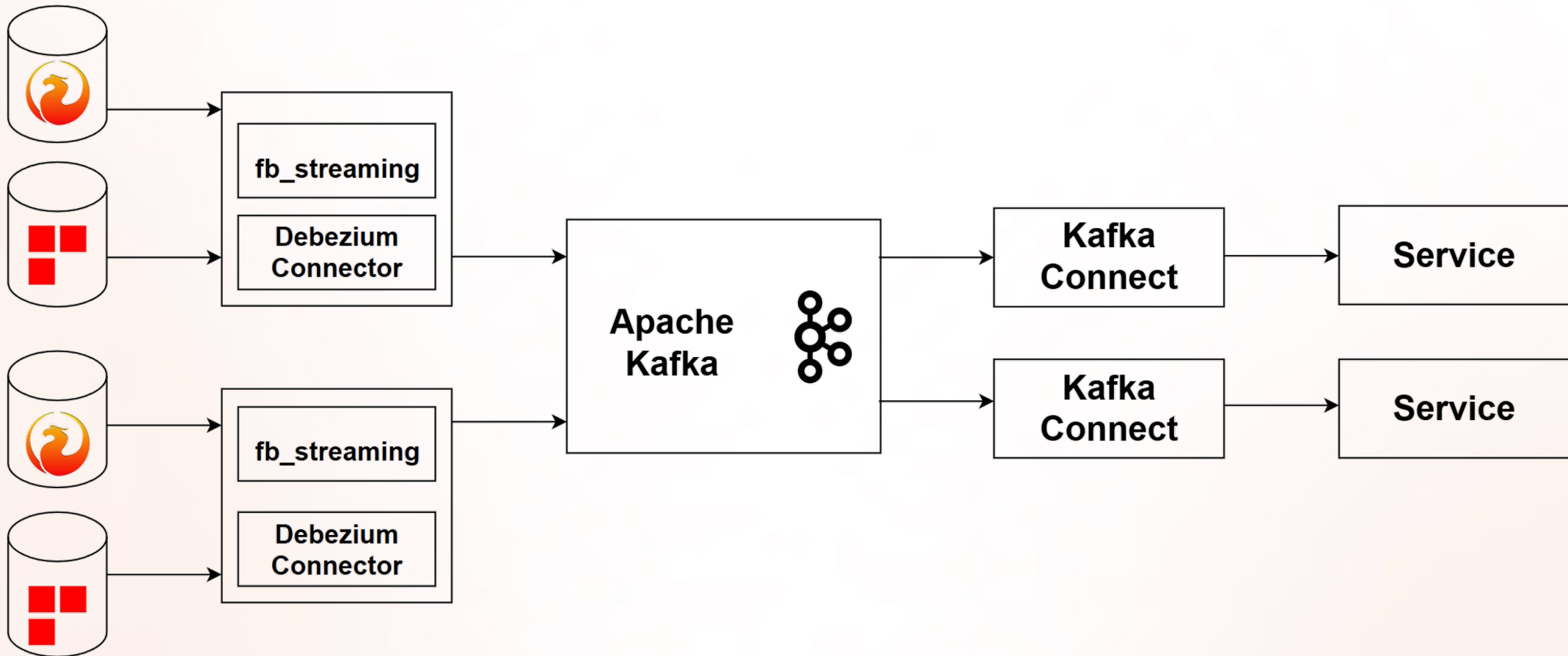
Firebird/РБД -> Data Warehouse



Реализация отказоустойчивости



Firebird/РБД -> микросервисы



HQBird

HQbird - fb_streaming

- Отдельная служба fb_streaming
- Плагин kafka_cdc_plugin

fb_streaming — replication.conf

```
database = c:\fbdata\5.0\test.fdb
{
  journal_directory = d:\replication\journal
  journal_archive_directory = d:\replication\kafka_source
  journal_archive_timeout = 10
}
```

Репликация таблиц

Добавляем таблицы в публикацию

```
ALTER DATABASE INCLUDE TABLE TABLE_1, TABLE_2,  
TABLE_3 TO PUBLICATION
```

```
ALTER DATABASE INCLUDE ALL TO PUBLICATION
```

HQbird - публикация

Состояние публикаций:

RDB\$PUBLICATIONS

RDB\$PUBLICATION_TABLES

HQbird - fb_streaming.conf

```
task = d:\fbdata\5.0\replication\kafka_source
{
    database = inet://localhost:3055/test.fdb
    username = SYSDBA
    password = masterkey
```

HQbird — fb_streaming.conf

deleteProcessedFile = true

plugin = kafka_cdc_plugin

dumpBlobs = true

kafka_brokers = kafka_server:9092

kafka_topic_prefix = fb_streaming

kafka_topic_cdc_event = \${topicPrefix}.cdc

HQbird — fb_streaming.conf

kafka_topic_transaction = \${topicPrefix}.transaction

key_cdc_events_enable = true

key_converter_schemas_enable = true

value_converter_schemas_enable = true

transaction_metadata_enable = true

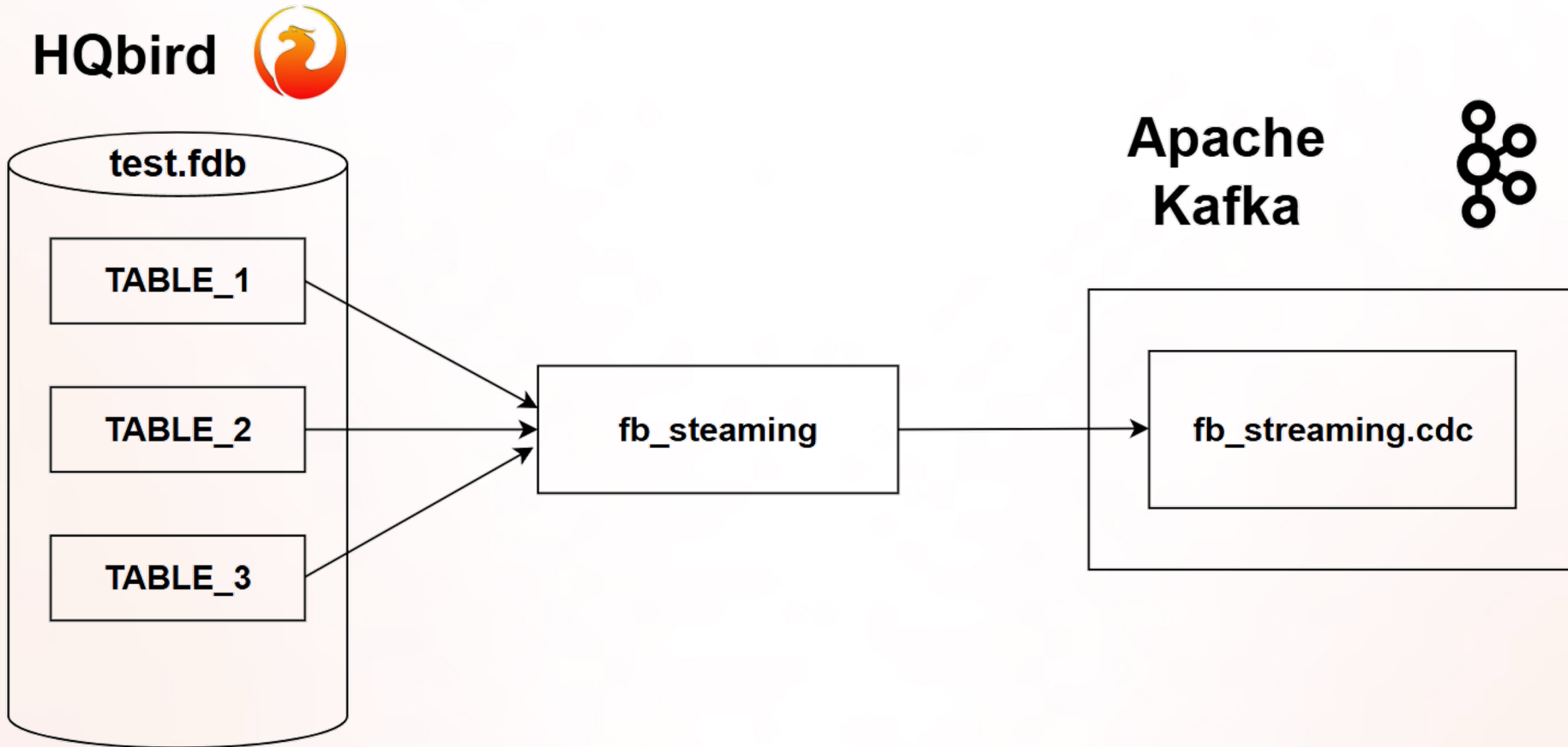
Hqbird — старт репликации

Установка и старт службы fb_streaming

Разрешаем публикацию

```
ALTER DATABASE ENABLE PUBLICATION;
```

Hqbird



Включение репликации для кластера HQbird

для создания отказоустойчивого кластера

replication.conf

```
journal_archive_command = "copy $(pathname) $(archivepathname)  
&& copy $(pathname)
```

```
d:\fbdata\5.0\replication\kafka_source"
```

Включение репликации для кластера HQbird

для создания отказоустойчивого кластера

fb_streaming.conf

include_tables

exclude_tables

Документация HQbird

Больше информации о настройке в HQbird User Guide

<https://www.ibase.ru/files/products/hqbird-2024-userguide-ru.pdf>



РЕД База Данных

РБД — plugin.conf

```
Plugin = Debezium_Connector {  
    Module = $(dir_plugins)/debezium_connector  
    Config = Debezium_Connector_config  
}  
Config = Debezium_Connector_config {  
    JarDirs = $(this)/debezium  
}
```

РБД — replication.conf

```
database = c:\DBASE\DEB_RDB.FDB  
{  
  plugin = Debezium_Connector  
}
```

reddatabase.connector.yaml

```
connector.name: reddatabase
connector.journal.directory: c:\\reddatabase-debezium-connector
connector.journal.check.interval.ms: 1000
topic.prefix: rdb_topic
kafka.producer.bootstrap.servers: kafka_server:9092
kafka.producer.retries: 3
kafka.producer.retry.backoff.ms: 1000
key.converter: org.apache.kafka.connect.json.JsonConverter
value.converter: org.apache.kafka.connect.json.JsonConverter
```

РБД - публикация

Добавляем таблицы в публикацию

```
ALTER DATABASE INCLUDE TABLE TABLE_1, TABLE_2, TABLE_3 TO PUBLICATION
```

```
ALTER DATABASE INCLUDE ALL TO PUBLICATION;
```

Разрешаем публикацию

```
ALTER DATABASE ENABLE PUBLICATION;
```

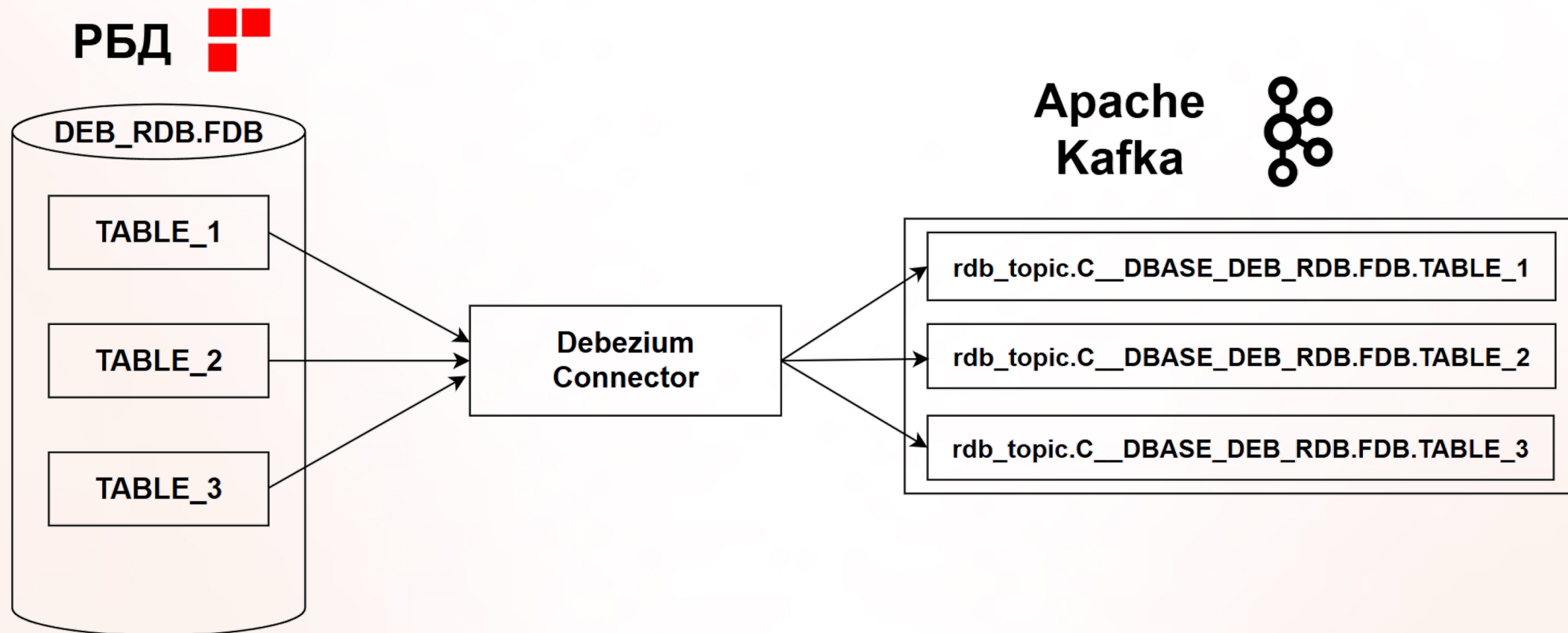

РБД - публикация

Состояние публикаций:

RDB\$PUBLICATIONS

RDB\$PUBLICATION_TABLES

РБД



Отказоустойчивый кластер РЕД Базы Данных

replication.conf

database = c:\DBASE\DEB_RDB.FDB

{

journal_directory = c:\reddatabase\journal

journal_archive_directory = c:\reddatabase\archive

journal_archive_timeout = 10

}

Отказоустойчивый кластер РЕД Базы Данных

reddatabase.connector.yaml

table.include.list

table.exclude.list

Документация РБД

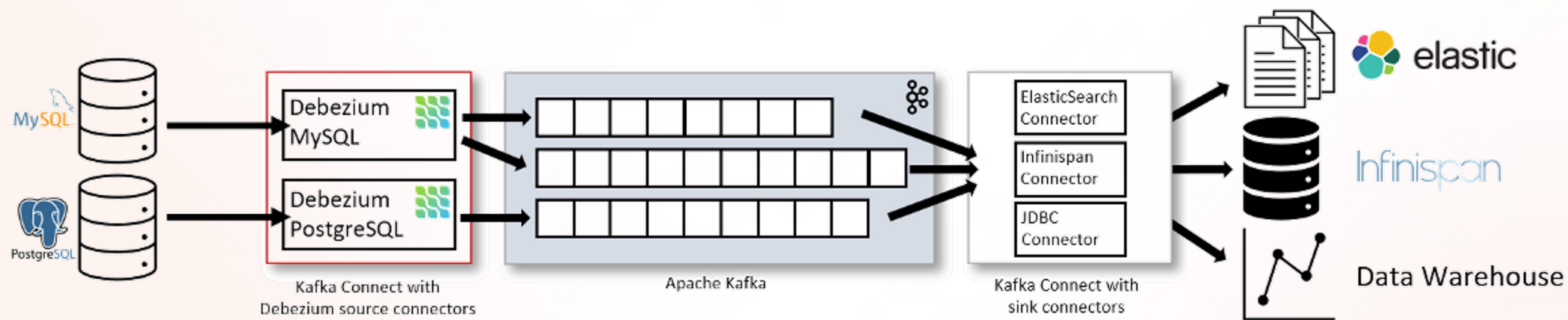
Больше информации о настройке в руководстве администратора
РЕД База Данных

<https://reddatabase.ru/ru/documentation/>



Топология использования CDC

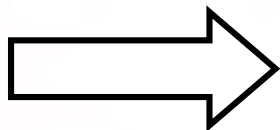
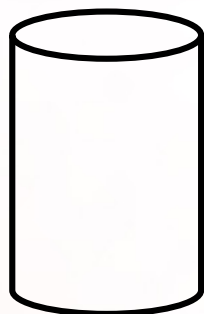
Топология использования CDC



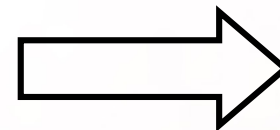
Топология использования CDC

Консолидация данных

Firebird
РЕД База Данных



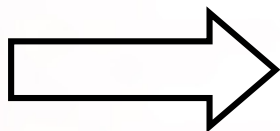
Apache Kafka



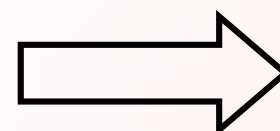
ClickHouse
GreenPlum
ELK
Любые СУБД



Firebird
РЕД База Данных

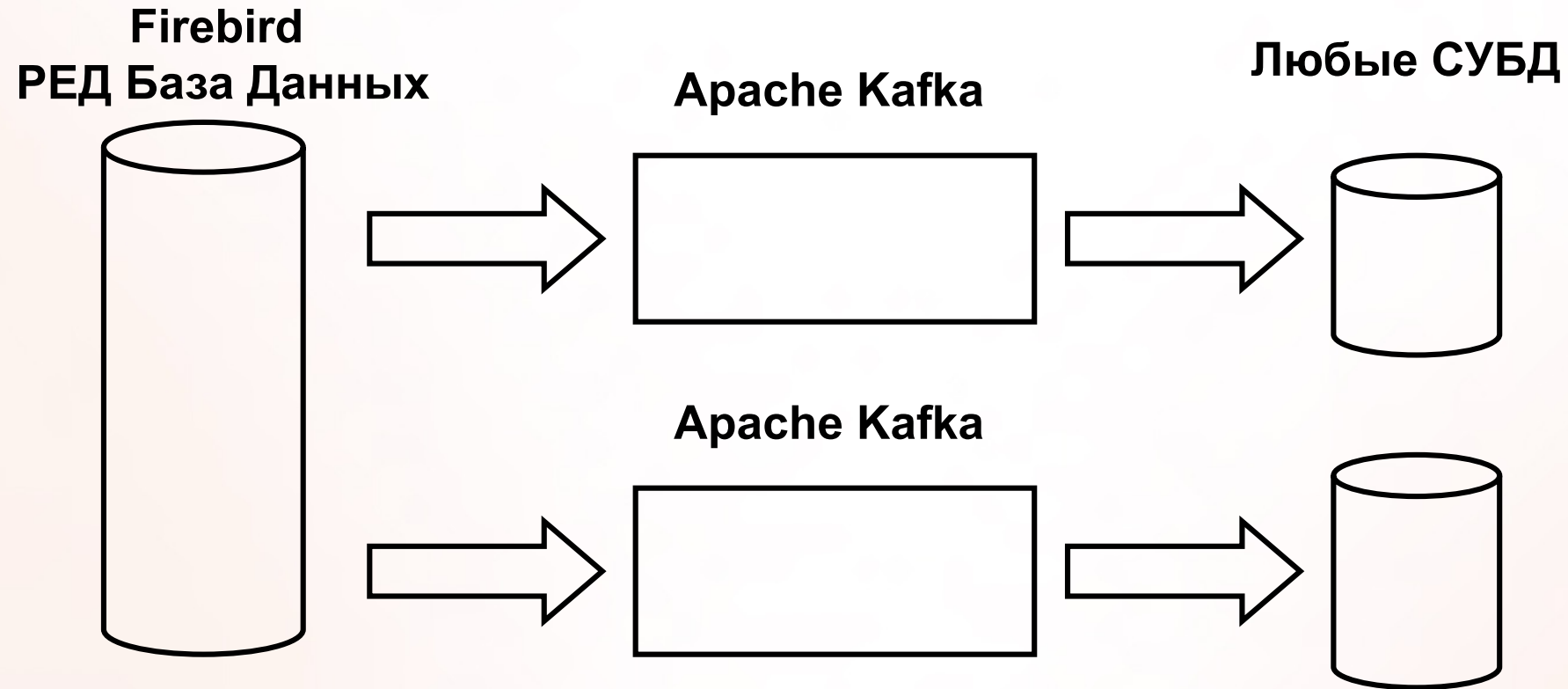


Apache Kafka



Топология использования CDC

Распространение данных



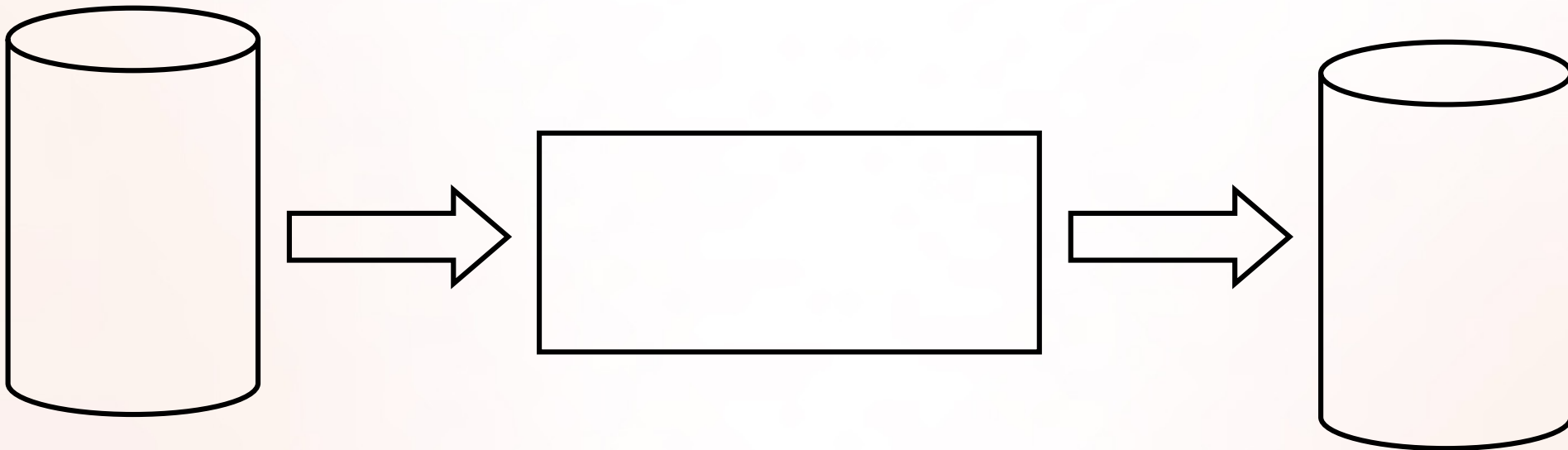
Топология использования CDC

Однонаправленная схема

Firebird
РЕД База Данных

Apache Kafka

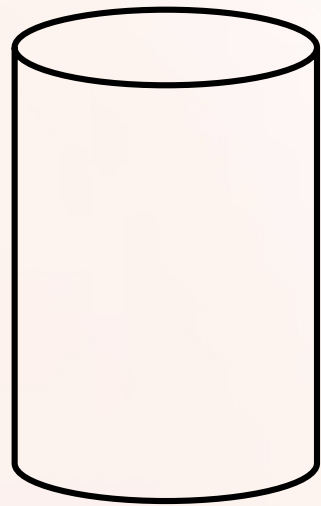
Firebird
РЕД База Данных



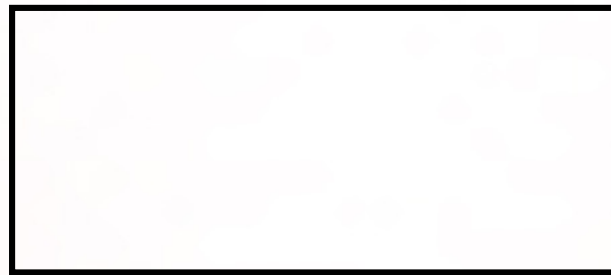
Топология использования CDC

Двунаправленная схема

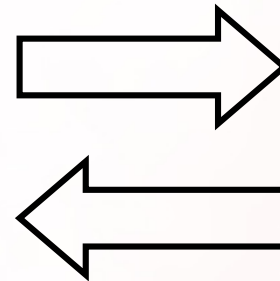
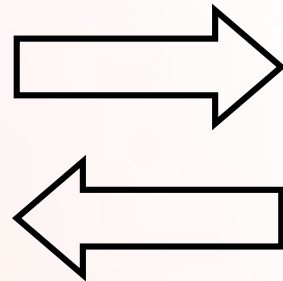
Firebird
РЕД База Данных



Apache Kafka



Firebird
РЕД База Данных



Вопросы?

Теплый Ламповый SQL

YouTube



RuTube



Boosty



29 МАЯ | МОСКВА

FIREBIRD CONF 2025