**Отчет по выполнению практической работы: Развертывание БД и настройка Airbyte для обмена данными**

**1. Цель работы**

Развертывание СУБД PostgreSQL и ClickHouse, настройка инструмента Airbyte для обмена данными между ними.

**2. Выполненные действия**

**2.1. Развертывание PostgreSQL**

1. Установил PostgreSQL на локальный компьютер:
2. Подключился к БД через DBeaver
3. Создал тестовую базу данных и таблицами:

``` sql

CREATE DATABASE test\_airbyte;

CREATE TABLE public.department (

dept\_no text NOT NULL,

dept\_name text NOT NULL,

CONSTRAINT department\_dept\_name\_key UNIQUE (dept\_name),

CONSTRAINT department\_pkey PRIMARY KEY (dept\_no)

);

CREATE TABLE public.employee (

emp\_no serial4 NOT NULL,

birth\_date date NOT NULL,

first\_name text NOT NULL,

last\_name text NOT NULL,

gender text NOT NULL,

hire\_date date NOT NULL,

CONSTRAINT employee\_gender\_check CHECK ((gender = ANY (ARRAY['M'::text, 'F'::text]))),

CONSTRAINT employee\_pkey PRIMARY KEY (emp\_no)

);

CREATE INDEX idx\_employee\_hire\_date ON public.employee USING btree (hire\_date);

CREATE TABLE public.dept\_emp (

emp\_no int4 NOT NULL,

dept\_no text NOT NULL,

from\_date date NOT NULL,

to\_date date NOT NULL,

CONSTRAINT dept\_emp\_pkey PRIMARY KEY (emp\_no, dept\_no),

CONSTRAINT dept\_emp\_dept\_no\_fkey FOREIGN KEY (dept\_no) REFERENCES public.department(dept\_no) ON DELETE CASCADE,

CONSTRAINT dept\_emp\_emp\_no\_fkey FOREIGN KEY (emp\_no) REFERENCES public.employee(emp\_no) ON DELETE CASCADE

);

CREATE TABLE public.title (

emp\_no int4 NOT NULL,

title text NOT NULL,

from\_date date NOT NULL,

to\_date date NULL,

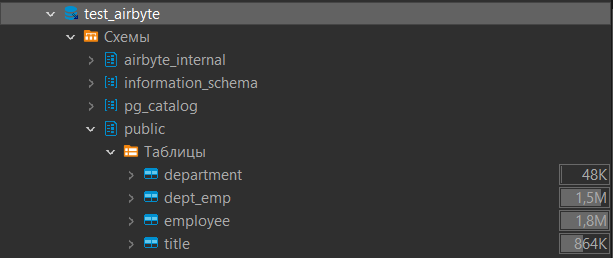
CONSTRAINT title\_pkey PRIMARY KEY (emp\_no, title, from\_date),

CONSTRAINT title\_emp\_no\_fkey FOREIGN KEY (emp\_no) REFERENCES public.employee(emp\_no) ON DELETE CASCADE

);

```

1. Заполнил таблицы тестовыми данными
2. Проверил наличие таблиц



**2.2. Развертывание ClickHouse**

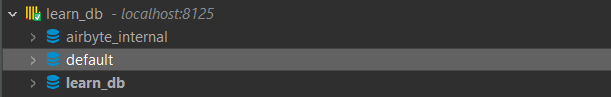
1. Запустил ClickHouse в Docker:

```bash

docker run --name clickhouse-full -e CLICKHOUSE\_DB=learn\_db -e CLICKHOUSE\_USER=admin -e CLICKHOUSE\_DEFAULT\_ACCESS\_MANAGEMENT=1 -e CLICKHOUSE\_PASSWORD=soil@WORK1 -p 8125:8123 -p 9000:9000/tcp -d -v clickhouse-logs:/var/log/clickhouse-server -v clickhouse-data:/var/lib/clickhouse clickhouse:25.4

```

2. Подключился через DBeaver



**2.3. Установка Airbyte**

1. Запустил Airbyte с помощью Docker Compose:

```bash

abctl local install

abctl local credentials --password YourStrongPasswordExample

```

2. Открыл веб-интерфейс по адресу `http://localhost:8000`.

**2.4. Настройка соединений в Airbyte**

**Источник (PostgreSQL) → Назначение (ClickHouse)**

1. В базе данных «test\_airbyte» создал пользователя и наделил его необходимыми правами

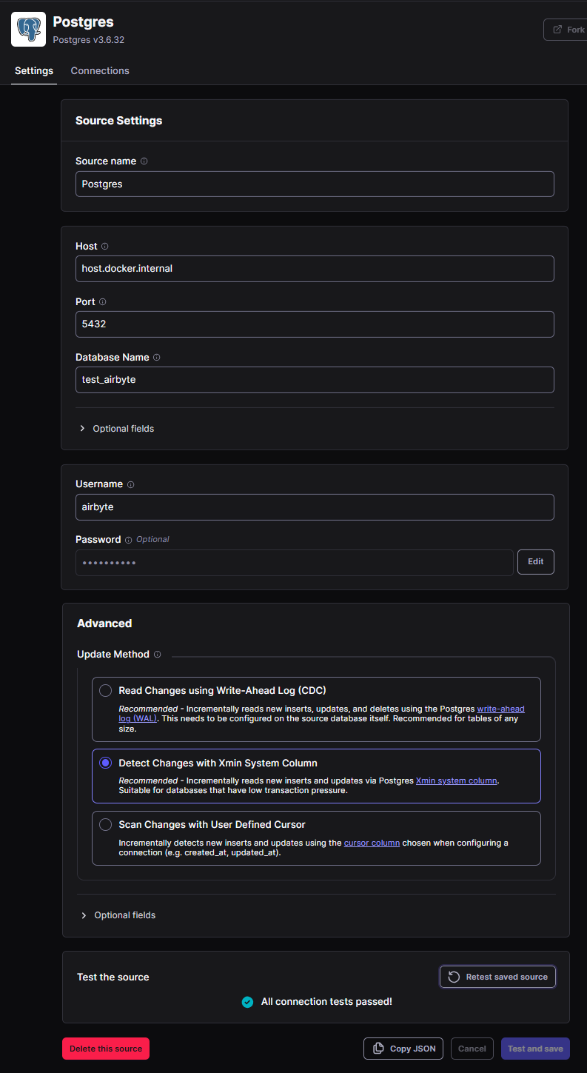
“”

CREATE USER airbyte WITH PASSWORD “my\_pass1”;

GRANT CREATE, TEMPORARY ON DATABASE "DestPostgrAirbyte" TO airbyte;

“””

1. Создал Source (PostgreSQL):

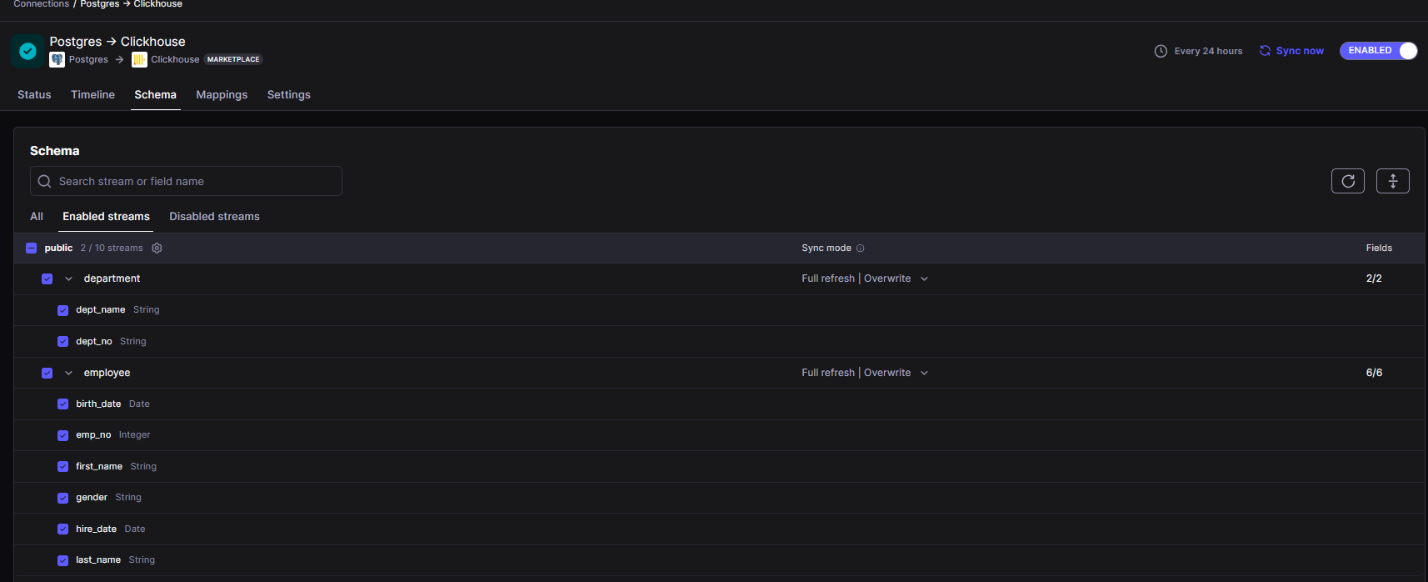


1. Создал Destination (ClickHouse):

Изображение выглядит как текст, снимок экрана, программное обеспечение, Мультимедийное программное обеспечение

Контент, сгенерированный ИИ, может содержать ошибки.

1. Настроил Connection таблиц «Employee» и «Department»:

- Выбрал режим полной синхронизации (Full Refresh | Overwrite).

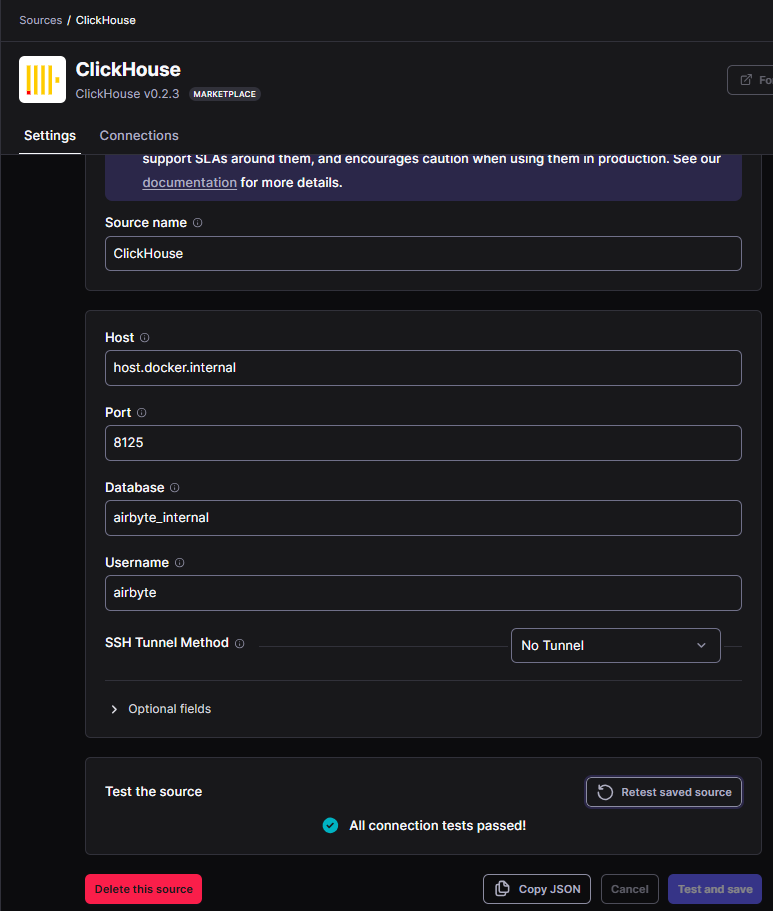
4. Запустил синхронизацию.

Изображение выглядит как текст, программное обеспечение, Мультимедийное программное обеспечение, Значок на компьютере

Контент, сгенерированный ИИ, может содержать ошибки.

**Обратная синхронизация (ClickHouse → PostgreSQL)**

1. Создал новый Source (ClickHouse).



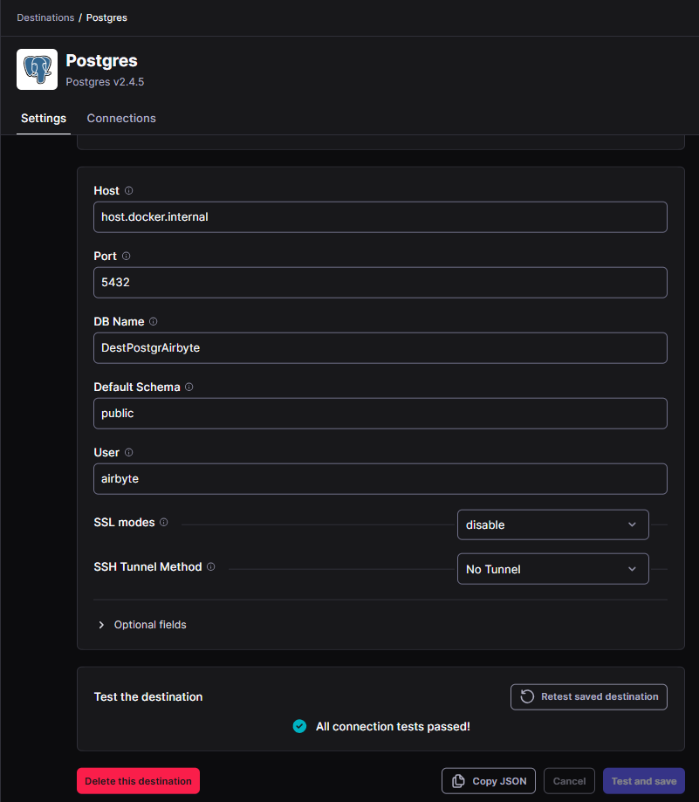
1. в базе данных «DestPostgrAirbyte» наделил пользователя необходимыми правами

“””

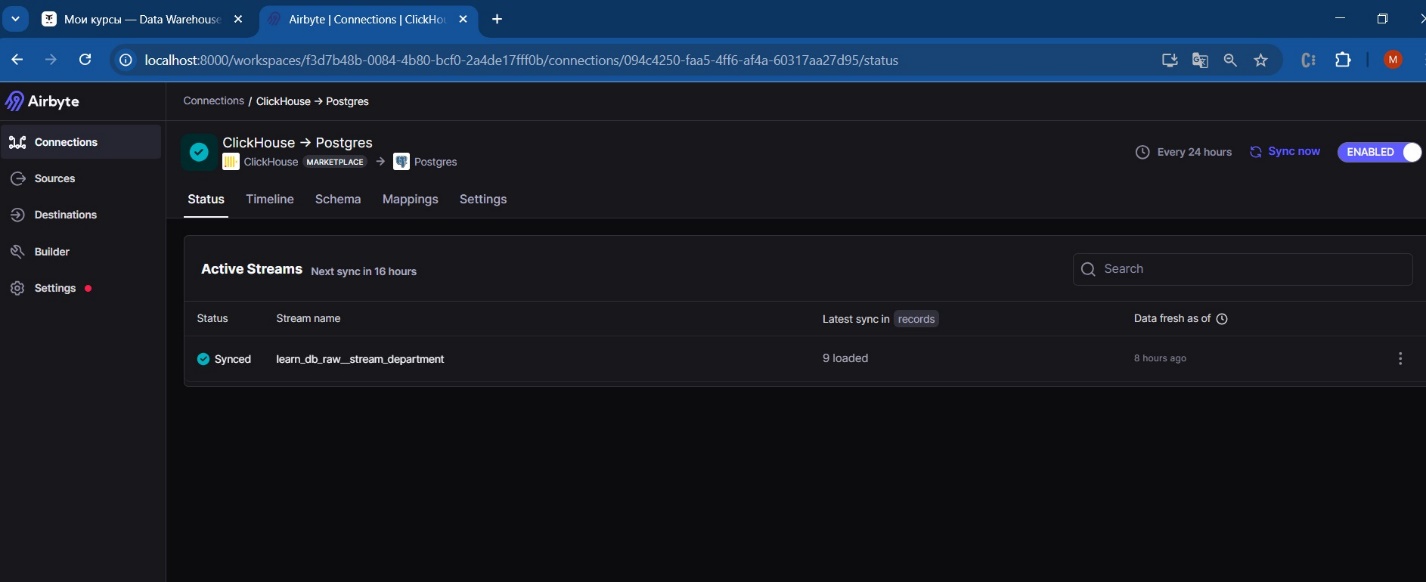
GRANT CREATE, TEMPORARY ON DATABASE "DestPostgrAirbyte" TO airbyte;

“””

1. Настроил Destination (PostgreSQL).



1. Запустил синхронизацию.



**3. Возникшие проблемы и их решение**

Проблема 1: Ошибка подключения к PostgreSQL из Airbyte

- Ошибка: `Connection refused` при попытке подключения.

- Решение: Проверил, что PostgreSQL слушает внешние подключения (`listen\_addresses = '\*'` в `postgresql.conf`) и разрешил подключения в `pg\_hba.conf`.

Проблема 2: Ошибка формата данных при синхронизации в ClickHouse

- Ошибка: Несоответствие типов данных (`String` vs `VARCHAR`).

- Решение: Привел схему таблиц к единому формату перед синхронизацией.

**4. Вывод**

- Успешно развернуты PostgreSQL и ClickHouse.

- Настроен Airbyte для двунаправленной синхронизации данных.

- Решены проблемы с подключением и типами данных.