

Департамент образования и науки города Москвы
Государственное автономное образовательное учреждение высшего
образования города Москвы
«Московский городской педагогический университет»
Институт цифрового образования
Департамент информатики, управления и технологий

ДИСЦИПЛИНА:

Проектный практикум по разработке ETL-решений

Лабораторная работа 6.1

Разработка полного ETL-процесса. Оркестровка конвейера данных

Выполнила: Шведова С.С., группа: АДЭУ-211

Преподаватель: Босенко Т.М.

Москва

2025

Задачи:

1. Запустить контейнер с Бизнес-кейсом «StockSense», изучить основные элементы DAG в Apache Airflow.
2. Спроектировать верхнеуровневую архитектуру аналитического решения Бизнес-кейса «StockSense»
3. Спроектировать архитектуру DAG Бизнес-кейса «StockSense»

На рисунке 1 показана верхнеуровневую архитектуру аналитического решения.

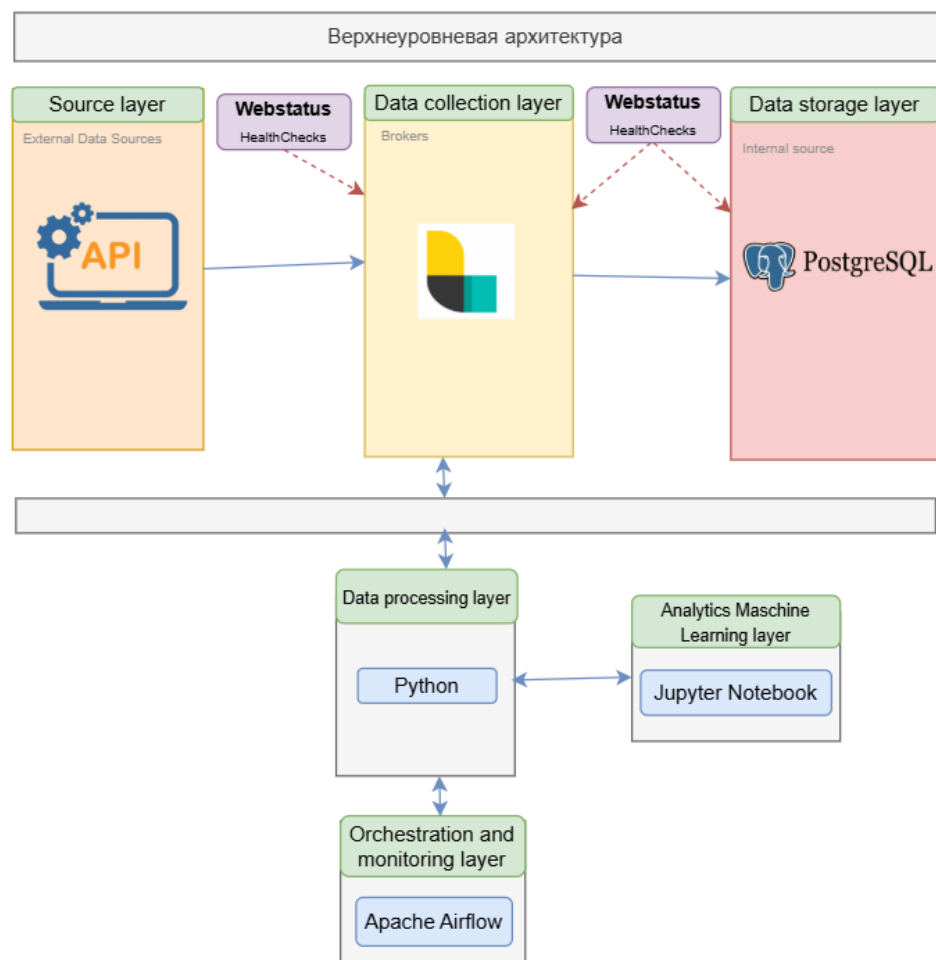


Рисунок 1. Верхнеуровневая архитектура

На рисунке 2 показана архитектура дага.

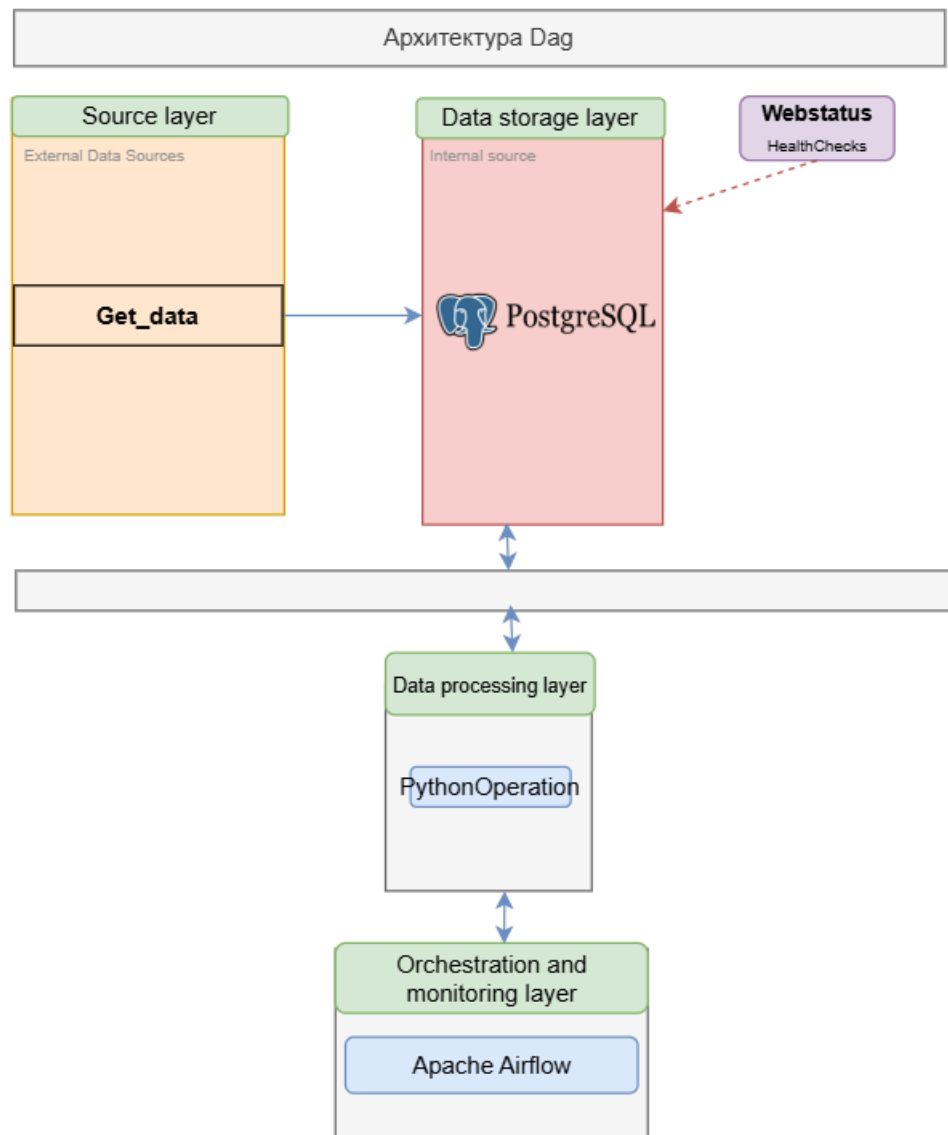


Рисунок 2. Архитектура дага

На рисунке 3 показано клонирование репозитория

```
dev@dev-vm:~$ git clone https://github.com/BosenkoTM/workshop-on-ETL.git
Cloning into 'workshop-on-ETL'...
remote: Enumerating objects: 637, done.
remote: Counting objects: 100% (30/30), done.
remote: Compressing objects: 100% (29/29), done.
remote: Total 637 (delta 18), reused 1 (delta 1), pack-reused 607 (from 1)
Receiving objects: 100% (637/637), 5.83 MiB | 7.18 MiB/s, done.
Resolving deltas: 100% (316/316), done.
dev@dev-vm:~$
```

Рисунок 3. Клонирование репозитория

На рисунке 4 показан запуск среды с помощью Docker Compose

```
dev@dev-vm:~/workshop-on-ETL/business_case_stocksense_25$ sudo docker compose up --build
[+] Running 8/8
✔ Network business_case_stocksense_25_default Created
✔ Volume "business_case_stocksense_25_postgres_data" Created
✔ Volume "business_case_stocksense_25_logs" Created
✔ Container business_case_stocksense_25-wiki_results-1 Created
✔ Container business_case_stocksense_25-postgres-1 Created
✔ Container business_case_stocksense_25-init-1 Created
✔ Container business_case_stocksense_25-webserver-1 Created
✔ Container business_case_stocksense_25-scheduler-1 Created
Attaching to init-1, postgres-1, scheduler-1, webserver-1, wiki_results-1
wiki_results-1 | The files belonging to this database system will be owned by user "postgres".
postgres-1 | The files belonging to this database system will be owned by user "postgres".
wiki_results-1 | This user must also own the server process.
postgres-1 | This user must also own the server process.
wiki_results-1 |
postgres-1 |
wiki_results-1 | The database cluster will be initialized with locale "en_US.utf8".
wiki_results-1 | The default database encoding has accordingly been set to "UTF8".
wiki_results-1 | The default text search configuration will be set to "english".
wiki_results-1 |
wiki_results-1 | Data page checksums are disabled.
wiki_results-1 |
wiki_results-1 | fixing permissions on existing directory /var/lib/postgresql/data ... ok
wiki_results-1 | creating subdirectories ... ok
wiki_results-1 | selecting dynamic shared memory implementation ... posix
postgres-1 | The database cluster will be initialized with locale "en_US.utf8".
postgres-1 | The default database encoding has accordingly been set to "UTF8".
postgres-1 | The default text search configuration will be set to "english".
```

Рисунок 4. Запуск среды

Как можно увидеть на рисунке 5, аирффлоу заработал

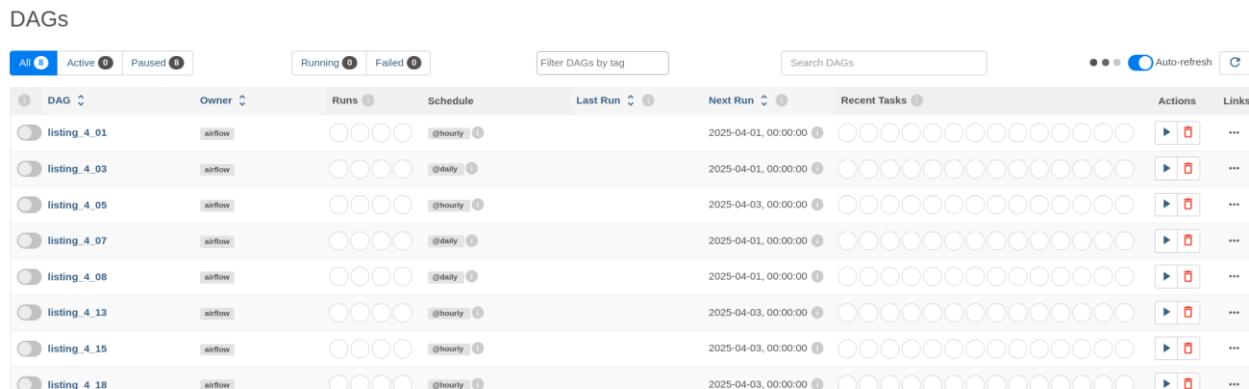


Рисунок 5. Аирфлоу заработал

На рисунке 6 показан граф дага

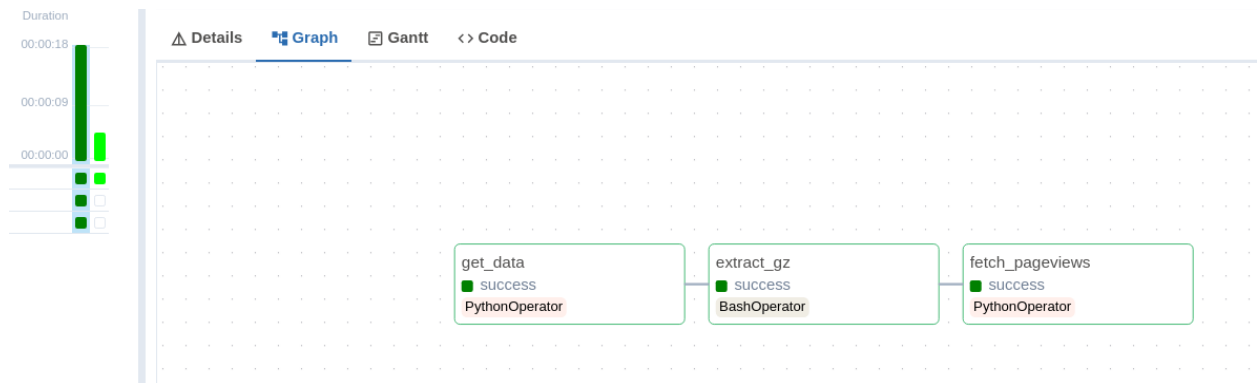


Рисунок 6. Граф дага

На рисунке 7 показана диаграмма Ганте графа

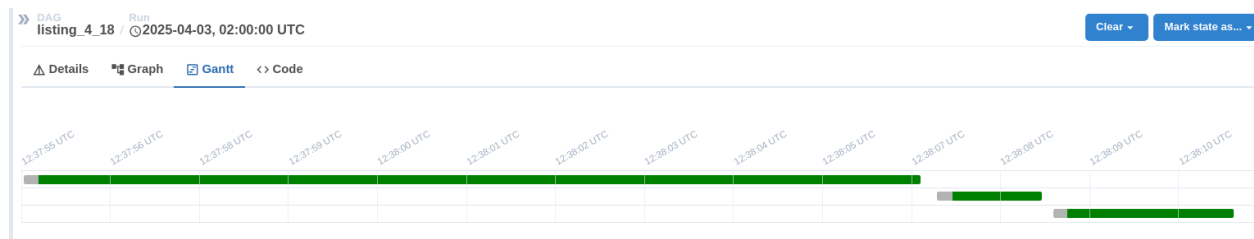


Рисунок 7. Диаграмма Ганта

На рисунке 8 показано подключение к постгрис

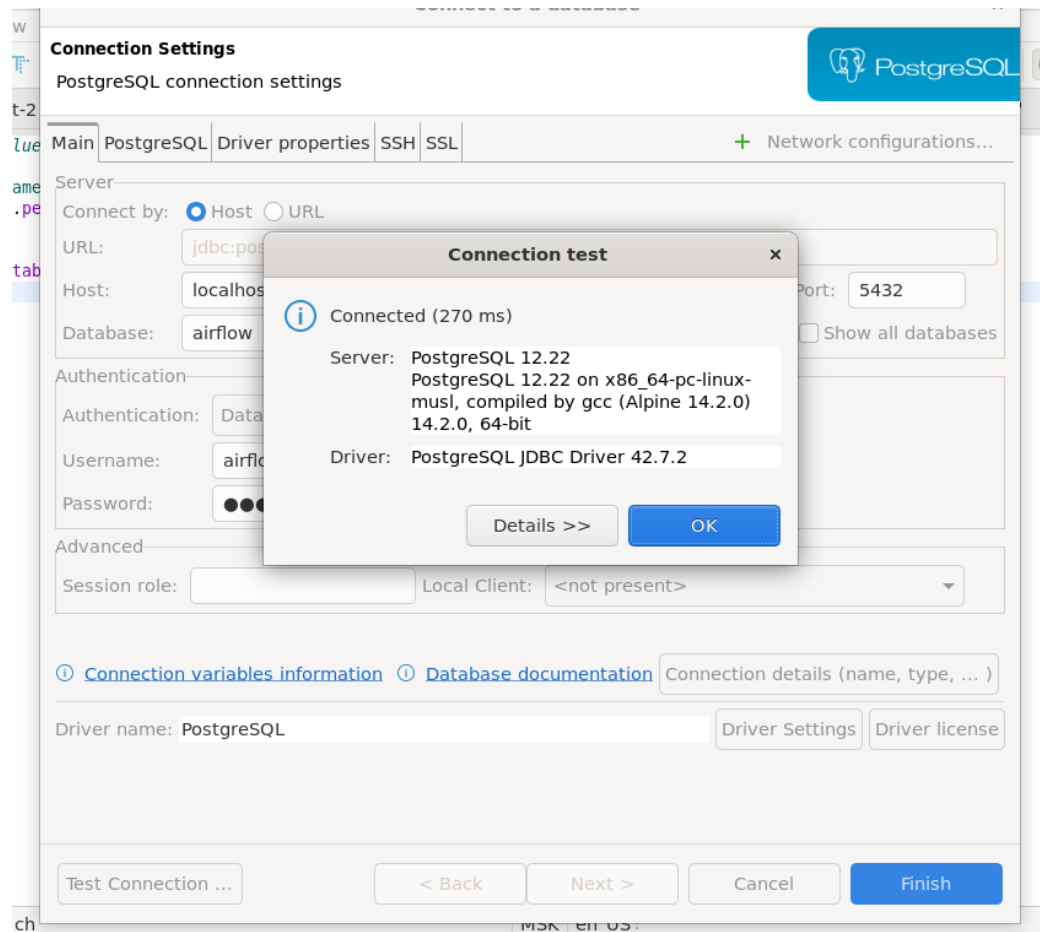


Рисунок 8. Подключение к постгрис

На рисунке 9 показаны таблицы в базе данных

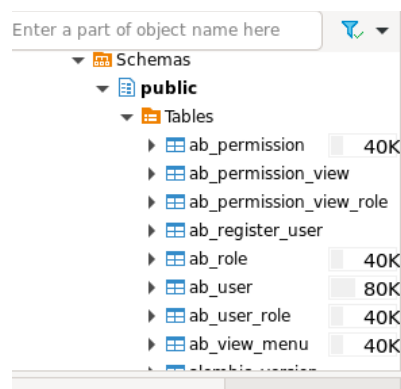


Рисунок 9. Таблицы в базе данных airflow

На рисунке 10 показана ERD-схема базы данных Postgre SQL;

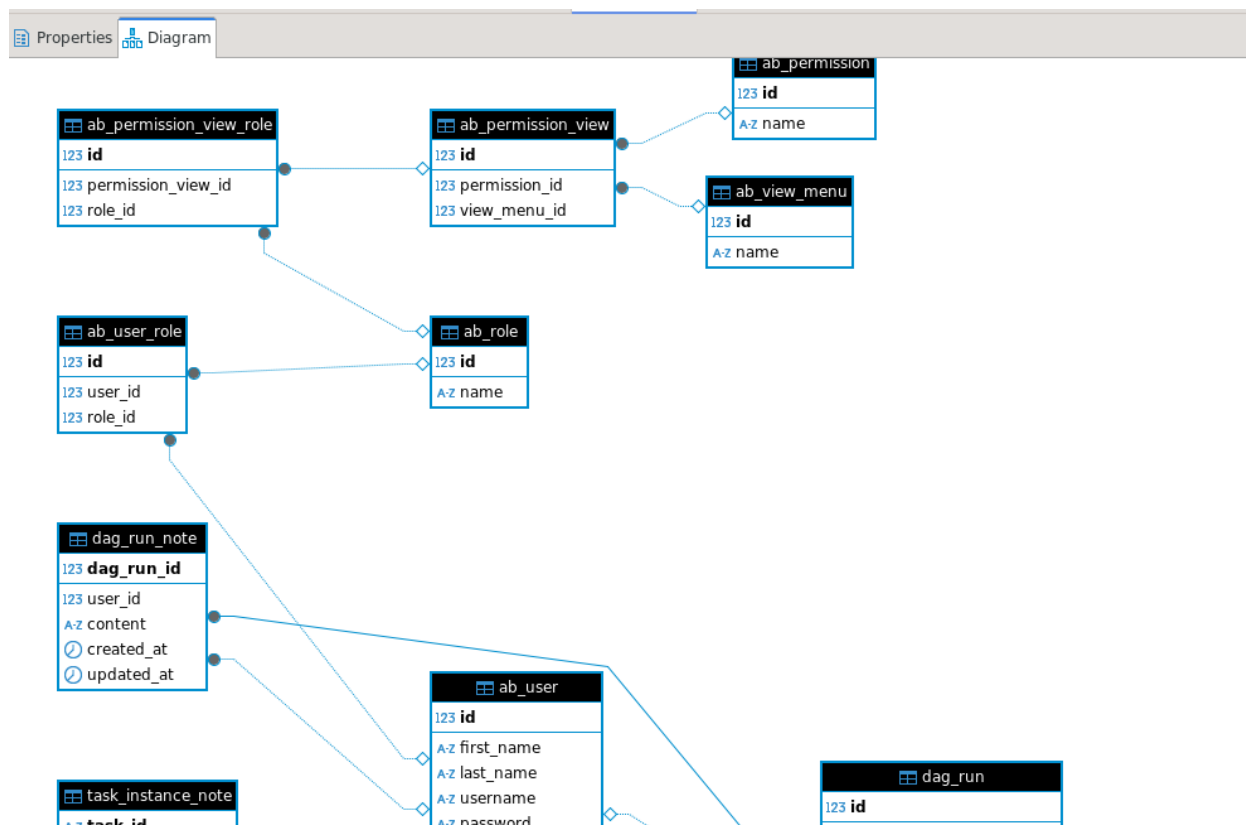


Рисунок 10. ERD-схема базы данных Postgre SQL

На рисунке 11 показано выполнение запроса к базе данных

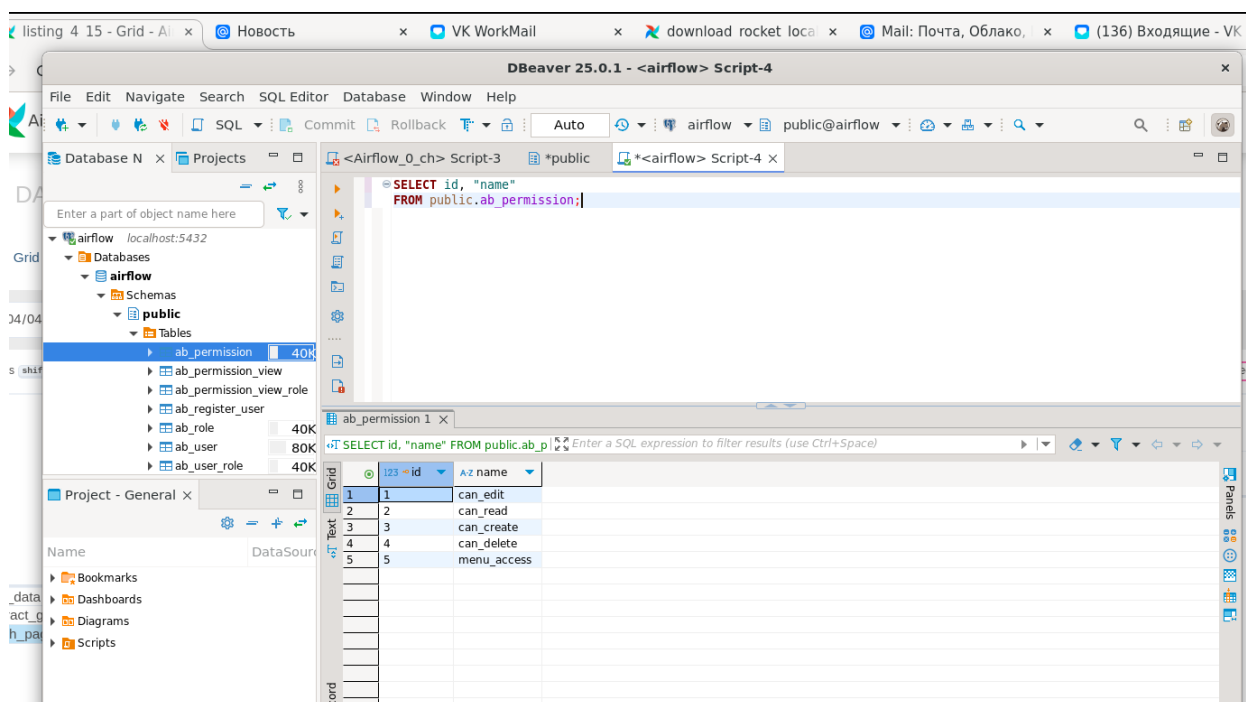


Рисунок 11. Выполнение запроса

Выводы:

1. Запущен контейнер с Бизнес-кейсом «StockSense», изучить основные элементы DAG в Apache Airflow.
2. Спроектирована верхнеуровневая архитектура аналитического решения Бизнес-кейса «StockSense»
3. Спроектирована архитектура DAG Бизнес-кейса «StockSense»