

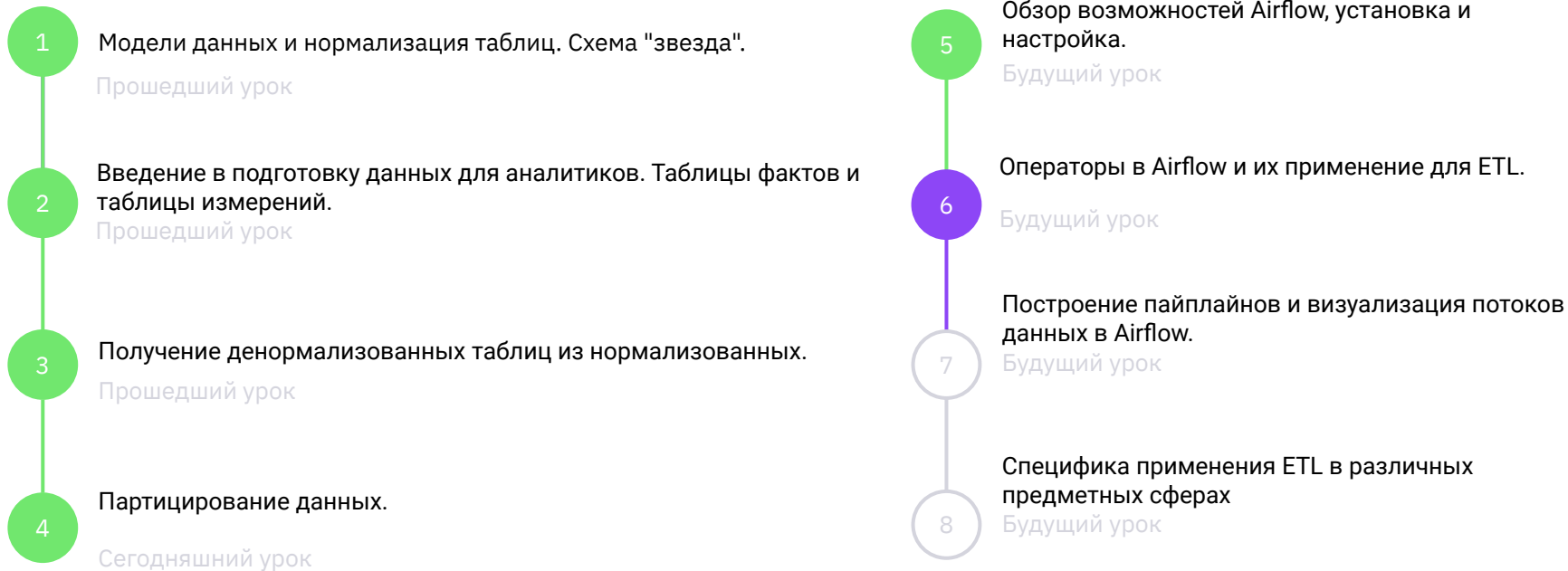
Операторы в Airflow и их применение для ETL

Урок 6





План курса (вертикальный)





Что будет на уроке сегодня



Что такое операторы в Apache Airflow



Типы операторов Apache Airflow



Викторина



Что такое операторы Airflow?

1. Части пайплайна отвечающие за выполнение команд
2. Команды, выполняемые асинхронным графом



Что такое операторы Airflow?

1. Части пайплайна отвечающие за выполнение команд
2. Команды, выполняемые асинхронным графом



Всегда ли оператор генерирует результаты выполнения?

1. Да обязательно
2. Нет чаще всего оператор не генерирует результатов



Всегда ли оператор генерирует результаты выполнения?

1. Да обязательно
2. Нет чаще всего оператор не генерирует результатов



Какие существуют типы операторов?

1. Оператор действия
2. Оператор перемещения
3. Оператор добавления
4. Сенсор оператор
5. Все варианты верны



Какие существуют типы операторов?

1. Оператор действия
2. Оператор перемещения
3. Оператор добавления
4. Сенсор оператор
5. Все варианты верны



Что такое оператор действия?

1. Любая команда выполняемая в системе
2. Оператор ожидающий наступление какого-то события чтобы выполнить команду
3. Команда выполняющая определенное действие
4. Оператор отвечающий за перемещение данных из одной системы в другую



Что такое оператор действия?

1. Любая команда выполняемая в системе
2. Оператор ожидающий наступление какого-то события чтобы выполнить команду
3. Команда выполняющая определенное действие
4. Оператор отвечающий за перемещение данных из одной системы в другую



Что такое оператор перемещения?

1. Любая команда выполняемая в системе
2. Оператор ожидающий наступление какого-то события чтобы выполнить команду
3. Команда выполняющая определенное действие
4. Оператор отвечающий за перемещение данных из одной системы в другую



Что такое оператор перемещения?

1. Любая команда выполняемая в системе
2. Оператор ожидающий наступление какого-то события чтобы выполнить команду
3. Команда выполняющая определенное действие
4. Оператор отвечающий за перемещение данных из одной системы в другую



Что такое сенсор оператор?

1. Любая команда выполняемая в системе
2. Оператор ожидающий наступление какого-то события чтобы выполнить команду
3. Команда выполняющая определенное действие
4. Оператор отвечающий за перемещение данных из одной системы в другую



Что такое сенсор оператор?

1. Любая команда выполняемая в системе
2. Оператор ожидающий наступление какого-то события чтобы выполнить команду
3. Команда выполняющая определенное действие
4. Оператор отвечающий за перемещение данных из одной системы в другую



Вопросы?

Вопросы?



Вопросы?





Практика



Типы операторов Airflow

Оператор действия

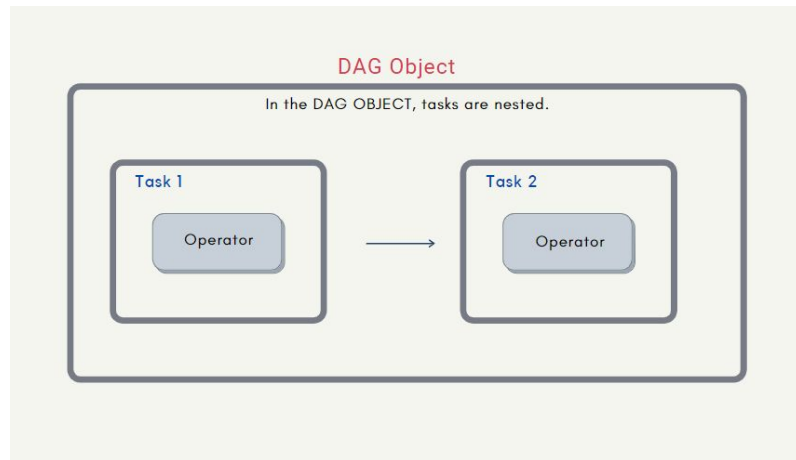
- Это программа, которая выполняет определенное действие.
- Например, *EmailOperator* и *BashOperator*.

Оператор перемещения

- Он отвечает за перемещение данных из одной системы в другую.
- Если вы работаете с большим набором данных, избегайте использования этого оператора.

Оператор sensor

- Оператор sensor ожидает поступления данных в определенное место.
- Это долгосрочные задачи.
- Они полезны для отслеживания внешних процессов, таких как загрузка файлов.





Задание 1

Создайте новый граф, добавьте в него два BashOperator, первый должен выводить на экран сообщение “Hello from Airflow”, второй должен брать код из bash файла. Создайте bash файл который будет выводить сообщение “Hello from Airflow bash script processor.” Создайте последовательную связь между первым и вторым оператором.

Добавьте между первым и вторым операторами еще один BashOperator который будет заканчивать свою работу в статусе skipped

A large teal circle containing the text "30 минут" in white.

30 минут

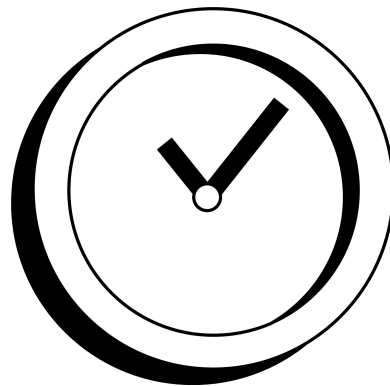


Задание 1

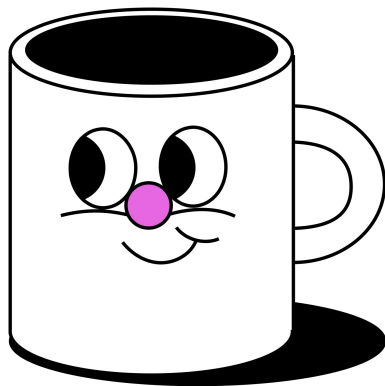
Создайте новый граф, добавьте в него два BashOperator, первый должен выводить на экран сообщение “Hello from Airflow”, второй должен брать код из bash файла. Создайте bash файл который будет выводить сообщение “Hello from Airflow bash script processor.” Создайте последовательную связь между первым и вторым оператором.

Добавьте между первым и вторым операторами еще один BashOperator который будет заканчивать свою работу в статусе skipped

<<30:00->>



Перерыв



<<5:00->>



Задание 2

В графе из предыдущего задания поменяйте BashOperator на pythonOperator функционал измениться не должен.

A large teal circle on the right side of the slide, containing the text '10 минут' in white.

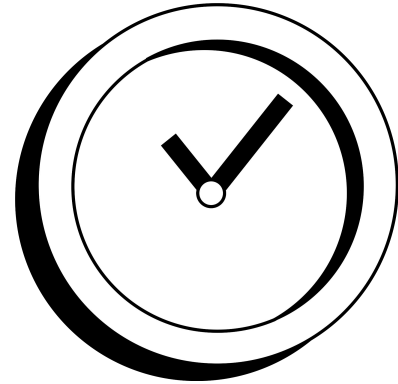
10 минут



Задание 2

В графе из предыдущего задания поменяйте BashOperator на pythonOperator функционал измениться не должен.

<<10:00->>





Задание 3

Добавьте в граф httpSensor который будет обращаться к сайту gb.ru.
Отправьте в чат скриншот кода и логи работы

A large teal circle containing the text '15 минут' in white.

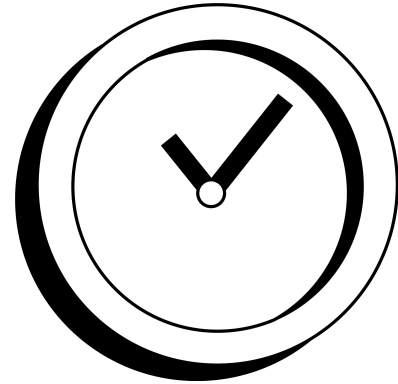
15 минут



Задание 3

Добавьте в граф httpSensor который будет обращаться к сайту gb.ru.
Отправьте в чат скриншот кода и логи работы

<<15:00->>





Задание 4

Создайте ShortCircuitOperator и измените граф таким образом чтобы было видно как выполняются два различных сценария в зависимости от результатов работы данного оператора.



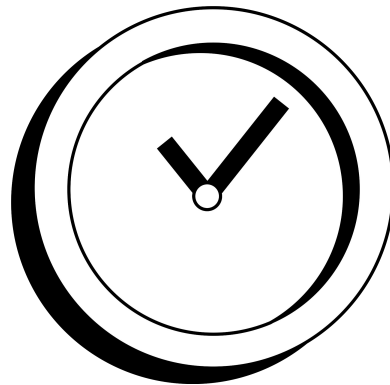
15 минут



Задание 4

Создайте ShortCircuitOperator и измените граф таким образом чтобы было видно как выполняются два различных сценария в зависимости от результатов работы данного оператора.

<<15:00->>





Вопросы?

Вопросы?



Вопросы?





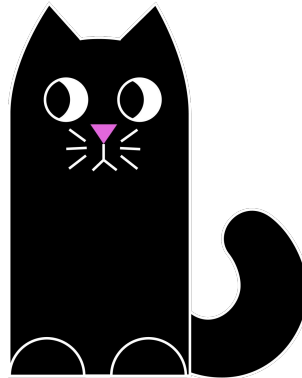
Домашнее задание



Домашнее задание

На основе сайта yandex.ru:

- Определите, на каком протоколе работает сайт.
- Проанализируйте структуру страницы сайта
- Внесите не менее 10 изменений на страницу с помощью инструмента разработчика и представьте скриншоты было/стало.
- Создайте прототип низкой детализации (дополнительное задание, если на семинаре дошли до задания №8)





Спасибо за внимание

