

1. Устанавливаем: `wsl --install`
2. Установим Ubuntu (18.04)
3. Установим MobaXterm
4. Добавим репозиторий в MobaXterm:
`sudo add-apt-repository ppa:deadsnakes/ppa`
5. `sudo apt update`
6. `sudo apt install python3.8`
7. `sudo apt install python3.8-distutils`
8. `wget https://bootstrap.pypa.io/get-pip.py`
9. `sudo python3.8 get-pip.py`
10. `sudo apt-get install mysql-server`
11. `sudo /etc/init.d/mysql start`
12. `sudo mysql_secure_installation` (yes в конце reload)
13. `sudo mysql -u root -p` (enter вместо пароля)
14. Проверим все ли у нас работает:
`select now(), 123 test;`
`exit`
15. `sudo mysql -u root -p`
`ALTER USER 'root'@'localhost' IDENTIFIED WITH mysql_native_password BY '1';`
`exit`
16. `sudo nano /etc/mysql/my.cnf`
`[mysqld]`
`port = 33061`
`ctrl+o – enter – ctrl x`
17. `sudo apt-get install python3.8-dev libmysqlclient-dev && sudo apt install pkg-config &&`
`pip install mysqlclient` Ошибка
18. Перезагрузим mysql
`sudo service mysql restart`
19. Запускаем heidsql (проверяем подключение, вводим 33061 и пароль 1)
20. Создаем схему для airflow в heidsql:
`CREATE DATABASE Airflow CHARACTER SET UTF8mb3 COLLATE utf8_general_ci;`
`CREATE USER 'Airflow'@'%' IDENTIFIED BY '1';`
21. Нажимаем пользователи airflow в heidsql и добавляем глобальные разрешения

22. `sudo apt install -y python3 python3-pip python3-venv`
23. `python3 -m venv airflow-venv && source airflow-venv/bin/activate`
24. `pip install mysqlclient`
25. `pip install apache-airflow==2.7.3 --constraint
"https://raw.githubusercontent.com/apache/airflow/constraints-2.7.3/constraints-
3.8.txt" --force-reinstall --upgrade`
26. `cd ..` - переходим к папке home
27. `sudo chmod -R 777 ./`
28. `export AIRFLOW_HOME=/home/airflow && airflow db init`
29. `ls`
30. Открываем папку home/airflow и создаем папку dags
31. В папке home/airflow меняем airflow.cfg на файл из облака к семинару
32. `airflow db init`
33. Перезагрузим `sudo service mysql restart`
34. `airflow db init`
35. `airflow users create --role Admin --username admin --email admin --firstname admin --
lastname admin --password admin`
36. `export AIRFLOW_HOME=/home/airflow && (airflow scheduler & airflow webserver -p
8080)`
37. Заходим в браузер и переходим <http://localhost:8080>
user: admin
password: admin

4. WSL-Ubuntu

```
.1.0.0 Safari/537.36"
127.0.0.1 - - [17/Apr/2024:10:16:35 +1000] "GET /static/appbuilder/js/ab_filters
.js HTTP/1.1" 304 0 "http://localhost:8080/home" "Mozilla/5.0 (Windows NT 10.0;
Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/120.0.0.0 YaBrowser/24
.1.0.0 Safari/537.36"
127.0.0.1 - - [17/Apr/2024:10:16:35 +1000] "GET /static/appbuilder/datepicker/bo
otstrap-datepicker.js HTTP/1.1" 304 0 "http://localhost:8080/home" "Mozilla/5.0
(Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/120.
0.0.0 YaBrowser/24.1.0.0 Safari/537.36"
127.0.0.1 - - [17/Apr/2024:10:16:35 +1000] "GET /static/appbuilder/js/bootstrap.
min.js HTTP/1.1" 304 0 "http://localhost:8080/home" "Mozilla/5.0 (Windows NT 10.
0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/120.0.0.0 YaBrowser
/24.1.0.0 Safari/537.36"
127.0.0.1 - - [17/Apr/2024:10:16:35 +1000] "GET /static/appbuilder/select2/selec
t2.js HTTP/1.1" 304 0 "http://localhost:8080/home" "Mozilla/5.0 (Windows NT 10.0
; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/120.0.0.0 YaBrowser/
24.1.0.0 Safari/537.36"
127.0.0.1 - - [17/Apr/2024:10:16:35 +1000] "GET /static/appbuilder/js/ab.js HTTP
/1.1" 304 0 "http://localhost:8080/home" "Mozilla/5.0 (Windows NT 10.0; Win64; x
64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/120.0.0.0 YaBrowser/24.1.0.0 S
afari/537.36"
127.0.0.1 - - [17/Apr/2024:10:16:35 +1000] "GET /static/dist/moment.0fcb6b41ff6a
87cf079e.js HTTP/1.1" 304 0 "http://localhost:8080/home" "Mozilla/5.0 (Windows N
T 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/120.0.0.0 YaBr
owser/24.1.0.0 Safari/537.36"
127.0.0.1 - - [17/Apr/2024:10:16:35 +1000] "GET /static/dist/main.bde72ea87585eb
c44fe9.js HTTP/1.1" 304 0 "http://localhost:8080/home" "Mozilla/5.0 (Windows NT
10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/120.0.0.0 YaBrow
ser/24.1.0.0 Safari/537.36"
127.0.0.1 - - [17/Apr/2024:10:16:35 +1000] "GET /static/dist/bootstrap-datetimep
icker.min.js HTTP/1.1" 304 0 "http://localhost:8080/home" "Mozilla/5.0 (Windows
NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/120.0.0.0 YaB
rowser/24.1.0.0 Safari/537.36"
127.0.0.1 - - [17/Apr/2024:10:16:35 +1000] "GET /static/dist/bootstrap3-typeahea
d.min.js HTTP/1.1" 304 0 "http://localhost:8080/home" "Mozilla/5.0 (Windows NT 1
0.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/120.0.0.0 YaBrows
er/24.1.0.0 Safari/537.36"
[2024-04-17T10:16:43.273+1000] {scheduler_job_runner.py:1605} INFO - Adopting or
resetting orphaned tasks for active dag runs
[2024-04-17 10:20:17 +1000] [13173] [INFO] Handling signal: winch
[2024-04-17 10:20:17 +1000] [13170] [INFO] Handling signal: winch
```


Войти

1

+

DAGs - Airflow

+

localhost:8080

DAGs - Airflow

Airflow

DAGs

Cluster Activity

Datasets

Security

Browse

Admin

Docs

00:17 UTC

AA

DAGs

All 0

Active 0

Paused 0

Running 0

Failed 0

Filter DAGs by tag

Search DAGs

Auto-refresh

DAG

Owner

Schedule

Actions

Links

No results

Showing 0-0 of 0 DAGs

Version: v2.7.3

Git Version: .release:f1243537838516b8bb8156130bc001595bfbeeb01

Войти

1

+

my_first_dag - Grid - Air

localhost:8080

my_first_dag - Grid - Airflow

перевести

00:35 UTC

AA

Airflow

DAGs

Cluster Activity

Datasets

Security

Browse

Admin

Docs

The scheduler does not appear to be running. Last heartbeat was received 9 minutes ago.

The DAGs list may not update, and new tasks will not be scheduled.

DAG: my_first_dag

Hello World DAG

Schedule: 0 12 * * *

Next Run: 2024-04-15, 12:00:00

Grid

Graph

Calendar

Task Duration

Task Tries

Landing Times

Gantt

Details

Code

Audit Log

17.04.2024 00:35:30

25

All Run Types

All Run States

Clear Filters

Auto-refresh

Press shift + / for Shortcuts

deferred

failed

queued

removed

restarting

running

scheduled

skipped

success

up_for_reschedule

up_for_retry

upstream_failed

no_status

«»

DAG

my_first_dag

Duration

00:00:00

00:00:00

00:00:00

hello_task

Details

Graph

Gantt

Code

Parsed at: 2024-04-17, 00:35:17 UTC

1 from datetime import datetime

2 from airflow import DAG

3 from airflow.operators.dummy_operator import DummyOperator

4 from airflow.operators.python_operator import PythonOperator

5

6 def print_hello():

7 return 'Hello world from first Airflow DAG!'

8

9 dag = DAG('my_first_dag', description='Hello World DAG',

10 schedule_interval='0 12 * * *',

Toggle Wrap

▼ Unnamed-1

▼ Airflow	1,2 MiB
ab_permission	32,0 KiB
ab_permission_view	48,0 KiB
ab_permission_vie...	48,0 KiB
ab_register_user	32,0 KiB
ab_role	32,0 KiB
ab_user	80,0 KiB
ab_user_role	48,0 KiB
ab_view_menu	32,0 KiB
alembic_version	16,0 KiB
callback_request	16,0 KiB
connection	32,0 KiB
dag	16,0 KiB
dagrun_dataset_ev...	32,0 KiB
dag_code	16,0 KiB
dag_owner_attribut...	16,0 KiB
dag_pickle	16,0 KiB
dag_run	64,0 KiB
dag_run_note	32,0 KiB
dag_schedule_data...	32,0 KiB
dag_tag	32,0 KiB
dag_warning	16,0 KiB
dataset	16,0 KiB
dataset_dag_run_q...	32,0 KiB
dataset_event	16,0 KiB
import_error	16,0 KiB
job	16,0 KiB

```

1 CREATE DATABASE Airflow CHARACTER SET UTF8mb3 COLLATE utf8_general_ci;
2 CREATE USER 'Airflow'@'%' IDENTIFIED BY '1';
3

```

```

44 SELECT * FROM information_schema.REFERENTIAL_CONSTRAINTS WHERE CONSTRAINT_SCHEMA='Airflow' AND TABLE_NAME='dataset' AND REFERENCED_TABLE_NAME IS NOT NULL;
45 SELECT * FROM information_schema.KEY_COLUMN_USAGE WHERE TABLE_SCHEMA='Airflow' AND TABLE_NAME='dataset' AND REFERENCED_TABLE_NAME IS NOT NULL;
46 SHOW CREATE TABLE `Airflow`.`dataset`;
47 SELECT tc.CONSTRAINT_NAME, cc.CHECK_CLAUSE FROM `information_schema`.`CHECK_CONSTRAINTS` AS cc, `information_schema`.`TABLE_CONSTRAINTS` AS tc WHERE tc.CONSTRAINT_SCHEMA='Airflow' AND tc.TABLE_NAME='dataset' AND tc.CONSTRAINT_TYPE='CHECK' AND tc.CONSTRAINT_SCHEMA=
48 SHOW VARIABLES;
49 FLUSH PRIVILEGES;
50 SHOW COLUMNS FROM `mysql`.`user`;
51 SELECT `user`, `host`, authentication_string AS `password` FROM `mysql`.`user`;

```