

ООО «Астра Консалтинг»

Астра ИС МД (Инфраструктурные Сервисы)

**Модуль Astra Apache Airflow**

**ИНСТРУКЦИЯ ПО УСТАНОВКЕ ЭКЗЕМПЛЯРА ПРОГРАММНОГО  
ОБЕСПЕЧЕНИЯ, ПРЕДОСТАВЛЕННОГО ДЛЯ ПРОВЕДЕНИЯ  
ЭКСПЕРТНОЙ ПРОВЕРКИ**

Москва, 2026

## Содержание

1 Область применения .....	3
2 Установка системы.....	4
2.1 Технические требования к оборудованию.....	4
2.2 Минимальные требования к сторонним компонентам.....	4
3 Архитектура системы.....	5
3.1 Характеристики системы .....	5
3.2 Структура развертывания Standalone .....	5
4 Установка и запуск.....	6
4.1 Шаг 1: Подготовка окружения.....	6
4.2 Шаг 2: Создание виртуального окружения .....	6
4.3 Шаг 3: Установка Apache Airflow.....	6
4.4 Шаг 4: Запуск экземпляра ПО .....	6
4.5 Шаг 5: Проверка статуса экземпляра ПО .....	7
5 Доступ к сервисам.....	8
5.1 Точки доступа .....	8
5.2 Аутентификация в веб-интерфейсе .....	8
5.3 Доступ через CLI.....	8
6 Проверка работы .....	9
6.1 Базовая проверка.....	9
6.2 Тестирование выполнения DAG .....	9
6.3 Проверка журналов и состояния компонентов.....	9
7 7. Управление экземпляром ПО .....	10
7.1 Основные команды управления.....	10
7.2 Повторный запуск экземпляра ПО .....	10
8 Просмотр логов работы системы.....	11
8.1 Журналы Airflow.....	11
8.2 Просмотр журналов в реальном времени .....	11

## **1 Область применения**

Настоящий документ предназначен для установки и запуска экземпляра программного обеспечения Астра ИС МД, модуль Astra Apache Airflow, в режиме Standalone для проведения экспертной проверки. Документ описывает порядок подготовки окружения, установки, первого запуска, проверки работоспособности и базового сопровождения экземпляра программного обеспечения.

В указанном режиме Astra Apache Airflow автоматически разворачивает веб-интерфейс, планировщик задач, встроенную базу метаданных SQLite и формирует учетную запись администратора.

## 2 Установка системы

### 2.1 Технические требования к оборудованию

Для локальной инсталляции требуется виртуальный сервер или физический сервер с минимальными техническими характеристиками, представленными в Таблице 1. В официальной документации Airflow рекомендуется минимум 4 ГБ оперативной памяти, при этом фактические требования зависят от нагрузки и конфигурации стенда.

Вид ресурса	Ед. измерения	Требуемый объем
Частота процессора	ГГц	2,0 и выше
vCPU количество процессоров и ядер	шт.	2
Объем оперативной памяти	ГБ	4
Объем жесткого диска	ГБ	20
Сетевой интерфейс	шт.	1
Скорость сетевого интерфейса	Гбит/с	1

### 2.2 Минимальные требования к сторонним компонентам

Для установки системы на стенде используется интерпретатор Python и стандартный пакетный менеджер pip. Официальная документация Airflow 3.1.8 указывает поддержку Python версий 3.10, 3.11, 3.12 и 3.13

Требуемые компоненты:

- Операционная система Linux (Ubuntu, Debian, Astra Linux или совместимая);
- Python версии 3.10–3.13;
- pip актуальной версии;
- python3-venv для создания виртуального окружения.

```
sudo apt update
```

```
sudo apt install -y python3 python3-pip python3-venv
```

## 3 Архитектура системы

### 3.1 Характеристики системы

Параметр	Значение
Версия Apache Airflow	3.1.8
Режим развертывания	Standalone
Порт веб-интерфейса	8080
Аутентификация	Включена
Учетная запись	admin (создается автоматически)
База метаданных	SQLite
Исполнитель задач	LocalExecutor
Каталог журналов	\$AIRFLOW_HOME/logs

### 3.2 Структура развертывания Standalone

В режиме Standalone экземпляр программного обеспечения запускается на одном сервере и включает следующие компоненты:

- webserver — веб-интерфейс управления Apache Airflow;
- scheduler — планировщик задач;
- metadata database — встроенная база данных SQLite;
- демонстрационные DAG — примеры рабочих процессов, создаваемые при первом запуске.

Команда `airflow standalone` автоматически создает рабочий каталог Airflow Home, выполняет инициализацию базы данных, запускает webserver и scheduler, а также создает пользователя admin.

## 4 Установка и запуск

### 4.1 Шаг 1: Подготовка окружения

Создайте рабочую директорию и перейдите в неё:

```
mkdir -p ~/airflow-standalone  
cd ~/airflow-standalone
```

Установите необходимые пакеты:

```
sudo apt update  
sudo apt install -y python3 python3-pip python3-venv
```

### 4.2 Шаг 2: Создание виртуального окружения

Создайте и активируйте виртуальное окружение Python:

```
python3 -m venv airflow_venv  
source airflow_venv/bin/activate  
pip install --upgrade pip
```

### 4.3 Шаг 3: Установка Apache Airflow

Для установки системы используется официальный способ установки через pip с constraint-файлом. Пример команд для последней стабильной версии документации Airflow 3.1.8:

```
AIRFLOW_VERSION=3.1.8  
PYTHON_VERSION="$(python -c 'import sys;  
print(f"{sys.version_info.major}.{sys.version_info.minor}")')"  
CONSTRAINT_URL="https://raw.githubusercontent.com/apache/airflow/  
constraints-${AIRFLOW_VERSION}/constraints-  
${PYTHON_VERSION}.txt"  
pip install "apache-airflow==${AIRFLOW_VERSION}" --constraint  
"${CONSTRAINT_URL}"
```

### 4.4 Шаг 4: Запуск экземпляра ПО

Для запуска системы выполните команду:

```
airflow standalone
```

При первом запуске автоматически выполняются создание базы данных SQLite, запуск scheduler и webserver, создание пользователя admin и генерация примеров DAG.

#### **4.5 Шаг 5: Проверка статуса экземпляра ПО**

После запуска системы необходимо убедиться, что экземпляр программного обеспечения запущен корректно и веб-интерфейс доступен по порту 8080.

```
ps aux | grep airflow  
ss -tulpn | grep 8080
```

При запуске в фоновом режиме рекомендуется контролировать статус процесса и журнал запуска.

## 5 Доступ к сервисам

### 5.1 Точки доступа

Сервис	URL / команда	Аутентификация
Web-интерфейс Apache Airflow	http://localhost:8080	требуется
CLI Apache Airflow	airflow <команда>	не требуется

### 5.2 Аутентификация в веб-интерфейсе

Откройте в браузере адрес `http://localhost:8080` или `http://<IP-адрес-сервера>:8080`.

Для входа используется учетная запись:

```
Логин: admin
```

```
Пароль: автоматически выводится в консоли после выполнения  
команды airflow standalone
```

### 5.3 Доступ через CLI

Примеры команд CLI:

```
airflow version  
airflow dags list  
airflow tasks list example_bash_operator
```

## 6 Проверка работы

### 6.1 Базовая проверка

Проверка версии установленного программного обеспечения:

```
airflow version
```

Проверка списка доступных DAG:

```
airflow dags list
```

### 6.2 Тестирование выполнения DAG

В составе standalone-экземпляра присутствуют демонстрационные DAG. Для тестирования необходимо активировать любой DAG в веб-интерфейсе и выполнить его ручной запуск.

Ожидаемый результат: в интерфейсе Airflow статус задач должен измениться на success после успешного выполнения.

### 6.3 Проверка журналов и состояния компонентов

Проверка каталогов журналов:

```
ls -la ~/airflow/logs
```

При необходимости следует проверить доступность порта 8080 и наличие запущенных процессов Airflow.

## 7.7. Управление экземпляром ПО

### 7.1 Основные команды управления

```
airflow version  
airflow dags list  
pkill -f 'airflow standalone'
```

### 7.2 Повторный запуск экземпляра ПО

Для повторного запуска необходимо активировать виртуальное окружение и снова выполнить команду **airflow standalone**. При запуске через `nohup` или `systemd` допускается выполнение в фоновом режиме.

## 8 Просмотр логов работы системы

Журналы дают возможность просматривать сообщения, выводимые программным обеспечением в процессе запуска и выполнения рабочих процессов.

### 8.1 Журналы Airflow

Журналы задач и компонентов хранятся в каталоге:

```
$AIRFLOW_HOME/logs
```

### 8.2 Просмотр журналов в реальном времени

При запуске в фоновом режиме с перенаправлением вывода в файл журнал может просматриваться в реальном времени следующей командой:

```
tail -f airflow.log
```

При интерактивном запуске основной поток сообщений отображается непосредственно в консоли.