

Загрузка данных в S3 хранилище Я.Облако

В нашем сервисе появилась возможность загружать данные в Я.Облако (S3)

S3 работает как файлохранилище. Про его настройки можно отдельно почитать в документации самого сервиса: <https://cloud.yandex.ru/docs/storage/quickstart>

S3 помечается отдельно, так как имеется ввиду технология облачного хранилища Amazon S3, на базе которого можно организовать собственное хранилище в Я.Облако

В данный момент для загрузки в Я.Облако доступны данные из следующих систем:

- AppsFlyer
- Загрузка из CSV
- Google BigQuery
- Yandex.Direct
- Yandex.Metrica

Особенность сервиса заключается в том что по сути это файло-хранилище, поэтому потоки данных в ходе работы будут формировать файлы и помещать их в облако, указанное пользователем.

Рассмотрим организацию потока на примере передачи данных из Директа в Я.Облако

1. Прежде всего нужно добавить подключение к Облаку (если оно уже организовано, переходим к следующему шагу)

× **Добавление: Подключение**

Хост хранилища *

Идентификатор ключа, который вы получили при генерации статического ключа *

Секретный ключ, который вы получили при генерации статического ключа *

ⓘ Не забудьте нажать кнопку сохранения, чтобы завершить настройку

Сохранить

Для работы рекомендуем пользоваться инструкцией <https://cloud.yandex.ru/docs/storage/operations/>

Что нужно вводить на этапе подключения:

Хост хранилища - расположение хранилища в сети, его настраиваете при работе с хранилищем и хостингом

Идентификатор ключа - ID ключа шифрования в вашем хранилище. Не сам ключ а именно ID, система по этому ID будет отправлять запрос на работу с данными.

Секретный ключ - Один из ключей шифрования, который генерируется при помощи статического ключа. Нужен для обращения к данным.

2. Выбираем источник и приемник данных, а так же набор данных

✕ Добавление: Подключение

Хост хранилища *

Идентификатор ключа, который вы получили при генерации статического ключа *

Секретный ключ, который вы получили при генерации статического ключа *

❗ Не забудьте нажать кнопку сохранения, чтобы завершить настройку

Сохранить

3. Настраиваем источник данных, указывая необходимый аккаунт для загрузки данных
4. Настраиваем Приемник данных

3 Настройка приемника данных

Выберите аккаунт для загрузки данных или добавьте новый *

 `https://storage.yandexcloud.net`  [+ Добавить](#)

Название бакета (корзины) *

1

Должно быть уникальным при создании или доступным под указанным аккаунтом. Название не должно содержать символ подчеркивания _ . Подробнее: <https://cloud.yandex.ru/docs/storage/concepts/bucket#naming>

☒ Публичный доступ на чтение содержимого бакета

☐ Использовать динамическое название файла

Если настройка активна, то при каждом запуске потока будет создаваться новый файл с выбранным форматом названия файла

Путь к файлу с учетом папок и названия файла *

1

Пример: папка1/папка2/название_файла.расширение_файла

Пресет для обработки и генерации файла *

Выберите вариант...

Время хранения объекта в днях

Требуется целое положительное число больше 0

Далее

- а. Выбираем ранее подключенный аккаунт, либо добавляем новый
 - б. выбираем название Бакета. с этим названием поток создаст Бакет и будет помещать в него файлы
 - в. публичный доступ к Бакету - опциональная настройка, необходима если нужен открытый доступ к бакету
 - г. путь к файлу
 - д. пресет для обработки и генерации файла - определяет в каком формате будет сгенерирован файл
 - е. время хранения объекта в днях(TTL) - опциональная настройка, которая позволяет определить количество дней с момента создания или последнего обновления таблицы, спустя это время таблица будет автоматически удалена.
5. Выбираем общие настройки потока - название, как часто будет запускаться и за какой период генерировать данные