Загрузка данных из BQ в S3

Garpun Feeds может выгружать файлы из Google BigQuery в S3.

S3 — сервис хранения неструктурированных данных. Внутри можно хранить все что захотим, но в основном используется как файловое хранилище. Каждый файл представлен в виде объекта, сами же объекты лежал в бакетах. Про его настройки можно отдельно почитать в документации самого сервиса: https://cloud.google.com/storage/docs/introduction

Перед началом работы нужно добавить подключения к ВQ и S3(если оно уже организовано, переходим к этапу созданию потока):

Авторизовавшись в системе Garpun переходим в раздел "Подключения" (ссылка), выбираем S3, нажимаем "+подключение".

| × Добавление: Подключение |
|--|
| |
| Хост хранилища * |
| Название подключения |
| 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 |

- Хост хранилища расположение хранилища в сети, его настраиваете при работе с хранилищем и хостингом
- Название подключения то, как будет называться наше подключение.
- 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.

Авторизовавшись в системе Garpun переходим в раздел "Подключения" (ссылка), Выбираем Google BigQuery, нажимаем "+подключение".

| × Добав | ление: Подключен | ие | |
|-----------------------------------|-----------------------------------|---------------------|--|
| Google BigQuery * 🗞 Подключить | | | |
| Не забудьте наж. |) ать кнопку сохранения, чтобь | завершить настройку | |
| • Нажимаем | | | |

- Выбираем необходимый аккаунт
- Нажимаем "Разрешить"

1. Источник данных > Приемник данных

В качестве источника выбираем Google BigQuery, в качестве приемника - хранилище S3, в качестве набора данных - "Стандартная выгрузка данных".

| | | | | 🔞 Не нашли нух | жную сис | тему? |
|---|---------------------------|---------------------|----------------------------------|----------------|----------|--------|
| Источник данных * | | | Приемник данных | (* | | |
| Google BigQuery | ~ | > | 💋 S3 | | | \sim |
| | | | | | | |
| Набор данных * | | | | | | |
| Стандартная выгрузка данных | | | | ~ | Q | + Доба |
| СТРОЙКа ИСТОЧНИКА ДАННЫХ Настройка источника данных | | | | | | |
| строика источника данных | | | | | | |
| Выберите аккаунт для получения данных или добавьте н | овый * | | | | | |
| Выберите вариант | | | Q + Добавить | | | |
| Resource concerns | | | | | | |
| Редактировать подключения | | | | | | |
| Project ID B BigQuery * | | | | | | |
| Project ID в BigQuery * Выберите вариант | | | ~ | | | |
| r едах провата гладоноченния Project ID a BigQuery * Выберите вариант Standard SQL Query * | | | ~ | | | |
| Project ID a BigQuery * BudGepure approach Standard SQL Query * 1 | | | ~ | | | |
| Project ID и BigQuery * Bui/Spurte вариант Standard SQL Query * | | | ~ | | | |
| Project ID и BigQuery * Bui/Spurte вариант Standard SQL Query * | | | ~ | | | |
| Project ID и BigQuery * BulSepurte вариант Standard SQL Query * | | | ~ | | | |
| Project ID и BigQuery * BulGepurte вариант Standard SQL Query * | | | × | | | |
| Project ID a BigQuery * BuGepure aaproart Standard SQL Query * 1 | | | ~ | | | |
| Project провеле наурночения Project Dia BigQuery * Выберите вариант Standard SQL Query * 1 | | | ~ | | | |
| r egekinjudel is inglooreenin Project Di a BigQuery * Выберите вариант Standard SQL Query * | | | ~ | | | |
| Project in Judie is Inglowershin Project Di a BigQuery * BidGepurte aapvart Standard SQL Query * 1 | | | ~ | | | |
| r сдех прове в наудночения Project Da BigQuery * Выберите вариант Standard SQL Query * 1 Используйте плейсколдеры \$(feed date_from) и \$(feed date_ | то), чтобы подставить в з | апрос пернод получе | ۷ чия данных | | | |

- Выбираем проект, из которого будем брать данные
- Прописываем SQL запрос

• Используйте плейсхолдеры \${feed.date_from}, \${feed.date_to}, \${feed.datetime_from} и \${feed.datetime_to}, чтобы подставить в запрос период получения данных. Например, WHERE date BETWEEN '\${feed.date_from}' AND '\${feed.date_to}'

3. Настройка приемника данных

| Bregelness services the realities the section to a the particulate control . | | | |
|---|----------------------------|-------------|----------|
| Dudeparte separant | | 0, | + Antone |
| Парвания бакота-(кордины) * | | | |
| | | | |
| Пресет для обработни и темериции фойла * | | | |
| mulieparte napatart | | | |
| Поли презенти выбрана Чанивная выпуска на БрДилеу (Раприе), ВАВНО, ченбы коер ВД на Муть в папка ная файку сучетам папка и названия файка.* | ets görtyn e filletty | | |
| | | | |
| . Проход для файла ната I іната Сілана Старание, файла ракшоренне, файла - Проход Аля натакі пана Unama I Создать содпатка для партиций/шордов | | | |
| Произтру для файлон гланал Патала.СПанаралине, райла ракциортине, райла Произтру для патала Ганала? Сездать подполки для партикуній/шардов Для | | | x v |
| Приня до офика на Пинастранана с дола до со до со обранот, дола приня до на конто на Пинастрани Пинастрани. Соцать подпаться для портичей нардата. Да Пострайка акточна, кото пратота вобрано Патана в котурала за Врублир Улагрит? - пот вобрана 1967 го дола Подпата Парата парата до обработо у Оказуна? - пот вобрана 1967 го дола Подпата Парата парата до портичуба нардата. | | | x v |
| Принят для прибля каная Гланизаться, райла датамрение, файле Принят для таких налаги бланай Соцать Боданиев, для зарачайной бланадар И Постоя Парила Салонан, как практом на переко Четонана для обремор Гатурия - постоя парила 25 со праконствико II друго парактор (16), учите у налагио на переко - постоя Приня Сарас и праконствико II друго парактор (16), учите у налагио на - Постоя Приня Сарас и отноже сарареность бале и - Постояние Принят отноже сарареность бале и | | | x v |
| Tyronni gran dalor nimi Turusztismanne, jakita garsunytene, jakita Tomany na nimi natita tihandi Catarin kogan kapita dalami Interplanta antona ncon reporteri andjalori Yitenisan nuryyosi ni BigDury Pirepert Interplanta antona ncon reporteri andjalori Yitenisan nuryyosi ni BigDury Pirepert Interplanta antona kapita dalami daga dalami dalami dalami daga dalami Interplanta antona daga dalami daga dalami daga dalami daga dalami Interplanta dalami dalami daga dalami daga dalami daga dalami Interplanta dalami dalami daga neuro Yitena antopia ni BigDury (Panyut) | | | X V |
| Приету для прило калена Гоника Толикански, райла дотимурном, файло Приету для полно калей Ульника? Стадать быратом для партирия бырадора. До Полнованијана: До са протогото на фарано Толико-то калуузана на барбину Рагириет? - пол вибрита: До са преколотото на фарано Толико-то Калуу полу улатаничу Вилика - пол вибрита: До са преколотото на фарано Толико, у Калуузана на барбину Рагириет? - пол вибрита: До са преколототото Калуузана са Барбину Рагириет? - Пол вибрита: До са полотототото Калика - Полотокалема, настоя калуузана на барбину Рагириет? - Полотокалем, настоя калуузана до калика. | | | X 9 |
| Принят для прило сління Тинира Тананансь, райна доконренно, райно Принят для пакино клічніка Стадать Каралина, рак партиранічноградня Ла Постанів Каналина, кото протото набулької Наганова пакурова на Проблем УШ-трибут паки пакурова VIII. Та філіон Бару со параваться Каралина для партирай накада на пакурова VIII. Та філіон Бару со параваться Каралину Панадан Проблемний філіор VIII. Та філіон Бару со параваться Каралину Панадан На законна, кото паравел притото Карали на Баруличу Митриної. — Полонаціона да укали карали на такурова на Табрану УМитриної. — Полонаціона да укаливанськи Калалина філіона Каралину Панадан. — Полонаціона да укаливськи Калалини філіона Параличу Панадан. | Dennengen Østilmetasione | - 104 (1-1) | X V |
| The energy and plant a result in security distance and plant approximately plant approximation of the energy of t | teroren (stimtater | - 130,0-0 | XV |
| . Приняд для пробле літина Гоннов Тоннован, дойла дополорение, дойло Товану для на полі палки Тоннова Стадать барганих дага пратанов надуран Натоння палурана на бірбану Улагриятії полі падіяних дага пратанов надуран Чатоння палурана на бірбану Улагриятії полі падіяно Чата са пратанов надурани Чатоння палурана на бірбану Улагриятії на падіяно Чата са пратанов надурання Тоннана палурана на бірбану Улагриятії На палонана, коло падіян падато Гангена Барани (Чанарикт) Полінованае дирого малон Каланова Барани Полінованае дирого малон Каланова Бараних Полінованае дирого малон Каланованих Полінованае дирого малон Каланования Саланования Полінованае дирого малон Каланования Саланования Саланования Полінованае дирого малон Каланования Саланования Полінованае дирого малон Каланования Саланования Полінованае дирого малон Каланования Саланования Полінования дирого малон Каланования Саланования Саланования Саланования Полінования дирого малон Каланования Саланования Саланования Полінования Саланования С | beerne (slostson | - []]] (0-0 | X V |
| (Figure 2) provide a series of transmission, planta persuspense, planta (Figure 2) persons and value of based 2. Charts to strands, person stype performance on any person to BigDury (Figure 2) interplant accesses, etc. in spectrum and plant (First-in-or any person BigDury (Figure 2) interplant accesses, and spectrum stype accesses and spectrum stype accesses ac | ренный фарматом н | 100 | × v |
| Thereary particular series of human and participance, patient generative equipance areas for segment and human and a series of the series of t | berren gabetabe | - 120 00-0 | × × |
| If priority provide intervent from the formation of pathole pressupport on the development, dashes Toporty para research ender Namedal Charante dashes statement and pathole Thread-in an anyyora so BigDoury Brogney Interpolar access, state repertures and pathole Thread-in anyyora so BigDoury Brogney Interpolar access and the spectrum statement and the statement a | Denorities & Distantia Den | - []]] []-0 | × × |

- Выбираем аккаунт для загрузки данных
- Название бакета. Инструкция по неймингу бакета для GCS: https://cloud.google.com/storage/docs/buckets#naming, для Yandex Object Storage - https://cloud.yandex.ru/docs/storage/concepts/bucket#naming
- Пресет для обработки и генерации файла. Если пресетом выбрано 'Нативная выгрузка из BigQuery (Parquet)', важно, чтобы юзер BQ имел доступ к бакету
- Путь к файлу

- Создавать ли подпапки для партиций. При использовании опции в указанной папке будут созданы подпапки, разбитые по датам.
- "Использовать динамическое название файла". При включенной опции каждый запуск потока будет создавать в приемнике

| Н | овый файл по выбранному шаблону Выберите формат динамического названия файла * | |
|---|---|---|
| | Добавить к названию файла префикс с диапазоном сбора данных.[диапозон сбора]_[название] | ~ |
| | Добавить к названию файла префикс с диапазоном сбора данных.[диапозон сбора]_[название] | |
| | Добавить к названию файл префикс с датой сбора данных.[дата запуска]_[название] | |

○ Время хранения в днях.

4. Общие настройки

На этом этапе вам необходимо изменить название потока если необходимо. Название потока генерируется автоматически.

Выбрать период сбора при автоматическом перезапуске. По умолчанию устанавливается "на основе внутренних правил", что означает, что пересбор потока будет происходить за последнии 30 дней.

Установить расписание

Нажать "Готово"