Arenadata™ Grid

Версия - v2.3.1-nsud

Процессор обработки входных данных

2020 Arenadata™

Оглавление

1 Публичные методы **3**
1.1 insert_messages_from_kafka . . . 3
1.2 load_csv_lines . . . 3
1.3 get_metric . . . 4

2 Приватные методы **5**
2.1 parse_avro . . . 5
2.2 load_avro . . . 5
2.3 prepare_kafka_message_for_insert . . . 5
2.4 get_function_by_name . . . 6

Используется для преобразования и обработки данных, приходящих из внешних систем.

Глава 1

Публичные методы

1.1 insert_messages_from_kafka

Метод, предназначенный для вставки данных из **Kafka** в спейсы роли *storage*. Требует для работы наличие `schema_registry`, которое предоставляет схемы *avro*.

Параметры:

- `messages` – массив сообщений Kafka, которые представляют собой lua-таблицу структуры:

```
{
   "topic":"имя топика",
   "partition":"номер партиции",
   "offset":"оффсет",
   "key":"ключ",
   "value":"значение"
}
```

- `parse_key_function_str` – имя метода, использующегося для парсинга ключа;
- `parse_value_function_str` – имя метода, использующегося для парсинга значений.

Возвращаемое значение: Возвращает результат обработки данных из Kafka. Выходной формат:

```
{
    "topic":"имя топика",
    "partition":"номер партиции",
    "offset":"оффсет",
    "key":"ключ",
    "result": "результат операции true/false",
    "error": "сообщение об ошибке или nil"
 }
```

Конфигурация:

Использует параметры из настроек топиков кафки и настроек schema_registry.

1.2 load_csv_lines

Метод, предназначенный для вставки строк *csv* в спейсы роли *storage*.

Параметры:

- `space_name` – имя спейса, куда производится вставка;
- `lines` – массив строк csv для вставки.

  Возвращаемое значение:

- *true* – при успешной вставке;
- *false,error* – при ошибке.

1.3 get_metric

Метод, предназначенный для передачи метрик данной роли в систему сбора метрик.

Возвращаемое значение: Набор метрик в формате *JSON*.

Глава 2

Приватные методы

2.1 parse_avro

Метод, предназначенный для разбора *avro* и преобразования его в формат, необходимый для вставки в *storage*.

Параметры:

- `schema` – имя схемы для запроса в schema_registry;
- `value` – значение для вставки в формате avro.

Возвращаемое значение:

- *true* – при успешной загрузке;
- *false,error* – при ошибке.

2.2 load_avro

Метод, предназначенный для загрузки данных из `parse_avro` в *storage*.

Параметры:

- `space_name` – имя спейса, куда производится вставка строк;
- `lines` – строки, которые необходимо вставить.

Возвращаемое значение:

- *true,разобранные данные* – при успешном разборе;
- *false,error* – при ошибке.

2.3 prepare_kafka_message_for_insert

Метод, предназначенный для предварительного разбора данных из `insert_messages_from_kafka`.

Параметры:

- `topic` – имя топика Kafka, откуда пришли данные;
- `data` – данные для вставки в *storage*;
- `parse_function` – функция для разбора данных.

Возвращаемое значение:

- *true,разобранные данные* – при успешном разборе;
- *false,error* – при ошибке.

2.4 get_function_by_name

Метод, предназначенный для возврата функции по ее имени.

Параметры:

- `function_name` – имя функции для поиска в скоупе;

Возвращаемое значение: Функция или nil.