

0.1. Рудов М.С., Харлампенков И.Е., Потапов В.П. Поточковая обработка данных с датчиков IoT

Решение задачи контроля концентраций пылевого аэрозоля на территории предприятия требует непрерывного мониторинга в автоматическом режиме. Измерительное устройство регулярно отправляет данные IoT. Принимающая сторона должна работать в режиме постоянной готовности, поэтому возникает потребность в системе потокового сбора данных.

На данный момент нет полностью готового средства для получения данных с датчика IoT, но его можно реализовать на базе ETL (extract-transform-load) решений из числа Apache Airflow, Apache NiFi, StreamSets и т.д. [1] Важно отметить, что создаваемая процедура обработки данных должна быть ориентирована на поддержку нескольких типов датчиков и поддерживать преобразование данных к общему для системы формату.

Выбран Apache NiFi [2], так как он позволяет запускать процедуры обработки не только по расписанию, но и по факту получения данных (поточковый режим). К плюсам Apache NiFi относится и возможность в графической форме быстро создать ETL-процесс, множество готовых процессоров (получение данных по протоколам MQTT, S7, PLC4X и т.д. [3]), отправка в очереди сообщений, преобразования данных, выполнение запросов к в базе данных), экспорт и импорта схемы решения, возможность реализовать свой процессор на Java.

Схема работы представлена следующим образом. Замеры принимаются по протоколу MQTT (ConsumeMQTT), данные конвертируются в JSON (ConvertRecord), из JSON по Jolt спецификации выделена и приведена к удобному виду именно та часть, которая нам необходима (Jolt TransformJSON). В зависимости от типа пришедшего сообщения данные отправляются в очередь сообщений AMQP, сообщения, которые не прошли проверку по типу, удаляются. Данные сохраняются в базу данных и используются для дальнейшей обработки в системе.

В результате с помощью ETL-средства Apache NiFi реализован приём данных с датчика IoT по протоколу MQTT, данные преобразованы к требуемому виду и отправлены в очередь сообщений.

Список литературы

- [1] ARTEM VARMIN Top 5 Enterprise ETL tools. [Электронный ресурс]. URL: <https://freshcodeit.com/freshcode-post/top-5-enterprise-etl-tools> (дата обращения 20.08.2022).
- [2] Apache NiFi. [Электронный ресурс]. URL: <https://nifi.apache.org/docs.html> (дата обращения 25.08.2022)
- [3] Language support for protocols. [Электронный ресурс]. URL: <https://plc4x.apache.org/users/protocols> (дата обращения 25.08.2022)