***Д.С. Молчанов, студ.; рук. М.А. Орлова, к.т.н., доц. (НИУ «МЭИ»)***

**ЭФФЕКТИВНАЯ СЕРИАЛИЗАЦИЯ В БАЗУ ДАННЫХ КОРПУСОВ ДАННЫХ, КОТОРЫЕ ИМЕЮТ НЕПОСТОЯННЫЙ ФОРМАТ**

Молчанов Д.С,. Орлова М.А.

Сериализация в базу данных является неотъемлемой частью анализа больших объемов информации. От скорости и оптимизации этого этапа зависит общая эффективность накопления и анализа данных.

Современные технологии предлагают большой выбор готовых решений для использования базы данных для хранения той или иной информации. Однако для конкретной задачи возникают специфические условия, которые накладывают ограничения на выбор и требуют отдельной реализации.

В рамках этой работы рассмотрены варианты использования как NoSQL [1-2], так и SQL [3] решений. При выборе базы данных ключевыми аспектами являлись необходимость универсального хранения корпусов данных с любым количеством столбцов (при условии использования определенного собирателя трафика), масштабируемость в перспективе и стек используемых технологий (язык программирования высокого уровня Go). В итоге разработан подход к сериализации трасс сетевого трафика, которые хранят в себе информацию о протоколах, которые использовались устройством во время сбора данных.

**Литература**

1. **Namiot D.E.** Time Series Databases// Program. Comput. Softw. – New York, 2022. – Vol. 48, № 7. – P. 413–423. Conference: XVII International Conference «Data Analytics and Management in Data Intensive Domains» (DAMDID/RCDL’2015), Obninsk, Russia, 2015

2. **Команда Yandex Cloud**. NoSQL: виды, особенности и применение [Электронный ресурс] ООО «Яндекс.Облако» – 28.10.2022 Режим доступа: https://cloud.yandex.ru/blog/posts/2022/10/nosql, свободный. – Загл. с экрана.

3. **Конференции Олега Бунина (Онтико)**. Time series данные в реляционной СУБД. Расширения TimescaleDB и PipelineDB для PostgreSQL [Электронный ресурс] Habr – 28.09.2019 Режим доступа: https://habr.com/ru/companies/oleg-bunin/articles/464303/, свободный. – Загл. с экрана.