

Анализ возможностей применения решения SAP HANA для оптимизации транспортной ситуации в крупном городе

Цыбульский Анатолий Сергеевич

Российский экономический университет имени Г.В.Плеханова

Студент

Родионова Дарья Фёдоровна

Российский экономический университет имени Г.В.Плеханова

студент

Аннотация

Объемы информации растут с каждым годом в несколько раз, появляются новые решения для хранения и анализа больших данных в режиме реального времени. В статье рассмотрена платформа SAP HANA, ее основные характеристики, а также процесс анализа больших данных платформой. Определены преимущества использования SAP HANA для транспортной инфраструктуры и выявлены основные положительные результаты, к которым приводит внедрение платформы.

Ключевые слова: SAP HANA, большие данные, анализ данных, транспортная ситуация.

Analysis of the possibilities to use SAP HANA for optimizing the transport situation in a large city

Tsybulskiy Anatoliy Sergeevich

Plekhanov Russian University of Economics

Student

Rodionova Daria Fedorovna

Plekhanov Russian University of Economics

Student

Abstract

Data volumes are increasing every year, new solutions for big data analysis and storage in real lifetime are developed. The article reviews SAP HANA platform, its key functions and the process of big data analysis. Reasons and favorable results of SAP HANA installation in transportation system are described in the article.

Keywords: SAP HANA, big data, data analysis, city transport, transportation system.

В разных отраслях объем и сложность создаваемых данных увеличивается с огромными темпами. Такая тенденция свойственна не

только наукоемким отраслям, но и низко технологичным, традиционным отраслям. Закономерно возникают проблемы хранения и анализа быстро увеличивающихся в размерах наборов данных. Традиционные технологии обработки данных делают расход человеко-часов и вычислительных ресурсов на анализ неэффективным. С одной стороны, существующие аналитические технологии, как например проект Hadoop, позволяют фильтровать, очищать и обрабатывать петабайты неструктурированных данных [1]. С другой стороны, способность существующих технологий выполнять аналитику для получения точных и своевременных результатов обработки информации не успевает за ростом объема и сложности новых данных.

Работа с огромными объемами информации создает не только определенные трудности, но и возможности, в первую очередь – поиск неочевидных закономерностей и взаимосвязей в данных.

Реализовать возможности и справиться с трудностями предлагается с использованием хранения и вычислений в оперативной памяти – так называемая технология вычислений in-memory [2]. Технология реализует масштабируемость вычислительных ресурсов в режиме реального времени, а также возможность выполнять визуализацию данных, специальный анализ и прогнозируемую аналитику структурированных данных. Использование вычислений в оперативной памяти, приводит к сокращению времени, необходимого для обработки больших данных, а также к обеспечению быстрого времени отклика, и достаточной глубине анализа и упрощению связанных бизнес-процессов [3].

Современным продуктом для анализа больших данных является платформа компании SAP – S/4HANA, использующая технологию In-memory. Эту платформу можно упрощенно описать как реляционную СУБД, размещающую базу данных по возможности целиком в оперативной памяти, благодаря чему достигается минимальное время отклика для выполняемых запросов к базе данных. Такой подход позволяет решать задачи обработки данных практически в реальном времени, используя встроенные OLAP-процессоры.

Получать информацию на обработку платформа способна от ERP-систем, от популярных СУБД (MySQL, Oracle, SQL server), таблицами в формате csv или XML-файлами. Платформа передает результаты обработки и анализа ETL-средствам и инструментам визуализации [4].

В общем виде процесс работы с большими данными можно описать в несколько этапов:

- Получение данных от различных источников
- Первичная обработка и очистка неструктурированных данных
- Структурирование данных
- Агрегация и первичный анализ данных
- Полноценный анализ под конкретные задачи
- Представление и визуализация данных в реальном времени.

Платформа SAP HANA может быть использована в целях анализа и улучшения транспортной инфраструктуры крупного города. Программный комплекс позволяет анализировать текущую ситуацию на дорогах, прогнозировать загруженность основных автомобильных магистралей, а также оценивать эффект от реализации дорожных и градостроительных проектов.

В первую очередь, платформа предоставляет доступ к данным, собираемым камерами, датчиками и прочими системами трекинга и мониторинга, в реальном времени, о ситуации на дорогах водителям и людям, пользующимся общественным транспортом. Таким образом, пользователи получают, к примеру, персонализированную информацию для оптимального выбора маршрута, основанную на их местоположении, и имеют возможность оценить временные затраты на дорогу. Прогнозирование с использованием имитационного моделирования и машинного обучения позволяет определить развитие ситуации на дорогах для составления рекомендаций участникам движения. Работа с данными в реальном времени позволяет предугадывать и предотвращать проблемы и аварии до момента их появления за счет гибкого изменения ограничений скорости на дорогах, управления светофорами и публикации рекомендаций на информационных стендах, установленных вдоль движения транспорта.

Сбор статистической информации, относящейся к перемещению транспортных потоков, позволяет проводить долгосрочный ретроспективный анализ. Такой анализ транспортной инфраструктуры позволяет выявить участки дорог, нуждающиеся в реконструкции, перепланировке или изменении правил дорожного движения.

Моделирование изменения транспортной системы после внедрения проекта или закрытия участка автомобильной дороги определяет нагрузку на магистрали и определяет возможные способы их разгрузки. Задача моделирования, основанного на больших вычислительных мощностях платформы SAP HANA заключается в выявлении последствий изменения некоторых отрезков магистралей [6].

С помощью специальных датчиков представляется возможным в реальном времени отслеживать доступность парковочных мест и формировать карту «свободных мест для парковки». Это позволит оценивать и контролировать загруженность городских парковочных зон — очень актуальная тема для городского исторического центра с узкими дорогами.

Используя аналитику транспортной ситуации в городе, городским жителям возможно предоставить ежеминутную сводку о лучших маршрутах и местах для занятий спортом, пробежек. В перспективе, возможна разработка оптимальных маршрутов, свободных от опасностей, связанных с автомобильным транспортом. Что в конечном итоге приведет к качественным улучшениям уровня жизни в городе — эффект на сферы, которые непосредственно не контактируют с прикладной областью, в которую внедряется решение от SAP [5,7,8].

Существует возможность агрегации данных из различных источников, в том числе данных о погодных условиях и расписании мероприятий в городе, и их анализ в реальном времени с помощью вычислительного движка In-Memory, который позволяет отражать полную картину сложившейся ситуации, а использование открытого API позволяет использовать эту информацию сторонним приложениям и разработчикам в своих проектах.

Использование SAP HANA для анализа и улучшения транспортной инфраструктуры может привести к следующим результатам: ежегодное снижение расходов на транспорт, сокращение среднего времени на проезд на 25-30%, увеличение средней скорости движения в городе, снижение расхода топлива, уменьшение количества аварий на 50% в связи с подбором оптимального скоростного режима на дорогах.

На рис.1 представлена схема работы платформы In-Memory SAP HANA (рис.1).

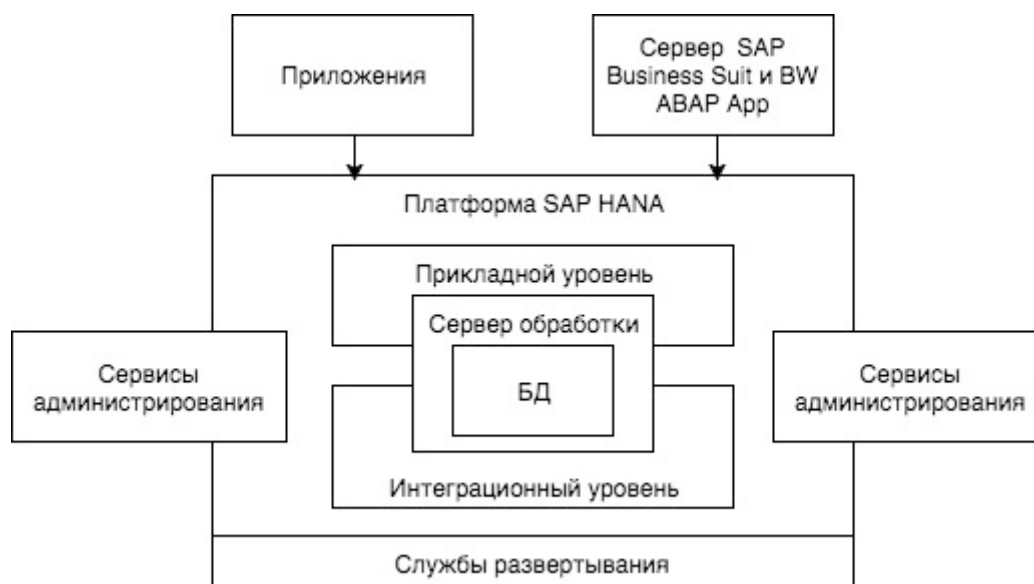


Рисунок 1 - Схема работы платформы In-Memory SAP HANA

Таким образом, можно сделать вывод о том, что использование вычислительных мощностей платформы SAP HANA имеет большие перспективы для развития. В частности, ее внедрение в транспортную систему больших городов может привести к значительным положительным результатам в достаточно короткие сроки.

Библиографический список

1. Hadoop Illuminated. Mark Kerzner and Sujee Maniyam, 2013
2. What is SAP HANA? // Wikibin.org URL: http://wikibon.org/wiki/v/Primer_on_SAP_HANA (дата обращения: 19.05.2017).
3. Plattner H.; Zeier A. In-Memory Data Management: Technology and Applications. Springer Science & Business Media, 2012.

4. Грушин. Н. Построение системы отчетности на базе SAP HANA// URL:<http://www.returnonintelligence.ru/wp-content/uploads/2015/02/SAP-Hana-Reporting.pdf> (дата обращения: 19.05.2017).
5. Попов А. А. Алгоритм формирования набора информационных систем для автоматизации жилищно-коммунального хозяйства региона, Электронный журнал «Известия РЭУ имени Г.В. Плеханова», №3 (17), 2014/ URL: http://www.rea.ru/Main.aspx?page=Nomer_3__17_
6. SAP HANA as a product // Sap.com URL: <https://www.sap.com/products/hana.html> (дата обращения: 19.05.2017).
7. Кокоулина О.П. Двигательная активность как важная составляющая здорового образа жизни студентов // Инновационное развитие российской экономики - международная научно-практическая конференция. Москва - 2016.
8. Кокоулина О.П. Анализ образа жизни и занятий физической культурой и спортом студенческой молодежи // Статистика и экономика Российский экономический университет имени Г.В. Плеханова. М., 2016.