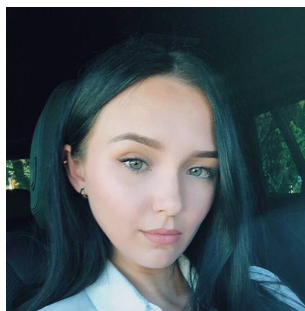


12 июня 2022

Как мы используем Luxms BI — российское ПО для аналитических отчетов



Полина Тесленко,
разработчик ГК «КОРУС
Консалтинг»

Тема импортозамещения последние месяцы не сходит со страниц изданий. Очевидно, что в области работы с данными компании также активно ищут новые возможности по использованию продуктов, которые не будут заблокированы в России, а производитель продолжит оказывать техническую поддержку. В этой статье я хочу поделиться практическим опытом работы с импортозамещающим ПО — Luxms BI.

Luxms BI – это платформа визуального контроля и аналитики, которая автоматически собирает, обрабатывает и визуализирует данные для отчётов.

Мы делаем проекты на этой системе уже больше трех лет. Их стек технологий воплотил датацентричный подход и уникальную методику подготовки данных, одновременно дает нам высокую скорость доступа к данным при разработке и возможности кастомизации.

Она состоит из нескольких компонентов. Поговорим сегодня о двух – импортере и оптимизационном модуле для загрузки данных в BI.

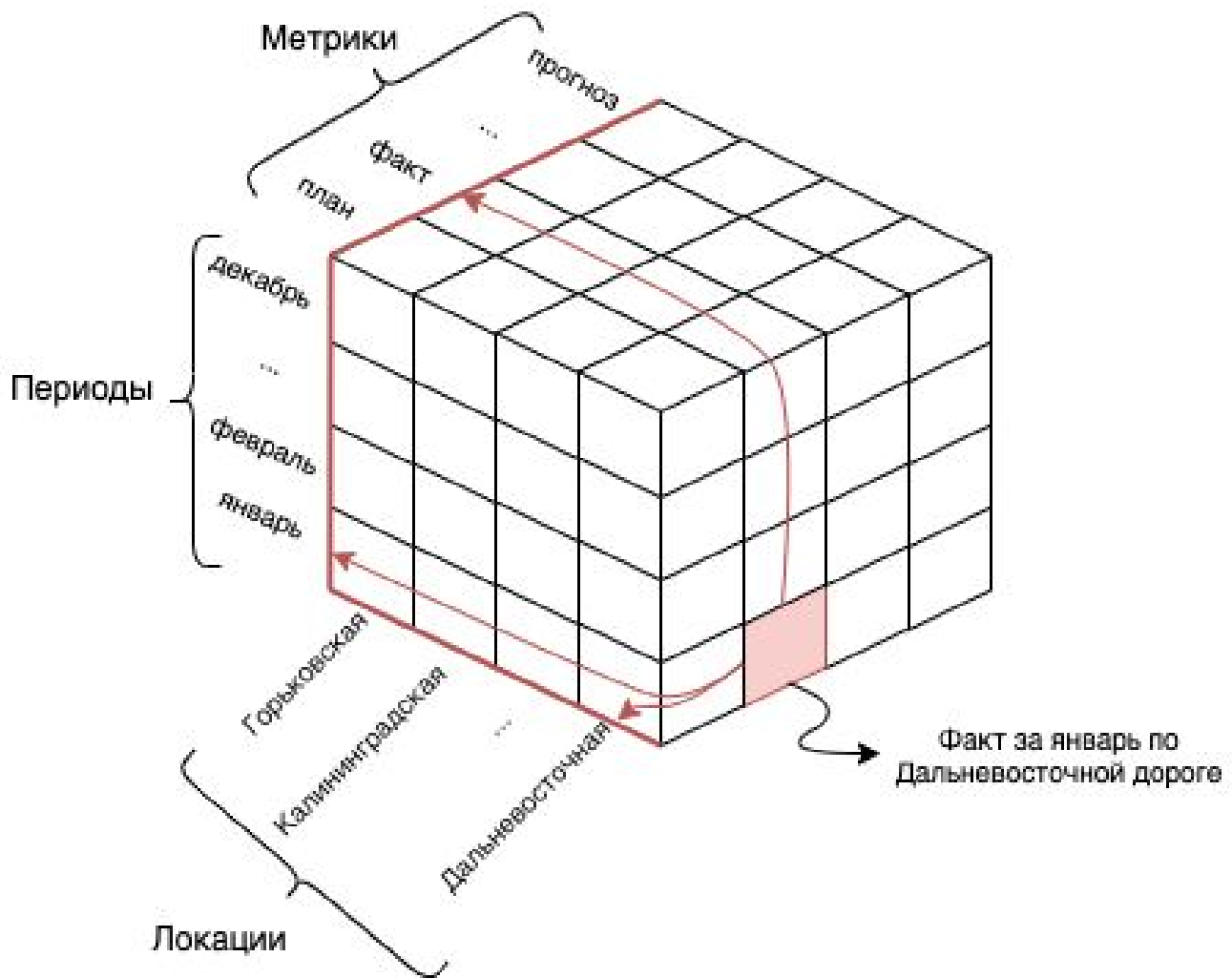
ИМПОРТЕР

Компонент импорта Luxms BI решает задачи по загрузке данных из внешних источников и является реализацией ETL-стека для их агрегации и трансформации.

Источником данных может являться любое хранилище, способное отдавать данные в виде плоских таблиц, поэтому для наиболее гибкого взаимодействия мы работаем с реляционными базами данных посредством SQL.

Импортер преобразует данные из плоской таблицы в трехмерный [OLAP](#) куб:





Для того, чтобы отобразить необходимые данные в целостный показатель платформы — датасет, в импортере хранятся так называемые батчи, которые представляют собой скрипт из SQL запроса, объединяющий набор метрик, периодов и локаций, а также настройки маппинга с применением процедур фильтрации, валидации и других преобразований, разделенных индивидуально для свойств каждого датасета.

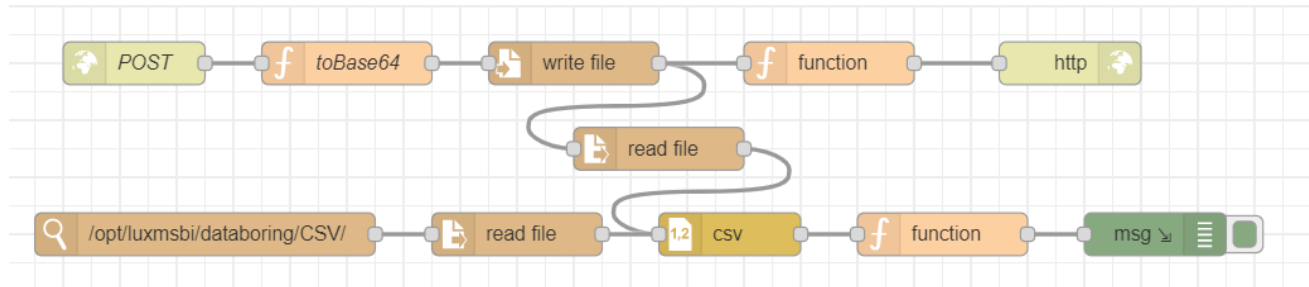
Результат выполнения SQL запроса берет данные из источника и помещает их в датафрейм-контейнер в виде плоской таблицы.

В итоге при общей загрузке импортера производятся агрегирующие операции над датафреймом и затем сохраняются в датасет, что в последствии позволяет отобразить визуальные данные на платформе Luxms BI.

LUXMS DATA BORING

Luxms Data Boring – это инструмент программирования на основе потоков. Мы используем его на уровне backend-части. Его главная задача — оптимизировать процесс загрузки данных в систему.

Потоковое программирование, представляет собой способ описания поведения приложения как «узлов», связанных между собой логической последовательностью. Каждый узел имеет четко определенную цель. Ему на вход даются определенные данные либо команды, он проводит заданные операции, а затем передает обработанные «сообщения» дальше по связям в другие узлы. Каждый узел принимает сообщение с данными через абстрактный интерфейс — порт, который работает так же, как сетевой порт в компьютерной сети. Данные отправляются на него через буфер, имеющий ограниченный размер. Один порт может относиться к нескольким экземплярам одного компонента, что упрощает использование структуры в распределенной системе или для параллельной обработки.



Большая часть кода инкапсулирована, то есть, изолирована на подобие «черного ящика», о внутренних процессах которого можно судить по его реакциям на внешнее воздействие. Поэтому исходный код имеет высокий потенциал повторного использования. Это также означает, что обновление или масштабирование приложения на базе потокового программирования может быть проще, чем при использовании более интегрированного приложения: систему обмена сообщениями, модули и систему портов можно изменять независимо, не затрагивая всю программу.

Для взаимодействия с Data Boring в качестве пользовательского интерфейса рабочая панель открывается в браузере, где можно построить необходимую цепочку узлов по принципу визуального конструктора, перетаскивая их из палитры готовых шаблонов в рабочее пространство и соединяя их вместе. Также есть возможность расширить палитру доступных модулей или доработать готовые путем добавления конечной и четко формируемой функцией, с входным/выходным портом. Логика работы узлов может быть реализована как на JavaScript, так и на других языках. Одним щелчком мыши приложение развертывается на сервер в среду выполнения, где запускает back-end часть для работы платформы.

Благодаря возможности параллельной загрузке данных и упрощенному операционному процессу, Data Boring оптимизирует работу с BI-платформой.

Визуально компонент довольно изящен и интуитивно понятен. Мое изучение потенциала Data Boring еще продолжается, но я уже оценила его универсальность и возможности, которые на самом деле намного обширнее, что вполне может определять его как мультизадачный компонент.

Работа с компонентами импорта данных Luxms BI продолжается, процесс взаимодействия и совместной разработки активно развивается, модернизируя и расширяя стек технологий, что является неотъемлемым показателем прогресса импортозамещения. Думаю, с нынешней динамикой ассортимент платформы пополнится новыми решениями и инструментами, с которыми каждый сможет ознакомиться лично.