

УДК 004.6

РАЗРАБОТКА АРХИТЕКТУРЫ СИСТЕМЫ МОНИТОРИНГА БИЗНЕС-ПРОЦЕССОВ НЕФТЕДОБЫЧИ В СИСТЕМАХ BUSINESS INTELLIGENCE

Процик И.И.

(Федеральное государственное автономное образовательное учреждение высшего образования «Национальный исследовательский университет ИТМО»)

Научный руководитель – к.техн.н. Иванов С.В.

(Федеральное государственное автономное образовательное учреждение высшего образования «Национальный исследовательский университет ИТМО»)

В работе рассмотрена проблема интеграции корпоративных информационных систем крупной нефтегазовой компании. Предложен выбор использования Business Intelligence системы QlikView в качестве инструмента построения ETL-процесса (процесс экстракции – выгрузки данных из источника, трансформации, загрузки данных), а также спроектирована многоуровневая архитектура данных и приложений процесса обработки данных.

Во многих крупных организациях находится в эксплуатации не одна информационная система, каждая из которых зачастую имеет узкую направленность, как система ведения планово-экономических данных или система мониторинга оперативных данных. Не являются исключениями крупные нефтегазовые предприятия, как «Газпром нефть», являющиеся владельцами терабайтов данных различного вида: от поиска и разработки месторождений до реализации готовых нефтепродуктов. Для таких компаний важным является высокий уровень качества данных, от чего зависит правильность принятия стратегически важных решений. Проанализированы проблемы, возникающие с источниками данных, как то, что методы сбора данных часто плохо контролируются, что приводит к появлению недопустимых значений, что сводит на нет преимущества даже продвинутых методов анализа данных.

Информационные системы могут быть использованы как для проверки, так и для дополнения друг друга. Так, на основе данных из нескольких систем строится отчетность, например, для сравнения оперативных и плановых ключевых показателей, при этом зачастую компаниям требуется иметь не только статичные значения в документе, но и возможности динамически анализировать данные в интерактивном формате. Множество видов отчетности в компании строится путем ручных выгрузок Excel-файлов из разных систем для дальнейшего ручного сопоставления данных и построения отчетности. Практика использования Excel-файлов для сопоставления данных нескольких систем трудозатратна, а также не исключен риск ошибок из-за человеческого фактора. При этом, не все системы поддерживают выгрузку данных в нужных разрезах, а выгрузка большого объема данных и построение необходимых разрезов в Excel – ресурсоемкая задача, с которой зачастую не могут справиться компьютеры бизнес-пользователей. Учитывая вышеуказанные факторы, корректность и скорость подготовки отчетности негативно влияют на качество принятия решений. Построение описанной отчетности требует, во-первых, консолидации данных. Во-вторых, необходим инструмент, позволяющий «на лету» изменять разрез данных. Данный функционал может быть реализован с помощью Business Intelligence системы.

В результате проведенного сравнения BI-систем по нескольким критериям (функционал ETL, возможность развернуть сервер в корпоративной сети, поддержка встроенной аналитики в корпоративные сайты) для решения задачи была выбрана система QlikView, которая имеет наиболее продвинутые возможности ETL. Нами была разработана функциональная многослойная архитектура, которая позволяет переиспользовать данные из систем для различных конечных приложений, представляющих собой интерактивные отчеты в форме таблиц и диаграмм с возможностью выгрузки статичных отчетов.

В процессе разработки ETL-процесса, были реализованы приложения-экстракторы, которые формируют первоначальный слой qvd-файлов (формат файлов QlikView Data), хранящий в себе сырые данные; приложения-трансформы, которые производят предварительную и основную обработку данных; приложение-модель данных, которое

объединят все необходимые данные и формирует итоговую модель, предназначенную для загрузки в итоговое приложение; приложение-отчет, включающее объекты навигации и визуализации, сгруппированные по листам.

Итоговое приложение может быть реализовано несколькими способами: с помощью основной системы QlikView, с помощью базовых визуализаций Qlik Sense, с помощью разработки mash-up'ов на базе Qlik Sense, которые реализуются на технологиях front-end разработки – HTML, CSS, JavaScript с возможностью использовать различные библиотеки и фреймворки. В работе проанализированы преимущества и недостатки применения различных подходов к разработке пользовательского интерфейса на системе Qlik. На текущем этапе принято решение визуализации с использованием системы QlikView как единого инструмента для построения ETL-процесса и визуализации в связи с меньшей оценкой трудозатрат и достаточными возможностями в повторении макетов из Figma – сервиса для разработки интерфейсов, используемого UI/UX-дизайнерами. В процессе реализации пользовательского интерфейса разработан подход к визуализации, основанный на создании алгоритмов, позволяющих строить ранее недоступные в QlikView диаграммы, и разработан алгоритм построения информационной диаграммы с перечнем бизнес-юнитов в разрезе этапов бизнес-процесса компании.

Рассматривая результаты проведенного исследования, можно отметить, что Business Intelligence системы позволяют решать большой спектр задач от хранения и обработки данных до статичных и интерактивных отчетов. Одним из наиболее оптимальных вариантов такой системы является набор продуктов от Qlik – QlikView и Qlik Sense, которые можно использовать независимо и в совокупности. В процессе была спроектирована функциональная архитектура, разработан подход к визуализации, а также описан и реализован алгоритм в рамках указанного подхода, что может быть применено в коммерческой и научной деятельности. В дальнейшем возможно рассмотреть интеграцию BI-систем с инструментами машинного обучения.

Процик И.И. (автор)

Подпись

Иванов С.В. (научный руководитель)

Подпись