

АНАЛИЗ ОТЧЕТНОСТЕЙ В АВТОМАТИЗИРОВАННЫХ СИСТЕМАХ



А.А. Навроцкий¹
Заведующий кафедрой
информационных технологий
автоматизированных
систем БГУИР, кандидат
физико-математических
наук, доцент



Р.В. Притула²
Директор департамента
по работе с
проблемными активами
розничного бизнеса ЗАО
«Альфа-Банк»



Н.В. Протченко¹
Студентка БГУИР,
специалист управле-
ния разработки биз-
нес-решений ЗАО
"Альфа-Банк"

¹Белорусский государственный университет информатики и радиоэлектроники, Республика Беларусь

E-mail: nadia.protchenko@gmail.com

²ЗАО «Альфа-Банк», Республика Беларусь

Abstract. Here was discussed issues of increasing the efficiency of the analysis of large amounts of data in the enterprises. Suggest option of a phased solution of data analysis problems. I has produced comparison of analytical models of information processing for analyzing the activity of the enterprise and selection of the optimal solutions

В настоящее время все острее становится проблема обработки накопленной за предыдущие годы информации. В течение многих лет на предприятиях накапливаются и хранятся огромнейшие массивы информации, но при этом ее большая часть не может быть использована аналитиками и руководителями. Зачастую применяемые технические средства оказываются слишком сложными либо вообще отсутствуют, что утяжеляет процесс проведения успешного анализа информации, наибольший результат можно достичь, лишь внедрив BI (Business intelligence) систему, позволяющую пользователям:

- задавать вопросы и получать ответы, без использования сложных SQL-запросов и посредников из ИТ-службы; добавлять новые источники данных в любой комбинации, необходимой для выполнения текущей задачи;
- быстро получать ответы;
- самостоятельно, без помощи ИТ-службы, изменять панель в процессе работы;
- использовать панель совместно с коллегами.

Всем этим требованиям удовлетворяет QlikView — аналитическая платформа, реализующая ассоциативную архитектуру с обработкой данных в оперативной памяти, позволяющая пользователям, не имеющим специальных технических навыков, самостоятельно формировать отчеты и анализировать

информацию о деятельности своего предприятия в самых разных разрезах, независимо от того, в каких учетных системах предприятия хранится эта информация (1С, Ахарт, Access, Excel, Oracle, SAP, SQL Server и т.д.).

Программный комплекс управляет сложными взаимосвязями между пакетами данных не на прикладном уровне, а на уровне внутренних инструментов платформы. С его помощью можно сохранять в оперативной памяти отдельные таблицы баз данных и манипулировать ассоциативными связями между ними. Любое из значений отдельного поля связано со всеми другими значениями, присутствующими во всей базе данных. При этом состав наборов данных может включать в себя тысячи таблиц, содержащих сотни тысяч полей.

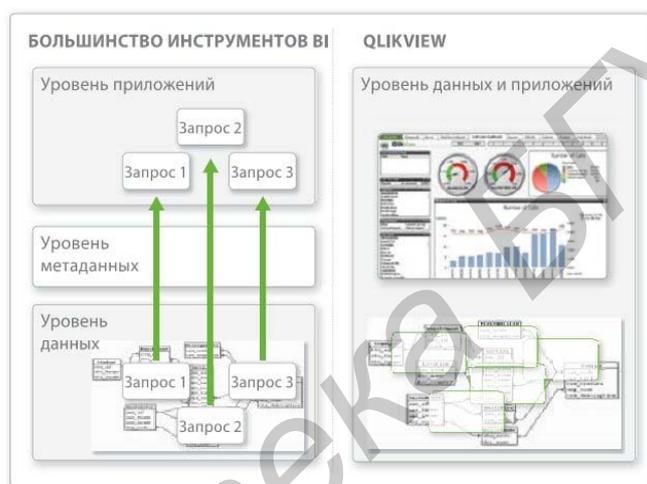


Рис. 1. Ассоциативная архитектура QlikView

Современные системы бизнес-анализа (Business Discovery) имеют следующие отличия от систем бизнес-анализа на базе OLAP-технологий:

- системы бизнес-анализа Business Discovery хранят транзакционные данные, что позволяет не изменять модель/архитектуру данных. Бизнес-задачи решаются и новая отчетность создается на уровне визуализации данных;
- в OLAP-кубах существуют избыточные данные (нули);
- OLAP-кубы сложно связывать;
- в куб сложно добавлять измерения (они стремительно увеличивают размер);
- при формировании новых запросов часто необходимо создание новых OLAP-кубов;
- так как в кубах хранятся суммы, поэтому для новых разрезов строятся новые кубы;
- модель данных может быть разработана гораздо быстрее: создание OLAP-куба для SAP BW – от 5 дней, решение такой же задачи в QlikView – всего 5 минут;

- любое поле в QlikView может быть и показателем и измерением одновременно;
- с точки зрения бизнес-пользователя в QlikView каждый следующий анализ делается на основе предыдущих;
- решение QlikView может поддерживать 1 человек и туда входит всё: ETL, модель данных, хранение и визуализация. В крупных системах этим занимается минимум 3 человека;
- в существующих системах бизнес-анализа очень много продуктов, за каждый продукт отвечает конкретный человек.

Онлайновая аналитическая обработка данных (OLAP) использует агрегированные данные для поддержки принятия решений. Существует множество разновидностей OLAP. Некоторые из них отличаются большей гибкостью, другие – высокой производительностью, но все они основаны на запросах и потому неспособны поддерживать взаимосвязи между элементами данных.

	QlikView	ROLAP	MOLAP	HOLAP
Гибкость	Высокий уровень	Высокий уровень	Низкий уровень	Высокий уровень
Производительность	Высокий уровень	Низкий уровень	Высокий уровень	Высокий уровень
Ассоциативность	Высокий уровень	Низкий уровень	Низкий уровень	Низкий уровень

Рис. 2. Сравнение технологий обработки данных

Решения на базе OLAP обеспечивает высокую скорость обработки запросов, но обладает рядом серьезных недостатков:

- 1 На кубах с большой размерностью, отношения объема многомерного куба к объему исходных данных, часто, составляет десятки и сотни раз.
- 2 Загрузка новой порции данных приводит к перестройке всего куба, что значительно снижает производительность. Кроме того во время загрузки пользователи не могут работать с данными.
- 3 Изменение в структуре куба требует к его полную перестройку.
- 4 Разрешены произвольные Пользователи не могут осуществлять произвольные запросы к системе, только те, которые были запланированы при проектировании куба. Произвольные запросы невозможны.

Также приведем основные преимущества OLAP системы:

- согласованность исходной информации и результатов анализа;
- проведение многовариантного анализа;
- управление детализацией;
- выявление скрытых зависимостей;
- создание единой платформы.

Реальность текущих бизнес сценариев выглядит таким образом, что большинство из них являются неоптимальными с точки зрения производительности. В настоящее время традиционная IT инфраструктура не может полностью удовлетворять потребностям одновременно транзакционных и аналитических систем. Это соответственно ведет к тому, что в архитектуре систем делаются уступки, тем или иным образом компенсирующие слабые стороны. Традиционные архитектуры практически не позволяют сочетать транзакционные среды (продажа товаров, обслуживание заказов, сбор денег) и аналитические (например доходность от продаж в разрезе заказчика или, например, конкретного региона).

Можно выделить несколько потенциальных областей, в рамках которых можно достигнуть потенциального повышения производительности и рассматривать QlikView как потенциальную замену каким либо другим средствам аналитики (OLAP):

- самоиндексация(система сама строит индексы по таблицам. таким образом, архитекторам и администраторам больше нет необходимости тратить время на создание, администрирование, анализ и оптимизацию индексов в схемах данных);

- высокая степень сжатия данных(сжатие данных до 20 раз по сравнению с их оригинальным состоянием);

- поколонное хранение данных, хранение данных в памяти, консолидация транзакционных и аналитических данных (объединение транзакционных и аналитических данных (расчета на лету) в едином хранилище;

- новые типы бизнес приложений (использование QlikView в качестве аналитического инструмента представляет возможности по разработке совершенно новых типов приложений, таких как анализ займов, онлайн оценка рисков, и т.д.).

Главным достоинством современных систем бизнес-анализа в отличие от технологии аналитической обработки информации в реальном времени OLAP (Online Analytical Processing), является их гибкость, высокая скорость, а также поддержка необходимых ассоциативных связей между данными. Система QlikView является реляционной платформой с кубами, создаваемыми по требованию пользователя, которая в состоянии обеспечить уровень гибкости аналогичный ROLAP при скорости MOLAP.

Литература

- [1]. Inmon W.H. Building the Data Warehouse. New York: John Willey & Sons, 1992. 576 p.
- [2]. Henschen D. Gartner BI Magic Quadrant: Winners & Losers // InformationWeek. Online Media Community. 2014. Available:
- [3]. Черняк Л. Визуальная аналитика и обратная связь // Открытые системы. Электрон. журн. 2013. №6.
- [4]. Свирцевский А. BI или бизнес-аналитика? Точка зрения SAS // Компьютерное Обозрение. Электрон. журн. 2014.
- [5]. А.А. Барсегян, М.С. Куприянов, В.В. Степаненко, И.И. Холод Методы и модели

анализа данных: OLAP и Data Mining

[6]. Салихов Е. Руководство по созданию и использованию аналитических приложений QlikView/ компания QlikView. М.: Консультационная группа АТК, 2010. 16 с.

[7]. Иткин В. Легенды и мифы о QlikView // Jet Info. Электрон. журн. 2013. №2. Режим доступа: <http://www.jetinfo.ru/author/vladimir-itkin/legendy-i-mify-o-qlikview>

[8]. Саврасов С.Б. Сравнительный анализ методов извлечения данных из источников данных информационной системы применительно к задаче агрегации данных // Молодежный научно-технический вестник.. Электрон. журн. 2013. № 12. Режим доступа: <http://sntbul.bmstu.ru/doc/641063.html>

Библиотека БГУИР