

Новые горизонты BI

Обзор основных тенденций на рынках программных и аппаратных решений для бизнес-аналитики, аналитики больших данных и построения хранилищ данных.



Лозовая Екатерина Александровна — директор ЗАО «Тери».

Эра бизнес-анализа

Мировой рынок систем бизнес-аналитики может вырасти в 2012 г. по сравнению с 2011 г. на 8,2% — до \$33,9 млрд, согласно данным IDC. Эксперты прогнозируют, что спрос на решения бизнес-аналитики будет стабильно увеличиваться, поскольку возрастает осведомленность компаний о преимуществах аналитических инструментов и их уникальных возможностях для принятия управленческих решений.

К 2015 г., по оценкам IDC, компании вложат свыше \$120 млрд в аппаратные средства, ПО и услуги, способные реализовать бизнес-потенциал аналитических решений.

В ближайшие несколько лет инструменты аналитики, ранее в основном использовавшиеся в сфере business intelligence для топ-менеджеров, станут стандартной компонентой большинства ИТ-систем. BI-опции будут использоваться для принятия решений, в основном, в онлайн-режиме, менеджерами всех уровней во взаимоотношениях с партнерами и клиентами. Одновременно клиенты при взаимодействии с ИТ-системами будут получать персонализированные услуги на основе динамически изменяющейся модели предпочтений и истории взаимодействия с компанией.

Более того, инструменты аналитики проникнут и в повседневную жизнь людей, делая ее разумнее, удобнее, лучше.

Изменения, которые произошли за последние несколько лет на рынке решений бизнес-аналитики, были самыми большими за всю его более чем 20-летнюю историю. Прежде всего, это связано со следующими факторами.

Высокопроизводительная аналитика

Это один из наиболее значимых векторов развития бизнес-анализа. Он поддержан почти всеми ключевыми игроками мирового рынка и появился вследствие предъявления новых требований к BI, которые сводятся к двум основным: максимальное приближение результатов бизнес-анализа к реальному времени и необходимость обработки очень больших массивов разнотипных данных. В настоящее время высокопроизводительная аналитика реализуется на базе трех технологий:

- аналитика in-memory;
- массивно-параллельная обработка запросов на уровне самого хранилища данных (DW);
- распределенная аналитика (grid-аналитика) по множеству независимых узлов обработки.

Все эти три технологии могут быть интегрированы как в составе одного решения, так и предлагаться в виде отдельных решений, дополняющих существующие методы бизнес-анализа. Интегрированные решения могут предлагаться как от одного вендора, так и от нескольких в виде совместных разработок.

Аналитика In-Memory

Благодаря новым поколениям процессоров, имеющих до 8 ядер на одном процессоре и до 1 Тбайт ОП в составе одного стандартного процессора, появилась возможность реализовать многолетнюю мечту — перенести всю обработку бизнес-данных в оперативную память (in-memory) на базе многоузловых высоконадежных с многотерабайтным объемом ОП SMP-кластеров (Symmetric Multi Processing). Это позволяет, по сути, революционизировать OLTP- и OLAP-обработку (Online Transactional Processing — OLTP, Online Analytical Processing — OLAP).

Благодаря этому удается организовать in-memory-работу сразу для многих приложений, значительно сокращая потребности в ОП, и не только от 10 до 1000 раз повысить скорость обработки транзакций и запросов, но и существенно улучшить управляемость и расширяемость приложений.

SAP HANA — одно из первых решений на рынке, полностью ориентированное на in-memory-обработку OLTP- и OLAP-приложений, активно поддерживаемое многими мировыми разработчиками серверов, включая HP, Fujitsu и др. SAP HANA предлагается как дополняющее решение к существующим BI-системам

и хранилищам данных. SAP HANA за счет многоузлового кластера с общей многотерабайтной ОП позволяет грузить в ОП хранилище данных и все приложения, которые с ним работают.

Объявлено об интеграции SAP HANA с SAP Sybase IQ server, что позволит создавать интеллектуальное хранилище устаревших или неактуальных данных, отличающееся малой совокупной стоимостью владения.

Примерами решений, которые осуществляют обработку запросов in-memory непосредственно в рамках самого хранилища, может служить, например, IBM Netezza и др.

Аналитика больших данных

В последний год стало активно развиваться направление, связанное с аналитикой больших данных (АБД), позволяющее использовать для бизнес-анализа не только транзакционные данные, но и большой объем неструктурированных данных, извлекаемых из корпоративных файловых хранилищ, интернета, социальных сетей, мобильных устройств, различных датчиков и др.

Аналитику больших данных, в основном, связывают с интеграцией традиционных BI-систем с Hadoop-кластерами. Последние, например, являются основой для хранения данных/логов многих поисковых систем и интернет-магазинов. Поддерживая единую распределенную файловую систему на множестве стандартных серверов, Hadoop-кластеры обеспечивают массовый параллелизм обработки при выборке метаданных из большого объема информации для дальнейшей их интеграции и анализа в составе традиционных хранилищ данных (DW — Data Warehouse).

Также в основе решений для АБД предполагается использование высокопроизводительных специализированных решений для развертывания DW, без которых достижение высоких скоростей (часто с требованием в режиме онлайн) при анализе больших данных и большого трафика становится невозможным.

АБД, прежде всего, привносит динамизм в традиционный BI, позволяя принимать решения не на основе заранее заданных сведений о клиенте/объекте/процессе, а на основе динамически меняющейся поведенческой модели. Одновременно АБД позволяет вскрывать тенденции и строить корреляции между данными/фактами/событиями, которые раньше было вообще невозможно выявить в рамках традиционных инструментов BI.

XIII ежегодная конференция компании “Терн”

Для банковского сектора решения данного класса ассоциируются, например, с борьбой с кибермошенничеством, которые позволяют выявлять факт злонамеренных действий в момент его совершения, а не после, как было возможно до недавнего времени, тем самым предупреждая все негативные его последствия.

Многие новые версии ПО, ориентированные на средний бизнес, сейчас уже предлагаются с учетом работы с большими данными. Так, IBM недавно выпустила новые версии ПО для управления базами данных и организации информационных хранилищ – DB2.10 и InfoSphere Warehouse 10. Новое ПО с легкостью интегрируется с системами обработки больших данных, автоматически сжимает данные в более компактные и поддающиеся обработке блоки, чтобы предотвратить разрастание ресурсов хранения, и делает “временной срез” данных, выделяя данные из прошлого, настоящего и будущего, избавляя от необходимости создания дорогостоящих дополнений к коду программ.

В ходе тестирования обработка запросов к системе хранения данных производилась до 10 раз быстрее, ускоряя принятие решений. С использованием нового ПО удалось увеличить объем свободного пространства в хранилище данных до 90%. С помощью нового ПО клиенты могут:

- обрабатывать большие данные и проводить их оперативный анализ в реальном времени для получения углубленного информационного представления из неструктурированных данных, таких как контент социальных медиа или текстовые сообщения с мобильных устройств. Информация, полученная из систем на базе платформы Nadoop, теперь с легкостью интегрируется с результатами анализа структурированных данных в информационном хранилище для поддержки принятия более оперативных и аргументированных решений;
- ускорять бизнес-процессы и снижать расходы на управление данными с помощью функции Adaptive Compression и Multi-Temperature Data Management, обеспечивая мгновенное сжатие данных, облегчая бизнес-приложениям их использование и размещение в наиболее эффективных устройствах хранения. Новое ПО автоматически определяет, насколько часто необходимы те или иные данные, и оптимально распределяет их по устройствам хранения с точки зрения экономической эффективности и текущей актуальности;
- изучать исторические данные и оценивать перспективу для улучшения принятия решений с помощью функции Time Travel Query, которая облегчает доступ к данным, относящимся и действительным на любой момент времени. Так, онлайн-туристическое агентство может автоматически определять противоречия в организации трансферов, когда, например, гостиница в Риме забронирована на восемь дней, а автомобиль зарезервирован в Нью-Йорке на три дня в течение этого периода. Раньше администраторам баз данных и разработчикам приложений приходилось создавать сложные программы для выявления таких взаимосвязей.

6 июня 2012 г. состоится XIII ежегодная конференция компании “Терн”, посвященная практике использования технологий business intelligence для управления бизнесом. С каждым годом число ее участников постоянно увеличивается, и не большой клуб людей, объединенных “по интересам”, вырос в крупное ИТ-мероприятие для топ-менеджмента, руководителей и специалистов по ИТ и финансам.

Традиционно конференцию компании “Терн” посещают профессионалы с самыми разными зонами ответственности:

- *руководители, отвечающие за принятие бизнес-решений*, могут узнать, какие из актуальных задач предприятия уже сейчас успешно решаются средствами business intelligence;
- *директора информационных служб и руководители ИТ-департаментов* с большим вниманием относятся к докладам наших клиентов и партнеров, в которых рассматриваются успешные внедрения;
- *руководители проектов, аналитики и специалисты* с интересом общаются с коллегами, обсуждая новые возможности программных продуктов, сравнивая версии и оценивая работу разработчиков;
- *бизнес-пользователи*, представляющие банковские и финансовые структуры, телекоммуникационные и торговые компании, промышленные предприятия, государственные учреждения и являющиеся непосредственными потребителями ВІ-услуг, имеют возможность задать детальные вопросы специалистам, изучить возможности продуктов на демонстрационных стендах, получить информацию об условиях обучения и технической поддержки.

За 12 лет конференцию посетило более 4000 представителей бизнеса и экспертов в области информационных технологий.

Гостями конференции в разные годы были компании SAP, BusinessObjects, IBM, Ascential, Informix, KXEN, Axiom EPM, Micropole Univer, GoldenSource.

В рамках развития направления АБД в конце марта 2012 г. IBM анонсировала решение IBM Smarter Analytics Signature Solution для выявления и сокращения случаев мошенничества, повышения эффективности финансовых операций и улучшения обслуживания потребителей.

По данным ФБР, ущерб от мошенничества только в сфере здравоохранения достигает \$250 млрд в год. Налоговое мошенничество обходится еще во многие миллиарды долларов. Адаптивные системы IBM, способные “обучаться” на обновляемых данных, помогают защитить-

Возможность прямого диалога с вендорами на конференции “Терн” существенно повысила степень доверия российских клиентов и партнеров к представляемым решениям, убедила их в правильности пути развития ИТ инфраструктуры предприятия через внедрение систем business intelligence.

Программа конференции 2012 года включает:

- доклад представителя компании KXEN, посвященный предсказательной аналитике;
- выступление представителя компании SAP о новых продуктах и технологиях SAP, тенденциях и перспективах их развития;
- выступление особого гостя – Пола Барбера, директора компании PROPHIX Software;
- доклады о новых проектах компании “Терн” с участием представителей организаций из банковской сферы, торговли, ТЭК и др.

При формировании программы конференции является обязательным отражение в докладах практического опыта реализации ВІ-проектов на предприятиях наших клиентов. Ведь человеку, стоящему перед выбором поставщика и партнера, отвечающему за внедрение, крайне важно услышать полезные рекомендации или советы именно от своих коллег.

Все клиенты “Терн” охотно откликаются на призыв поделиться опытом, неизменно участвуя с докладами и отвечая на вопросы из зала. Хотелось бы выразить благодарность всем нашим спикерам со стороны клиентов, среди которых необходимо особенно отметить представителей Банка России, Национального банка Казахстана, ОАО “Внешторгбанк”, ОАО “ФСК ЕЭС”, ОАО “МТС”, ОАО “Сургутнефтегаз”, АКБ “Росбанк”, ОАО “Вымпелком”, ОАО “Ростелеком”.

Мероприятие состоится в московском “Президент-Отеле”. Участие бесплатное, но требует обязательной регистрации на сайте – www.tern.ru.

ся от новых способов мошенничества. Решение IBM встраивает передовые аналитические алгоритмы непосредственно в бизнес-процессы, давая государственным учреждениям и страховым компаниям возможность выявлять случаи мошенничества в режиме реального времени – до незаконного получения злоумышленниками выплаченных средств. Применяя сложные аналитические методики, решение рекомендует пользователям наиболее эффективное средство защиты в каждом конкретном случае, оптимально используя ограниченные ресурсы организаций. Например, в одном случае система может

установить и сообщить, что получено обычное письмо с законной просьбой об оплате, а в другом случае — рекомендовать проведение полного расследования.

Всего к 2015 г. IBM планирует получить доход на рынке бизнес-аналитики в размере \$16 млрд.

Специализированные решения для DW

Все большее распространение в составе BI-решений получают специализированные устройства для DW. В основном они позиционируются для крупного бизнеса и в России стали активно продвигаться с середины 2011 г. В настоящее время решения данного класса в России представлены от: IBM — Netezza, Oracle — Exadata, HP — Vertica, EMC — Greenplum, SAP — SAP Business Suite/Sybase ASE, Teradata — Teradata Active Enterprise Data Warehouse 6650/6680.

Одни из основных преимуществ, которые предлагают подобные решения:

- многократно (до 100 раз и выше) увеличенную производительность в сравнении с конфигурациями на стандартных (enterprise-класса) компонентах;
- поддержку многотерабайтных и петабайтных DW без снижения производительности и высокой масштабируемости;
- минимизацию усилий на развертывание и дальнейшее сопровождение DW;
- поддержку рядом решений одновременную работу OLTP- и OLAP-приложений, а также обработку запросов in-memory на уровне самого специализированного DW;

Многие из вендоров уже анонсировали поддержку таких решений с NooP-кластерами.

Поколоночное хранение

К началу 2012 г. уже все ключевые игроки на рынке корпоративных реляционных СУБД стали предлагать поколоночное хранение данных — наиболее предпочтительное (в сравнении с построчным хранением) практически для всех OLAP- и BI-приложений, обеспечивающее более высокую производительность на том же самом оборудовании. Одновременно поколоночное хранение позволяет гораздо проще организовать параллельную обработку колонок в сравнении с построчным хранением¹⁾. Также поколоночное хранение в среднем обеспечивает в 10 раз более высокую компрессию данных в сравнении с построчным хранением²⁾.

Аналитика для всех и в любом месте

В настоящее время модули бизнес-анализа стали включаться многими вендорами во все компоненты бизнес-приложений, что дает возможность использовать предикативный и аналитический инструментариум служащими всех уровней и по-

зволяет вывести бизнес-компания на качественно новый уровень. Помимо этого, модули бизнес-анализа стали активно использоваться и в системах взаимоотношений с клиентами (например, на основе АБД), что приводит к повышению качества услуг, а это, в свою очередь — к росту бизнеса компании.

В этом контексте аналитика по запросу, предоставляемая, например, средствами SAP BusinessObjects Web Intelligence XI, для формирования отчетности по запросу (ad-hoc отчетности) и позволяющая самостоятельно создавать отчеты и проводить анализ без привлечения сотрудников ИТ-отдела, является иллюстрацией приближения BI “к народу”.

Расширяют эти возможности инструменты типа SAP Voyager, которые позволяют бизнес-пользователям самостоятельно проводить навигацию в OLAP-кубах, осуществлять построение нестандартных аналитических запросов без обращения в ИТ-службу.

Инструмент SAP BusinessObjects Mobile еще в большей степени увеличивает самодостаточность бизнес-пользователей для получения существующего BI-контента на мобильных устройствах.

Недавно SAP объявила о своих планах по приобретению Syclo — ведущего поставщика мобильных приложений и технологий для крупных предприятий.

Этот шаг компании SAP направлен на удовлетворение потребностей крупных компаний в расширенном использовании сотрудниками интеллектуальных и мобильных устройств в повседневной работе. Согласно отчету компании Gartner Inc., к 2013 г. мобильные телефоны станут основным средством доступа к интернету по всему миру, обойдя персональные компьютеры. Приобретение Syclo — один из последних шагов компании SAP, направленных на стимулирование роста и внедрение инноваций в своих основных приложениях, аналитических средствах, мобильных решениях для крупных предприятий, облачных средах, а также СУБД и технологиях, применяемых компаниями.

Платформы для совместной аналитической работы

Современный бизнес-анализ требует совместной работы большого коллектива аналитиков, позволяющих оперативно принимать правильные решения в точках роста бизнеса. Однако существующие традиционные BI-решения уже не позволяют полноценно реализовать это взаимодействие, прежде всего, из-за ограниченности доступных ресурсов, которые можно использовать в рамках существующих архитектур BI-систем. К этому следует добавить трудности: по совместному анализу витрин данных, создаваемых в разных подразделениях; по исследованию больших данных (до десятков и сотен терабайт) структурированных и неструктурированных; в производительности и др.

В настоящее время подобные платформы активно разрабатываются и уже предлагаются, включая их использование и на мобильных устройствах. Например, решение SAP StreamWork позволяет объединить людей, информацию и бизнес-

подходы для достижения быстрых и значимых результатов. Благодаря комбинированию оперативного социального взаимодействия и BI (в том числе и за счет интеграции SAP StreamWork с SAP BusinessObjects Mobile), достигается ускорение и повышение качества принимаемых решений.

*Екатерина Лозовая,
директор ЗАО “Терн”*

IBM: роль директоров по ИБ возросла

Май 2012 г. — Новое исследование IBM (ibm.com/smarter/cai/security) показывает отчетливую эволюцию корпоративных служб информационной безопасности и их руководителей: происходит переход от исполнения функций технической поддержки к стратегическому управлению бизнесом — несмотря на то, что в настоящее время лишь один из четырех опрошенных директоров по ИБ играет стратегическую роль в своей компании.

В рамках первого исследования роли CISO центр IBM Center for Applied Insights опросил свыше 130 директоров по ИБ и выявил три типа руководителей этой категории, с точки зрения готовности к устранению уязвимостей защиты, а также общего уровня развития корпоративной системы безопасности. Тип директора по ИТ-безопасности, получивший название “Influencer” и представленный почти четвертой частью всех респондентов, характеризуется непосредственным влиянием на бизнес-стратегию компании. Руководители этой категории, как правило, показывали себя более уверенными и подготовленными, чем их коллеги, отнесенные к типам “Protector” (занимаются исключительно техническими и организационными вопросами информационной защиты) и “Responder” (несут ответственность за ИТ-безопасность, но не имеют влияния на принятие решений).

Сегодня всем директорам корпоративных служб безопасности приходится решать сложнейшую задачу по защите наиболее ценных активов компании — денег, данных о клиентах, интеллектуальной собственности и бренда. Почти две трети опрошенных CISO заявили о том, что за последние два года внимание к вопросам ИБ выросло: серии широкомасштабных атак и утечек конфиденциальной информации убеждают высший менеджмент компаний в ключевой роли безопасности на современном предприятии. Более половины респондентов сочли обеспечение мобильной безопасности самой приоритетной технической задачей в ближайшие два года. Почти две трети ожидают увеличения расходов на информационную безопасность в следующие два года, а 87% из них предполагают, что этот рост будет выражаться двузначными цифрами. Если раньше задача CISO заключалась в реагировании на случаи нарушения безопасности, то сегодня роль директора по безопасности меняется в направлении более разумного и всестороннего управления рисками — от “тушения пожаров” к их предотвращению.

1) При построчном хранении данные таблиц БД хранятся блоками, каждый из которых представляет строку таблицы БД. При поколоночном хранении данные таблицы БД хранятся блоками поколоночно. При этом в каждой колонке хранятся и идентификаторы строк.

2) A Common Database Approach for OLTP and OLAP Using an In-Memory Column Database, Hasso Plattner (<http://www.sigmod09.org/images/sigmod11ktp-plattner.pdf>).