



Наталья Жилкина/
[nzhilkina@computerra.ru/](mailto:nzhilkina@computerra.ru)

BIG DATA: по закону больших чисел

Многие руководители ИТ-подразделений пытаются управлять «большими данными» традиционными средствами, упуская из виду специфику их обработки и откладывая решение проблем на потом. Наиболее часто проблемы замедления и даже остановки приложений возникают на уровне доступа к хранилищам данных (Data Warehouse). Это ведет к пробуксовке бизнес-процессов и деградации результатов бизнеса в целом.

Уточняя определение BIG DATA, специалисты компании Microsoft выделяют две категории задач, отличающихся друг от друга на несколько порядков по объему обрабатываемых данных.

Алексей Халяко, специалист в области программных решений группы SQL Customer Adviser Team в компании Microsoft, подчеркивает, что определение BIG DATA относится исключительно к категории мегахранилищ. На мировом рынке существует не много компаний, которые хранят и «перемальвают» в распределенной среде не просто большие, а громадные объемы данных, в основной своей массе неструктурированные. К их числу относятся Google, Amazon, Yahoo, Facebook, Microsoft, NASA и еще несколько им подобных.

Платформа для BIG DATA

Такие данные обрабатываются в сильно распределенной среде, параллельным способом, с помощью специальных методов и инструментов; наиболее известным из них является платформа Hadoop передового проекта, в развитии которого участвует сообщество разработчиков со всего мира.

Система Hadoop основывается на распределенном подходе к вычислениям и хранению информации. Она отличается горизонтальной масштабируемостью; экономичностью (информация и вычисления распределяются по кластеру, состоящему из большого числа узлов и построенному на самом обыкновенном оборудовании); эффективностью (распределение данных позволяет выполнять их обработку параллельно на множестве компьютеров, что существенно ускоряет этот процесс); надежностью (при хранении данных возможно обеспечение избыточности — благодаря хранению нескольких копий).

В основе Hadoop лежит распределенная файловая система (Hadoop Distributed File System, HDFS), использующая для хранения данных вычислительные узлы кластера, а также модель MapReduce для проведения масштабных вычислений при обработке огромных объемов данных. Согласно этому подходу приложение разделяется на большое количество небольших заданий, каждое из которых может быть выполнено на любом из узлов кластера. Вкупе с распределенной файловой системой это позволяет приложениям легко масштабироваться практически неограниченно. Источником вдохновения для разработчиков Hadoop послужили материалы по Google File System.

В настоящее время главной сферой интереса Microsoft является среда корпоративных данных, в том числе создание и управление хранилищами, которые содержат петабайты данных, в том числе и неструктурированных. Компания также предлагает средства интеграции и с технологиями HADOOP. «Мы предлагаем высокопроизводительные коннекторы и стандарты типа SQOOP, которые обеспечивают требования, касающиеся высокопроизводительного ввода-вывода, для обеспечения загрузки и выгрузки данных в HADOOP», — поясняет Александр Гвоздев, специалист по платформам и приложениям в компании Microsoft.

Впрочем, если исходить из определения большого хранилища, предложенного основателем концепции Data Warehouse Биллом Инмоном (Bill Inmon), который еще несколько лет тому назад осторожно уточнил, что большим хранилищем данных можно считать такое, в котором объем запросов на несколько порядков выше, чем в традиционном хранилище, — то далее в статье мы не станем придерживаться строгого определения BIG DATA в категориях Hadoop и будем рассматривать их с точки зрения подходов к обработке корпоративных данных, которые не могут быть реализованы современными, традиционными, стандартными, широко распространенными и имеющимися под рукой инструментами.

Корпоративные BIG DATA

Какие проблемы возникают при росте данных в обычных компаниях, не относящихся к категории таких супергигантов, как Google или Facebook?

Данные, которые приходится хранить и обрабатывать в корпоративной среде, множатся. И хотя это не мегаразммеры, но многие компании оказались не готовы к столь стремительному росту. Поступает много неструктурированной информации, под нее надо создавать хранилище, которое быстро разрастается.

По мере накопления данных доступ к ним замедляется: поиск данных становится медленным из-за того, что он ведется в больших диапазонах. «Рост объема существующих СУБД приводит к увеличению времени обработки запроса и поиска нужной информации. С ускорением бизнес-процессов время выполнения сервисных операций должно сокращаться», — отмечает Владимир Кирсанов, технический директор департамента корпоративных систем компании «Астерос».

Как поступают на практике в случае замедления СУБД? Нередки сценарии, когда создается второе, третье хранилище; они перестают быть консистентными, и наступает в каком-то смысле хаос.

Денис Серов, ведущий технический консультант компании «ЕМС Россия», описывает типичный сценарий: компания использует обычную базу данных. Со временем возникает идея разделить боевую базу данных и хранилище, из которого строятся отчеты. Такое разделение хранилищ можно наблюдать, например, на уровне различных департаментов.

— Обычно для этой цели берут инструменты того же производителя, которые использовались для транзакционных баз, и начинают с этим работать, — поясняет Серов. — По такому пути идут прежде всего потому, что он почти не требует инвестиций, специального развития экспертизы. Но сам по себе подход к выбору транзакционной базы данных в качестве информационного хранилища представляет собой мину замедленного действия, потому что эти хранилища начинают увеличиваться и по достижении некоторого размера (иногда до 10 терабайт) скорость обработки данных очень сильно замедляется. Для выхода из ситуации, чтобы ускорить анализ, создают еще одно хранилище. А на практике получается, что данные рассредоточены по разным хранилищам, которые могут оказаться неконсистентными, а отчеты — неточными, неактуальными.

Такие системы очень сложны в администрировании. Администраторы вынуждены тратить основное время на их сопровождение. А на разработку новых функциональных модулей, которые решают бизнес-задачи, времени уже не остается. Подход, который на первый взгляд не требует обучения нового персонала, специальной экспертизы и новых инвестиций, в итоге оборачивается тем, что все ресурсы оттягиваются на сопровождение. Возникает патовая ситуация, когда и новых сотрудников нет, и все имеющиеся заняты на 100%.

Решение проблемы специалисты видят в создании отказоустойчивой системы, которая наряду с требованиями к надежности отвечала бы также целям распределения нагрузки по серверам, повышения производительности и масштабируемости систем, создания специализированных информационных «решеток» (grid) для параллельной обработки баз данных.

Желательно, добавляет Денис Серов, чтобы у администраторов корпоративных систем был выбор скорости, которую предпочтительно использовать для обработки тех или иных данных. «Если в текущий момент, — говорит он, — скорость обработки не устраивает, у специалистов должны быть инструменты для легкого преобразования системы — с тем чтобы добиться приемлемой скорости».

Проблема динамического добавления компонентов в систему, как поясняет Серов, связана со сложностью отслеживания всех изменений: «Эта проблема не решается в традиционных системах. А если и решается, то с ограничениями, которые неприемлемы для выполнения бизнес-процессов. А вот в информационных системах нового поколения можно добавлять новые компоненты и перераспределять нагрузку уже автоматически, без необходимости сложной перенастройки, миграции данных».

Правильное решение, по рекомендации Дениса Серова, заключается в следующем: надо как можно раньше выбирать единое хранилище Enterprise DataWarehouse, которое может поглотить все данные и затем быстро их анализировать — не дожидаясь, когда проблема станет уже неразрешимой, как гордиев узел.

И еще один момент, на который обращает внимание Серов: для разработчиков нового функционала хранилищ тоже нужна специальная среда разработки, тестирования, отладки:

— На боевой системе это делать нельзя. Многие специализированные хранилища построены на нестандартном оборудовании, так что ПО для среды разработки нельзя просто скачать из Интернета и поставить на сервер. Придется приобретать еще одну систему для разработчиков, а они не дешевы. Эти проблемы снимаются, если хранилище построено на архитектуре открытых стандартов — как, например, GreenPlum.

Сезон покупок

Среди имеющихся на российском рынке производителей решений для работы с большими объемами данных в корпоративной среде наиболее известны (по алфавиту) EMC, IBM, Intersystems, Microsoft, Oracle, Teradata и еще несколько других.

Относительно недавно компания Teradata Software приобрела бизнес Aster Data, американского разработчика кластерной

программных решений Data Warehouse, основанных на СУБД с открытым кодом PostgreSQL и решающих «корпоративные» задачи в области бизнес-аналитики, а также ориентированных на применение в рамках облачных платформ.

Корпорация Microsoft приобрела и поглотила Data Allegro, в результате чего в ее линейке появился PDW (Parallel Data Warehouse) — главный DWH-продукт компании Microsoft в сегменте BigData.

Три подхода

Компания Microsoft для работы с большими объемами данных в корпоративной среде предлагает несколько решений. Во-первых, это Microsoft SQL Server 2008 R2 — традиционное ХД, которое развивается от версии к версии.

Для хранения и обработки больших массивов данных (от 5–10 до 80–100 Тбайт) Microsoft позиционирует решение SQL Fast Track на основе сбалансированной SMP-архитектуры.

— За небольшим исключением, очень крупными массивами данных, от 25–50 Тбайт, российские компании не располагают, — отмечает Алексей Халяко. — Большая часть оперирует объемом не более 1–2 терабайта. Российские заказчики пока еще не подготовлены к таким объемам, поэтому переход на узкопрофилированные, специализированные продукты с массивно-параллельной обработкой для многих из них является экзотикой.

Третий подход к решению задачи обработки больших массивов данных (от 100 Тбайт до единиц Пбайт) — Parallel Data Warehouse, решение на архитектуре MPP (массивно-параллельной обработки данных).

Независимо от вендора на рынке существуют два подхода к работе с быстро растущими объемами данных: технология Scale-up (вертикальное масштабирование) и Scale-in (горизонтальное масштабирование).

Принцип Scale-up в случае решения проблемы роста данных состоит в приобретении более мощного аппаратного обеспечения. Проблема данного подхода в том, что в какой-то момент возможности аппаратного и программного обеспечения обработать определенную нагрузку исчерпываются — ведь количество ядер, вычислительная способ-

ность одного узла ограничены. Поэтому рано или поздно SMP-архитектура упирается в пределы сотни Тбайт.

Максим Гончаров, специалист по платформе приложений компании Microsoft, отмечает еще один недостаток вертикального масштабирования: на каком-то этапе наращивания вычислительных мощностей эксплуатация такой системы становится существенно дороже — из-за того что стоимость против производительности начинает расти нелинейно. «В моей практике, — поясняет Гончаров, — была ситуация (лет восемь тому назад), когда одна крупная компания приобрела сервер, который на 5–10% увеличивал производительность (вышла следующая версия, новые ядра), но денег они заплатили вдвое с лишним больше. Проблема еще и в том, что если остаться на старом сервере, то все равно придется докупать дисковые емкости для хранения данных. Если же докупать ресурсы хранения, то сервер в какой-то момент перестает „переваривать“ объемы данных, которые распределены по всем дисковым массивам. Простая, казалось бы, задача добавления нового диска не всегда дает желаемый результат из-за

Проблема динамического добавления компонентов в систему связана со сложностью отслеживания всех изменений и не решается в традиционных системах. А если и решается, то с ограничениями, которые неприемлемы для выполнения бизнес-процессов.

системы управления базами данных nCluster, поддерживающей MapReduce-расширение SQL.

Компания IBM приобрела компанию Netezza, поставщика аппаратно-программных хранилищ данных, с целью расширить портфель продуктов для бизнес-аналитики на основе массивных распределенных хранилищ.

Для работы с большими хранилищами данных Oracle позиционирует решение Exadata. С кластерной базой данных, расположенной на общих дисковых устройствах, одновременно может работать множество экземпляров СУБД Oracle, запущенных на различных узлах кластера.

Hewlett-Packard приобрела компанию Vertica, занимающуюся разработкой решений для хранения и анализа данных, вскоре после того, как HP сообщила о прекращении работ над собственной платформой для хранилищ данных большого объема Neoview.

EMC, один из лидеров на мировом рынке решений по хранению данных, поглотила компанию Greenplum, известную на рын-

проблемы равномерного распределения данных на носителях. Добавление нового массива может привести к тому, что нагрузка на этот массив повысится, в то время как старые будут недогружены запросами на чтение и запись. Узкое место приводит к тому, что количество обрабатываемых данных в секунду может быть очень низким».

Поумерить эти финансовые аппетиты позволяет архитектура scale-up — за счет использования большего количества узлов для решения такого рода задач. Реализация решений на базе MPP-архитектуры может быть разной, но суть в том, что добавление нового массива может привести к тому, что нагрузка на этот массив повысится, в то время как старые будут недогружены запросами на чтение и запись. Узкое место приводит к тому, что количество обрабатываемых данных в секунду может быть очень низким.

Для работы с очень крупными массивами данных у компании Microsoft есть специализированное программно-аппаратное решение Parallel Data Warehouse (PDW), поставляемое совместно с HP, Dell или IBM. PDW — это классический пример MPP-архитектуры.

■ **Максим Гончаров:** «При вертикальном масштабировании на каком-то этапе наращивания вычислительных мощностей эксплуатация системы становится существенно дороже из-за того, что стоимость против производительности начинает расти нелинейно».



Как поступают на практике в случае замедления СУБД? Нередки сценарии, когда создается второе, третье хранилище, они перестают быть консистентными, и наступает в каком-то смысле хаос.

На принципах массивно-параллельной обработки, горизонтального и вертикального масштабирования построена и система GreenPlum, вошедшая в линейку решений EMC вместе с приобретением одноименной компании. «Система практически не имеет пределов масштабирования, — поясняет Денис Серов. — Добавление новых узлов ограничено лишь бюджетом заказчика или местом в серверной комнате».

— Решение GreenPlum масштабируется очень быстро и может анализировать данные объемом до нескольких петабайт практически безгранично, — подтверждает Алексей Пилипчук, технический директор компании «Техносерв». — Его использование может стать конкурентным преимуществом для банков: во-первых, за счет высокой производительности, а во-вторых, по причине небольшой стоимости (Greenplum заметно дешевле аналогов). При этом любой банк, который решит установить у себя такую систему, уже не будет первоиспытателем некой «экзотики». Это промышленная платформа на открытых стандартах, она успешно работает в России, в частности в одной из ведущих российских инвестиционных компаний. На западных рынках она отлично рекомендовала себя на внушительных объемах данных.

Все решает скорость

Валерий Юринский, директор отделения технологического консалтинга компании «ФОРС», отмечает, что в некоторых случаях хранение данных средствами СУБД не является первоочередной задачей. СУБД хорошо справляются с обработкой относительно редко изменяемых данных, но что делать с потоками событий, которые представляют собой постоянно поступающие данные, чрезвычайно редко запрашиваемые повторно?

«Такие потоки могут генерироваться различными датчиками, системами GPS, инструментами мониторинга сетей, биржевых торгов, действий пользователей и т. п., — рассказывает Юринский. — Эту информацию нужно быстро обработать, подать сигнал

■ **Александр Гвоздев:** «Microsoft не производит коммерческие продукты, реализующие подход HADOOP, но предлагает средства интеграции с этими технологиями».



кому следует или выполнить некоторое заранее определенное действие. Например, если в течение любых 20 секунд цена акций „Б“ выросла более чем на 2%, а цена акций „А“ не изменилась, автоматически купить акции „А“. Задача решается при помощи особого способа обработки данных Complex Event Processing, для реализации которого корпорация Oracle, например, создала специальный язык Continuous Query Language (CQL)».

Денис Соловьев, руководитель практики SAP компании «СИТРОНИКС Информационные Технологии» в России, отмечает, что мировые производители БД постоянно развивают свои решения в части увеличения скорости обработки и объемов хранения данных. «Определенно перспективным трендом сейчас являются решения класса In-memory БД, обеспечивающие очень высокую скорость выполнения обращений к БД, но требовательные к

аппаратному обеспечению, — говорит он. — Производители пробают также использовать другие модели данных, теория которых не так полно проработана, как для реляционных БД, но при этом они обладают рядом дополнительных преимуществ».

Соловьев рекомендует работу систем в режиме реального времени базировать на архитектурных принципах комплекса информационных систем предприятия:

— Различные OLTP-системы, из которых данные загружаются в некое хранилище, способны обеспечить режим «около реального времени» (Near real-time). Но извлечение данных из транзакционной системы и загрузка в хранилище все равно требуют времени, пусть даже это происходит в течение нескольких секунд. Если БД, используемая OLTP-системой, способна играть роль хранилища данных и параллельно выдавать информацию в аналитические отчеты или приборные панели для руководителей — можно говорить о режиме реального времени. Некоторые из наших партнеров (крупные поставщики решений), уже сегодня способны предложить такие системы.

Соловьев обращает внимание на то, что необходимость в срочном принятии мер в компаниях, столкнувшихся с пробле-

мами обработки больших объемов данных, зависит во многом от требований к оперативности предоставляемых данных из ХД. «Если сбор определенного отчета из большого объема данных по времени удовлетворяет клиента, то необязательно вкладывать значительные инвестиции в использование мощных дата-центров, — подчеркивает он. — Пусть этот отчет формируется три дня, но если он нужен раз в полгода, с модернизацией можно и подождать».

Инструменты интеграции

Для решения проблемы BIG DATA большую популярность приобрели инструменты интеграции данных. По оценкам IDC, за последние пять лет прирост цифровых данных в мире составил в среднем более 50% в год, а в России, принадлежащей к числу развивающихся рынков, темпы роста объема данных еще выше. Алексей Пилипчук отмечает изменения к лучшему на рынке в целом: «Сейчас уже на стадии формулировки бизнес-задачи решается, что нужно сделать с точки зрения данных, нормативной информации, движения документов, чтобы бизнес получил требуемый результат. Один из эффективных подходов состоит в том, чтобы подобрать необходимый уровень хранения и просчитать стоимость этого хранения, исходя из IML-методик.

Лавинообразный рост данных — это только одна сторона медали. Инструменты интеграции данных не поспевают за ростом требований к производительности, которые становятся все более сложными и разнообразными. Говоря о несоответствии между предлагаемыми инструментами интеграции данных сегодня и проблемами, с которыми сталкиваются компании, Алек-

сей Пилипчук отмечает, что использование дорогих брендовых решений для устранения проблем производительности, в том числе выполнение преобразований в базе данных, зачастую оказывается неэффективным: «Компании получают высокую стоимость всего решения в целом, раздутый штат программистов и администраторов, решение лишь тактических, но не стратегических задач, что, в свою очередь, существенно тормозит бизнес».

Эксперты отмечают, что новые решения из-за высокой стоимости могут себе позволить не все компании. «Поэтому производители ищут выход во внедрении промежуточных буферов хранения информации, используя твердотельные накопители, — объясняет Владимир Кирсанов. — Это позволяет увеличить скорость доступа к наиболее актуальной в настоящий момент информации («горячие данные») и снизить время поиска в больших массивах. Наблюдается также тенденция к внедрению «умных» систем хранения, которые самостоятельно решают проблемы фрагментации данных, на уровне системы хранения выполняют их дедупликацию и т.п.

Однако до сих пор, отмечает Кирсанов, остается значительный разрыв между стоимостью подобных решений и необходимой скоростью работы системы хранения, объемом и надежностью.

— На наш взгляд, — говорит он, — сегодня стоит сосредоточить внимание на создании приложений, которые умели бы самостоятельно принимать решения: понимать, к какой базе подключиться, чтобы получить «горячие» данные, из какой — извлечь менее актуальные и нужные, а какая база является архивом. ◀