

EMC Greenplum — новое поколение BI-хранилищ

Активная интеграция BI-аналитики (business intelligence) в онлайн-бизнес-процессы и системы оперативного принятия решений — один из основных трендов ИТ-индустрии последнего времени с самыми высокими ежегодными темпами роста. В публикации рассмотрены преимущества и особенности построения современных специализированных хранилищ данных (Data Warehouse — DW) и, в частности, — EMC Greenplum Database 4.0 — одного из немногих DW-решений, активно продвигаемых на российском рынке.



Денис Серов — руководитель направления технического консультирования “EMC Россия и СНГ”.

Введение — аналитика Больших Данных

Писатель Дуглас Адамс, в книге “Автостопом по галактике” рассказывает историю о создателях суперкомпьютера, который мог дать ответ на любой вопрос. В него решили загрузить всю доступную информацию с целью получить ответ на главный вопрос жизни и вселенной. Для того чтобы загрузить и проанализировать большой объем данных, потребовалось 7,5 миллиона лет. Увы, ответ оказался бессмысленным для далеких потомков создателей компьютера. Мораль — загрузка данных и обработка сложных запросов должны выполняться быстро. Короче — “дорога ложка к обеду”.

У тех, кто знаком с проблемами традиционных BI-хранилищ (Data Warehouse), ирония автора вызовет множество ассоциаций. Удивительно, но книга, написанная более 30 лет назад, поднимает проблемы, беспокоящие ИТ сегодня, в эпоху Больших Данных.

Ситуация на современном рынке сложилась до боли похожая на сюжет Адамса. Вычислительные мощности процессоров неуклонно растут, объем и производительность систем хранения едва ли не удваиваются каждые три года, однако хранилища данных не удовлетворяют нынешним потребностям. Почему?

Дело в том, что на рынке наметился новый тренд — аналитика Больших Данных (Big Data). Его суть заключается в том, что объемы информации (структурированной и неструктурированной) достигли той критической массы, когда из них можно извлечь ре-

альную бизнес-выгоду. Разработчики ПО, оценившие потенциал этого рынка, стали предлагать решения для их анализа (“Why Are Data Warehouses Growing So Fast?” <http://www.b-eye-network.com/view/7188>)

Как только подобные предложения поступили в эксплуатацию, производительности традиционных хранилищ данных стало недостаточно. Дело в том, что архитектура большинства систем хранения построена с прицелом на транзакционные нагрузки, доминировавшие десятилетиями и еще несколько лет назад считавшиеся главным элементом нагрузки. Аналитика Больших Данных сместила баланс в сторону потоковых нагрузок. Традиционные инфраструктуры оказались к этому не готовы. Поэтому те компании, которые хотят извлечь максимальную выгоду из информации, ищут новые подходы, все больше обращая внимание на поставщиков, предлагающих аппаратные и программные решения, специализированные именно для Больших Данных.

Проблемы традиционных хранилищ данных

Медленная загрузка

Медленная загрузка — это основной бич традиционных хранилищ данных, построенных на базе транзакционных СУБД. Однако он присущ не только им. Это может быть любая монолитная система, которая загружает данные через узкое горлышко “мастера”. При ограниченных окнах для загрузки это приводит к тому, что загрузить все данные, которые хочется, сложно. В итоге, загружают не все данные, а лишь те, без которых существующим аналитическим инструментам обойтись вообще невозможно (как правило, это всего 10–15% от желаемого объема). В результате, о загрузке дополнительных данных, которые позволяют взглянуть на бизнес под новым углом, уже не задумываются. Главное — поддержать статус-кво. Такой подход рано или поздно приводит к стагнации. Извлекать новые инсайты и разрабатывать инновационные подходы в таких условиях исключительно сложно.

Если же ситуация становится невыносимой для поддержания статус-кво, создается новое хранилище данных и в него перемещается часть функций. Так, два (а затем и более) хранилища начинают жить параллельной жизнью, и часть решений

принимается на основе данных, загружаемых в одно, другая часть, соответственно — основана на другом. Это ведет к неконсистентности связанных друг с другом отчетов, и снижению качества принимаемых решений. Управление и поддержка иногда разнородных хранилищ — становится дополнительной головной болью для ИТ-служб. Неслучайно в профессиональном сообществе с прошлого века не стихают дебаты на тему “Enterprise DWH vs Data Mart” (<http://www.information-management.com/issues/19980201/815-1.html>). При этом аргументы в пользу скорости и удобства работы противопоставляются аргументам в пользу консистентности отчетов и “единой версии правды”. Идеально было бы совместить перечисленные характеристики в одной системе.

Длительная обработка запросов

Аналитические запросы принципиально отличаются своим характером от транзакционных. Им присущи сканирование таблиц, большие выборки и математические операции с большим объемом данных, в то время как транзакционные запросы больше работают с отдельными строками и выполняют чтение и вставку небольших количеств данных. Наиболее подвержены этому пороку, хранилища уходящие корнями в транзакционные СУБД, — их пользователи страдают от того, что не могут получить ответы на аналитические запросы так же быстро, как можно было бы делать при помощи специализированных систем. Но и специализированные системы так же не все одинаково успешно решают проблему обработки сложных запросов. Если архитектура обработки запросов завязана на узкое место — то в результате там рано или поздно возникнет затор. Если узкое место нельзя расширить — то проблема традиционных хранилищ проявляется на более поздней, запущенной стадии. Огромное хранилище данных, неспособное быстро обрабатывать запросы, похоже на чехол без ручки. И бросить жалко, и нести неудобно.

Невозможность мониторинга и управления производительностью

Это одна из серьезных проблем, присущая как замороженным, так и специализированным аналитическим хранили-

шам. В первом случае у администратора нет инструментария из-за того, что он принципиально не предусмотрен в системной архитектуре.

Во втором случае — управление производительностью защиты глубоко в систему и спрятано от администратора. Второй случай неплохо работал до недавнего времени и работает до сих пор, но начинает давать сбои, так как он был создан в парадигме аналитических хранилищ вчерашнего дня.

Когда же возникает проблема, то администратор как в первом, так и во втором случае может лишь констатировать факт, что хранилище работает медленно. Изменить что-либо не представляется возможным, что, конечно, вызывает недовольство бизнес-пользователей.

Ограниченная масштабируемость

Объем данных растет. Время построения отчетов ограничено. Таким образом, чтобы успевать делать отчеты, надо вместе с ростом объемов данных наращивать и скорость их обработки.

И это еще полбеды. Хранилища ограниченной масштабируемости могут банально упереться в потолок производительности контроллера дисковой системы или главного вычислительного узла. В этом случае приходится либо развертывать рядом новое аналитическое хранилище, либо искусственно уменьшать объемы хранения, снижая количество анализируемых данных до более приемлемого уровня. В обоих случаях ухудшается качество принимаемых бизнес-решений.

“Проприетарная” архитектура

Еще один подводный камень традиционных хранилищ — проприетарная аппаратная архитектура — в один прекрасный момент из благословения превращается в проклятье. Подобно заколдованному мечу, который помогал своему хозяину разить врагов, он обрушивается на него самого. Дело тут в следующем: любая аналитическая система активно развивается. С одной стороны, появляется новый функционал. С другой, — каждое предприятие ведет разработку и совершенствование своих аналитических инструментов. Однако при выборе проприетарной архитектуры хранилища на этом пути, подобно Сцилле и Харибде, встают две скалы.

Во-первых, для разработчиков нужна своя, отдельная система. Желательно, чтобы эта система была целиком идентична основной, чтобы можно было полноценно оценивать эффективность оптимизации и время выполнения запросов. **Во-вторых**, модернизация функционала хранилища не всегда возможна без замены железа, т.е. для того, чтобы получить последние “вкусности” от мирового лидера, надо заменить железо, а вместе с ним — и заново приобрести все лицензии. Что при этом делать со старой системой? Обычно ее не списывают, а продолжают использовать. Так, на предприятиях начинают плодиться островки с разными версиями истины. Это не только сложно в управлении и поддержке, но и приводит к ухудшению качества принимаемых бизнес-решений.

Неспособность обрабатывать неструктурированные данные

Традиционные хранилища данных подразумевают анализ структурированных

данных. Однако все данные привести к такой форме не всегда возможно. Благодаря интернету и технологиям Web 2.0, сейчас накопился огромный пласт данных, которые крайне сложно структурировать и едва ли возможно анализировать методами SQL. Например, общие тексты, разнообразные логи интернета, наполнение социальных сетей, блогов и т.п.

В нынешних условиях многие компании хотели бы иметь возможность оперативно извлекать выгоду из этих данных. Это позволяет, например, раньше конкурентов понять предпочтения пользователей и поведение потребителей, выйти на новые рынки, предложить новые услуги.

Все эти проблемы, в конечном счете, приводят к существенным сложностям в создании единого хранилища данных корпоративного класса. Рекомендованные окружные пути сводятся к использованию нескольких независимых хранилищ, каждое из которых обслуживает свою категорию задач, создаются т.н. витрины данных.

Как мы видим, ситуацию на рынке традиционных хранилищ можно охарактеризовать как революционную и способствующую возникновению так называемых “market disruptive” технологий — т.е. таких, которые способны в короткое время сломить сложившийся статус-кво и привести к появлению новых лидеров отрасли.

Новая эра — аналитика Больших Данных

В настоящее время все мировые вендоры активно поддерживали новый тренд — аналитика Больших Данных. Среди ключевых игроков следующие: Teradata, Oracle (Exadata), SAP/Sybase, IBM (Netezza), EMC (Greenplum), HP (Vertica). Есть также множество других разработчиков ПО, предлагающих свои решения для хранилищ данных для среднего и малого бизнеса, но они не входят в контекст аналитики Больших Данных.

EMC/Greenplum

В 2003 году, выйдя из Teradata, тогда еще безусловного лидера отрасли BI-хранилищ, основали компанию Greenplum.

Для нового продукта выбрали имя, символизирующее на Востоке конец зимы и начало нового года. Основатели уверены, что

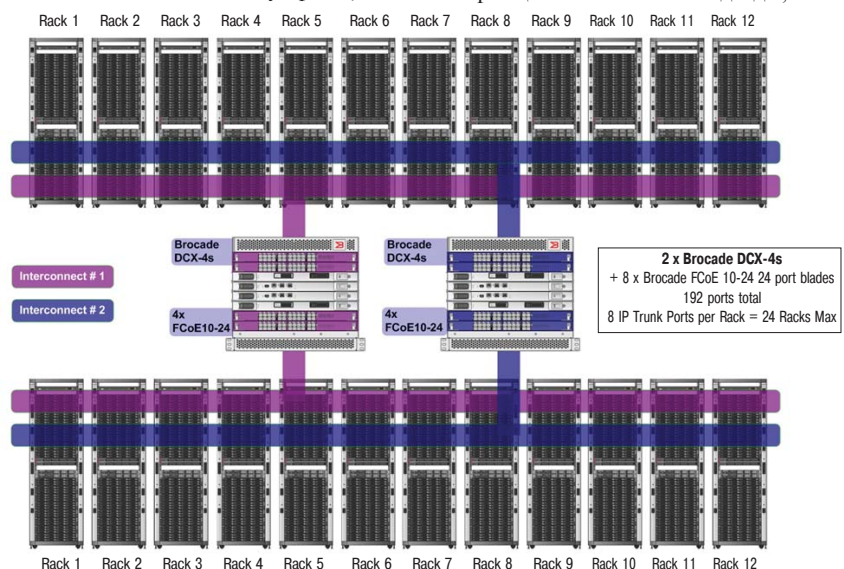
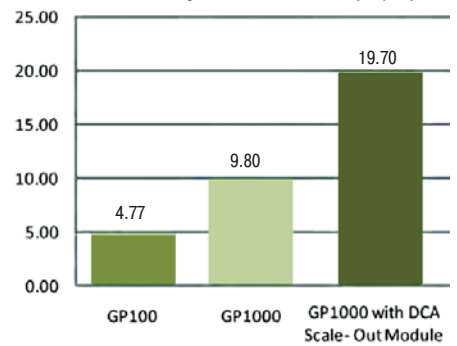


Рис. 2. Масштабируемость Greenplum практически не ограничена.

DCA scalability: Data load rates (TB/hr)



DCA scalability: Scan rates (GB/s)

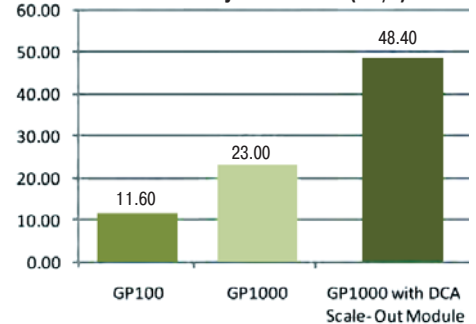


Рис. 1. Сравнительные скорости загрузки (вверху) и обработки запросов (внизу) для разных вариантов комплектации EMC Greenplum.

появлением Greenplum началась новая эра в мире хранилищ данных. Эра аналитики Больших и очень Больших Данных.

Бизнес оказался успешным, развивался очень уверенно. За 7 лет последовательно сменили друг друга 4 поколения систем. Были завоеваны позиции более чем в сотне очень крупных заказчиков во всех секторах рынка. Независимые аналитики оценили возможности предлагаемой технологии очень высоко, и назвали Greenplum ведущим визионером рынка BI-хранилищ. Единственным опасением были размер компании и ограниченность финансовых возможностей для поддержки и разработки передового продукта.

В 2010 году компания Greenplum была приобретена EMC, и начался новый этап развития, подкрепленный мощным финансированием.

Для решения комплекса проблем аналитических хранилищ, напоминающего гордиев узел, дальновидные компании выбирают принципиально новые подходы, не

подверженные наследованным болезням и лучше соответствующие реалиям сегодняшнего дня. Именно к таким технологиям относится EMC Greenplum.

Платформа Greenplum представляет собой специализированное аналитическое хранилище корпоративного класса, позиционируемое как самое быстрое из представленных на рынке. Так скорость загрузки данных в одностоечной конфигурации (GP1000) достигает 10 Тбайт/час (рис. 1), а максимальное число ядер процессора, которое может быть выделено для параллельной обработки (в конфигурации 24 стойки), составляет 4608 (рис. 2).

Преимуществами Greenplum являются:

- массивно параллельная обработка данных;
- горизонтальная масштабируемость;
- онлайн модернизация и расширение;
- полная поддержка серверов стандартной архитектуры;
- совместимость с облачными средами;
- поддержка сред обработки структурированных (SQL) так и неструктурированных данных (MapReduce);
- встроенная аналитика;
- универсальная платформа для хранилищ и витрин данных, ELT, text mining, статистического моделирования;
- совместимость с ключевыми поставщиками ПО интеграции данных и бизнес-аналитики;
- доступность для тестирования;
- низкая начальная стоимость;
- быстрое развертывание;
- функции надежности enterprise класса.
- партнерство с ведущими разработчиками ПО для анализа и интеграции данных.

Среди ключевых направлений, в которых работает EMC/Greenplum на этом этапе, можно назвать:

- развитие локальных инженерных ресурсов;
- обучение партнеров на местах;
- установление более прочных контактов с поставщиками ПО для анализа данных.

Позволяет обрабатывать данные как при помощи SQL так и при помощи MapReduce

Мастер узел строит планы выполнения запросов для получения наивысшего быстродействия

Прямой интерконнект между узлами для быстрого взаимодействия

Сегментные узлы производят обработку данных максимально близко к их дислокации

Массивно параллельная загрузка для быстрой загрузки данных

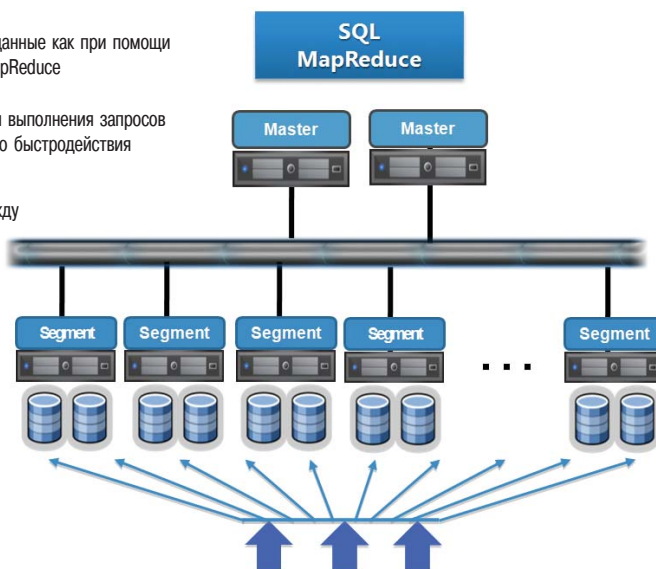


Рис. 3. Архитектура Greenplum DCA.

Архитектура Greenplum

Становым хребтом архитектуры хранилища данных является аналитическая СУБД (DBMS). База данных Greenplum – это система массивно-параллельной обработки данных, т.е., распределенная система из 2-х или более индивидуальных серверов, выполняющих операции параллельно. У каждого сервера есть свои процессор, память и диски. Все серверы общаются друг с другом по общей сети. В этом случае единая СУБД может эффективно использовать комбинарованные вычислительные ресурсы всех индивидуальных серверов, обеспечивая большую мощность и масштабируемость.

СУБД Greenplum построена на основе PostgreSQL 8.2.14 и в большинстве случаев неотличима от PostgreSQL с точки зрения поддержки SQL, опций настройки и функциональности конечного пользователя.

СУБД Greenplum способна обрабатывать хранение и обработку больших объемов данных, распределяя нагрузку по множеству серверов. При этом мастер является входной точкой, к которой клиенты подключаются и делают SQL-запросы. Обработка данных происходит на сегментных серверах.

Основными инфраструктурными компонентами архитектуры Greenplum являются следующие (рис. 3).

Мастер Сервер, принимающий и распараллеливающий пользовательские запросы. Для отказоустойчивости используются 2 сервера в конфигурации Active/Passive с синхронной репликацией метаданных.

Сегментные Серверы, выполняющие хранение и обработку данных. Для повышенной отказоустойчивости данные зеркалируются между сегментами. Всего в каждом шкафу может быть до 16 сегментных серверов. Всего может быть до 24 шкафов (маркетинговое ограничение).

Сетевой интерконнект (gNet), обеспечивающий скоординированное взаимодействие узлов друг с другом. Физический интерконнект построен на базе 10GE – коммутаторов Brocade. Для отказоустойчивости используется 2 коммутатора в каждом шкафу.

Коммутаторы соединяют мастер-серверы, сегментные серверы, и к ним также может быть подключена система резервного ко-

пирования. Важной частью интерконнекта является логика поэтапного выполнения запросов, которая позволяет осуществлять обмен данными между сегментами, а мастеру при этом возвращается окончательный ответ на запрос.

Эталонная архитектура Greenplum – Data Computing Appliance (DCA)

DCA представляет варианты стандартных поставок, которые предусматривают комплектацию полной стойки – 40U шкаф (модель GP1000), 1/2 стойки (модель GP100) и 1/4 стойки (модель GP10). В шкафу может быть 4, 8 или 16 сегментных серверов. В каждом сегментном сервере установлено по 12 дисков (SAS 600 Гбайт 10k rpm или SATA 2000 Гбайт 7.2k rpm). Диски сегментных серверов сконфигурированы в RAID5(5+1), т.е. каждый узел имеет полезную емкость от 2,6 Тбайт до 9 Тбайт несжатых зеркалированных данных. Шкаф – до 40 Тбайт. С учетом коэффициента сжатия – полезная емкость может быть примерно в 4 раза выше.

Для обеспечения отказоустойчивости как модель GP100, так и модель GP1000 комплектуются двумя мастер-серверами. В каждом мастер-сервере – 6 дисков, сконфигурированных в группу RAID5(4+1), плюс диск горячей замены.

Характеристики/состав сегментных/мастер-серверов и интерконнекта:

- сегментный сервер (Dell R510):
 - два CPU Intel X5670 2.93 GHz (6 ядер);
 - память – 48 GB DDR3 1333 MHz;
 - один FCoE-адаптер – Brocade CNA 1020 (2 порта x 10Gbps);
 - один NIC – 2 порта x 1Gbps;
 - один RAID-контроллер – Dual channel 6 Gb/s SAS;
 - 12 HDD x 600 GB 15k rpm SAS;
- мастер-сервер (Dell R610):
 - два CPU Intel X5680 3.33 GHz (6 ядер);
 - память – 48 GB DDR3 1333 MHz;
 - один FCoE-адаптер – Brocade CNA 1020 (2 порта x 10Gbps);
 - один RAID-контроллер – Dual channel 6 Gb/s SAS;
 - один NIC – 4 порта x 1Gbps;
 - 6 HDD x 600 GB 10k rpm SAS;
- межузловой интерконнект:
 - два 32-портовых коммутатора (Converged Enhanced Ethernet – CEE, Fibre Channel over Ethernet FCoE) – 24 x 10 GbE порта и 8 FC порта в каждом;
 - один коммутатор для администрирования – 24-порта x 1Gb.

Всего может быть до 24 40U шкафов, т.е. более 10 Пбайт (см. рис. 2). Максимальная известная продуктивная инсталляция Greenplum – хранилище с полезной емкостью 6 Пбайт в очень большом интернет-магазине.

Преимущества подхода EMC

Высокая скорость загрузки и обработки запросов

Архитектура Массивно Параллельной Обработки (MPP – Massive Parallel Processing), позволяет давать ответы даже на сложные запросы с максимальной скоростью. Достигается это следующим обра-

Табл. 1. Сравнительные характеристики двух типов эталонных архитектур Greenplum

	GP1000 (шкаф)	GP1000C (шкаф)
Мастеров	2	2
Сегментов	16	16
Ядер ЦПУ*	192	192
Память*	768 GB	768 GB
Диски*	192 SAS 600 GB 10к грп	192 SATA 2000 GB 7.2к грп
Полезная емкость (не сжатая)	36 TB	124 TB
Полезная емкость (со сжатием)	144 TB	496 TB
Сканирование	24 GB/с	14 GB/с
Загрузка	10 TB/час	10 TB/час

* Без учета мастер-серверов

зом: пользовательские запросы приходят на мастер-сервер, который анализирует структуру запроса и составляет оптимальный по стоимости план его выполнения (т.н. cost based query optimization). После этого запрос разбивается на более мелкие подзапросы, которые рассылаются по максимальному количеству сегментных узлов. Вся дальнейшая обработка происходит на сегментных узлах, где данные хранятся максимально близко к процессорам. План обработки запроса может быть поэтапным, т.е. разбитым на несколько шагов. На последнем шаге, каждый узел возвращает ответ на свою часть запроса мастер-серверу, который дает цельный ответ на изначальный запрос пользователю. Такая архитектура обработки запросов не имеет узкого места. С ее помощью можно получить практически любую скорость обработки, добавляя новые сегментные узлы. Каждый узел при этом сканирует данные со скоростью от 0,9 до 1,5 Гбайт/с.

Сравнительные скорости загрузки и обработки запросов для разных вариантов комплектации EMC Greenplum см. рис. 1.

В табл. 1 приведены сравнительные характеристики архитектур с разными типами дисков (SAS и SATA).

Стандартная архитектура

Программное обеспечение Greenplum может быть развернуто на любых серверах стандартной архитектуры.

Это открывает перед потребителями целый ряд возможностей. *Во-первых*, можно развернуть демо-систему просто скачав из интернета программное обеспечение Greenplum. *Во-вторых*, потребитель не привязывается жестко к производителю серверов. В качестве поставщиков аппаратной платформы для Greenplum могут быть выбраны любые производители (IBM, HP, Dell, и другие). *В-третьих*, данная программная архитектура полностью готова к существованию в облачной среде (т.е. является Cloud Ready).

Неограниченная масштабируемость объема и производительности

Добавление новых узлов приводит к пропорциональному росту не только объема, но и производительности всего комплекса. Крупнейшая инсталляция Greenplum в мире более 6 Пбайт.

Полиморфное хранение

Традиционные реляционные БД хранят данные строчного типа. Этот метод уходит корнями в OLTP системы. Однако аналитические СУБД имеют другой способ обращения с данными. Вместо того чтобы выполнять чтение и запись множества отдельных строк, они обрабатывают большие и более сложные запросы, захватывающие большие объемы данных. В основном, это чтение сканирующего типа и крайне редкая пакетная конкатенация.

Существует множество подходов, позволяющих оптимизировать скорость такого типа доступа. Основные – колоночное хранение с вертикальной декомпозицией (Sybase IQ и другие) и оптимизация на уровне дисковых операций (Exadata и другие). СУБД Greenplum позволяет выбрать любой из этих вариантов – для каждой таблицы (или ее раздела) DBA может выбрать хранение, выполнение и коэффициент сжатия, наилучшим образом соответствующие типу доступа к данным (рис. 4).

Это позволяет достичь в Greenplum максимальной эффективности используемой емкости при сохранении оптимальной скорости обработки данных.

Обработка больших объемов неструктурированных данных

Парадигма параллельной обработки больших объемов неструктурированных данных – MapReduce, а также ее реализация Apache Hadoop полностью поддерживаются в Greenplum. Доступными примерами такой обработки является программа GREP, выполняемая параллельно на десятках или сотнях узлов, чтобы быстро обработать петабайтные объемы данных.

Другие примеры – параллельная индексация больших объемов текста, анализ логов посещения, статистический анализ пользователей с разными атрибутами, поиск ключевых фраз. При этом может выполняться массово параллельная аналитика как данных находящихся на внешних источниках (в том же интернете), так и данных находящихся на распределенной файловой системе (Hadoop Distributed File System).

Популярно использование языка Python для создания основных логических модулей MapReduce – собственно MAP- и REDUCE-функций. Эта популярность обусловлена простотой разработки на этом языке, а также богатыми библиотеками, наиболее яркой, из которых является Python NLTK (Natural Language Tool Kit), позволяющая производить синтаксический и, порой, даже смысловой анализ поистине огромных объемов текста. В отличие о других реализаций

MapReduce фреймворка, реализация его в Greenplum позволяет использовать одновременно два популярных сегодня подхода – SQL и NoSQL, – используя гибридные конструкции из SQL выражений и MapReduce заданий.

Это открывает перед многими компаниями новые возможности для повышения своей конкурентоспособности и быстрому выходу на новые рынки.

Встроенная аналитика

Для ведущих аналитических приложений существует возможность выгрузки части математической логики на уровень хранилища. Таким образом, есть возможность избежать излишней передачи данных и разгрузить процессоры аналитического сервера. В итоге, это дает ускорение ответа на наиболее сложные запросы, такие как, например, кластерный и регрессивный анализ, нейронные сети и т.д. Наиболее свежий пример – SAS Scoring Accelerator.

"Кто ест Зеленые Сливы"?

Среди пользователей Greenplum есть весьма выдающиеся представители всех секторов рынка. Вот лишь некоторые публичные примеры:

- большие телекоммуникационные операторы (TMobile, Nokia Siemens Network, Reliance Telecom);
- крупные финансовые институты (Zions Bancorporation, Deutsche Bank, NASDAQ, NYSE и ФРС США);
- ведущие розничные сети (Walmart);
- интернет-гиганты (Skype);
- государственные организации (Офис Планирования Бюджета США, Финансовый Регулятор США).

Как попробовать?

Greenplum DataWarehouse существует в исполнении Community Edition, доступном любому зарегистрировавшемуся пользователю. Скачать полнофункциональную версию можно с сайта сообщества: <http://community.greenplum.com>.

Можно скачать Greenplum в виде сконфигурированной виртуальной машины, чтобы упростить процесс развертывания и сразу после скачивания приступить к функциональному тестированию.

Нагрузочное тестирование можно провести как с себя на площадке, так и в Центре Демонстрации Решений EMC.

Внедрение

Внедрение происходит по выбору заказчика, либо силами EMC, либо силами заказчика. Первые внедрения в России делали у себя энтузиасты из ведущей российской инвестиционной компании. Внедрение прошло успешно, и организация давно и успешно использует Greenplum для биржевой аналитики.

Эксплуатация

Для многих пользователей, в итоге, именно преимущества в скорости загрузки и обработки затмевают все и склоняют в сторону выбора Greenplum.

Однако как и любая система, Greenplum требует к себе некоторого внимания со стороны администратора. Так что аспект управляемости тоже очень важен.

Загрузка данных

Для этого лучше использовать профессиональные ETL-инструменты, такие как Informatica. Однако Greenplum может



Рис. 4. Пример модели гибкого хранения в рамках одной таблицы.

обходиться и без таковых. Для ускорения загрузки данных можно использовать механизм ELT (Extract Load Transform), который происходит значительно быстрее, чем ETL, так как трансформация уже загруженных данных выполняется значительно более мощной аппаратной и математической средой узлов Greenplum.

Мониторинг производительности

В Greenplum очень хорошие возможности мониторинга и оптимизации под любые нагрузки, которые превосходят то, что могут предложить некоторые другие системы (например, Teradata).

Помогает лучше понимать систему инструментарий Performance Monitor, собирающий информацию со всех сегментов при помощи агентов. Он представляет собой browser-based приложение, которое позволяет наблюдать производительность на уровне системных метрик и деталей выполнения запросов. Предлагаемая панель приборов позволяет мониторить утилизацию системы в реальном времени, параллельно с выполнением пользовательских запросов. При этом существует возможность углубиться в план выполнения запроса и лучше его понять.

Управление качеством обслуживания

Разумеется, мониторинг производительности Greenplum дополняется возможностями оптимизации и управления вычислительными ресурсами. Это жизненно необходимо для любого хранилища корпоративного класса, на которое приходится множество запросов от разных по важности групп пользователей и бизнес-подразделений.

Для того чтобы обеспечить необходимый уровень сервиса, Greenplum позволяет разграничивать приоритеты выполнения запросов на ролевой основе. Назначая роли различным группам запросов, администратор может контролировать число одновременных запросов того или иного типа, чтобы избежать чрезмерной загрузки системы. Приоритизация влияет на долю вычислительных ресурсов (процессор, память, дисковые операции), которые может использовать та или иная "ресурсная очередь" (рис. 5).

Резервное копирование и катастрофоустойчивость

Бизнес-критичная аналитика нуждается в регулярном резервном копировании. Здесь принципиальным моментом является все та же архитектурная особенность, что и при загрузке данных — полный параллелизм. Каждый сегмент-узел Greenplum мо-

жет "заливаться" на устройство резервного копирования параллельно с остальными. Это может дать максимальную скорость, если, конечно, целевое устройство сможет выдержать такой темп записи. Но у EMC для такой цели есть свой "рояль в кустах" — скоростная система резервного копирования DataDomain, которая, кроме всего прочего, умеет дедуплицировать данные по ходу их резервного копирования и реплицировать их на другую площадку, чтобы обеспечить катастрофоустойчивость.

Для создания консистентной резервной копии серверы, составляющие кластер Greenplum, по установленному расписанию выполняют одновременный согласованный дамп своих данных на систему резервного копирования, как на точку монтирования NFS.

Резервное копирование может выполняться в горячем режиме, т.е. в это же время могут обслуживаться пользовательские запросы. Далеко не все ВІ-хранилища могут похвастаться функцией быстрого резервного копирования и катастрофоустойчивости.

Миграция данных

В некоторых случаях может иметь смысл мигрировать данные из старого хранилища в Greenplum. Данные могут быть перенесены на Greenplum при помощи свободно распространяемых или коммерческих инструментов, таких, например, как Informatica (универсальный ETL инструмент) или Wisdomforce (специализируется на извлечении данных из Oracle).

Однако в большинстве случаев заказчики предпочитают оставить имеющееся хранилище данных и установить Greenplum в качестве перспективного, быстрого и масштабируемого дополнения.

Во всех случаях степень сложности миграции или организации сосуществования двух разных хранилищ зависит от конкретной реализации. Могут быть отличия на уровне загрузочных инструментов (ETL), а также на уровне построения отчетов. Оценка степени усилий может быть легко произведена благодаря простоте развертывания пилотного стенда Greenplum.

Миграция с Teradata

Проще всего происходит миграция с хранилищ Teradata. Секрет прост — главный архитектор Greenplum — выходец из Teradata. Поэтому миграция практически автоматизирована.

Миграция с Netezza

Хранилище Netezza, так же как и Greenplum, является ответвлением от PostgreSQL, поэтому синтаксис у этих хранилищ довольно схожий. Миграция на Greenplum облегчается, благодаря возможности прямого копирования данных (команда COPY) с Netezza на Greenplum.

Примеры внедрений

Пример 1 — финансовый сектор

Профиль компании: крупная банковская сеть, объединяющая 17 банков с более чем 500-ми офисами в 10-ти регионах страны. Стратегия роста компании состоит в активном использовании Больших Данных — большой объем данных из разных источников — для того чтобы

взаимодействовать с заказчиками разными путями и создавать спрос на новые услуги, увеличивая таким образом финансовые потоки.

Проблема. Возросший объем данных, разнообразие запросов и предъявленные требования к скорости загрузки и обработки данных оказались выше возможностей традиционного хранилища. Директор по бизнес-аналитике говорит: "Наше монолитное хранилище с трудом справлялось с обработкой растущих объемов данных, поэтому нам пришлось разбить его на витрины данных меньшего размера, которые начали быстро плодиться повсюду. Эта децентрализация привела к тому, что аналитики начали испытывать трудности с оценкой рисков, связанных с бизнес-операциями. Выполнение аналитических запросов длилось часами. Иногда — всю ночь. Администраторы проводили большую часть времени, решая возникающие проблемы."

Решение. Результатом внедрения Greenplum стало значительное ускорение загрузки и анализа данных. Хранилище Greenplum обеспечивает скорость обработки, требуемую более 600-ми пользователями, включая 70 продвинутых пользователей, которые выполняют анализ и создают отчеты ежедневно. Бизнес-данные загружаются в хранилище Greenplum за 30 минут в сравнении с 24 часами на предыдущей системе.

"Раньше аналитические запросы занимали часы, а Greenplum выполняет ту же работу за несколько секунд или минут. Это делает продвинутых пользователей гораздо более продуктивными. Фактически, они получили на 30% больше времени на изучение и анализ результатов отчетов. Это оказывает прямое влияние на повышение качества принимаемых ими решений."

"Быстрая обработка данных, которую обеспечивает Greenplum, делает возможным выполнение глубокого и упреждающего анализа специфического поведения клиентов. Результаты таких исследований служат основой для разработки новых предложений, маркетинга и повышения объемов продаж. Это дает выход на совершенно новый уровень взаимоотношений с клиентами (CRM)."

Сегодняшний бизнес компании сильно завязан на систему оценки доходности — приложение, работающее на Greenplum для расчета прибыльности займов, инвестиций или депозитов клиентов.

"Когда разразился банковский кризис, мы использовали систему оценки доходности для анализа исторических данных и создания новых наборов предположений для наших менеджеров по работе с займами. Обработка данных глубиной в месяц занимала семь часов. Сейчас, используя Greenplum, мы можем "перемолотить" данные за 2 года примерно за то же время (7,5 часов)."

Интересно заметить, что развертывание системы анализа доходности было невозможно до внедрения Greenplum по той простой причине, что весь штат администраторов был занят поддержкой традиционных аналитических хранилищ данных.

Раньше все, чем мы занимались, — было решение проблем и поддержка. С Greenplum же, мы можем не только нормально функционировать без больших затрат времени на поддержку, но и прикладывать больше усилий на развертывание новых бизнес-приложений".

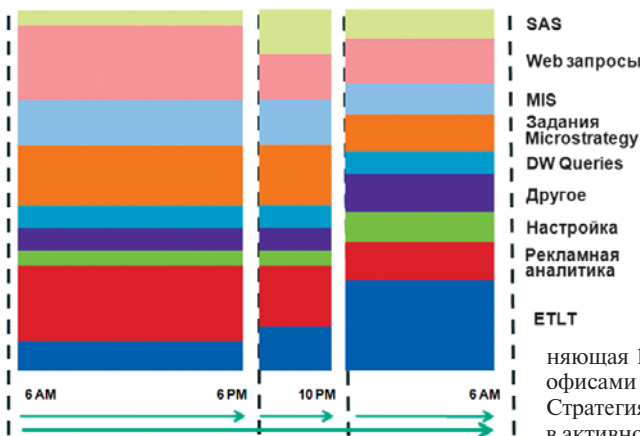


Рис. 5. Пример динамического управления качеством сервиса по расписанию.

Пример 2 — телекоммуникационный сектор

Профиль заказчика. Крупная телекоммуникационная компания с более чем 40 миллионами абонентов, с темпами роста — 50% в год. Компания быстро добавляла новые сервисы и функции к своим традиционным предложениям мобильной связи с целью увеличить долю рынка в среде жесткой конкуренции.

Проблема. Рост спроса привел к взрывному росту количества систем и инфраструктуры, необходимой для управления бизнесом. Нужда предоставлять точную и своевременную аналитику всем бизнес-подразделениям становилась все более острой. Одна из частей компании, требовавшая быстрого предоставления качественной аналитики, — Департамент Контроля Соблюдения Закона. Телекоммуникационная компания должна быстро давать детальные записи телефонных разговоров по запросам органов правосудия и государственной безопасности. Слишком быстрый рост абонентской базы и неспособность традиционных систем справиться с возрастающими требованиями привел к тому, что компания перестала успевать вовремя предоставлять необходимые отчеты. Удовлетворение запросов стало занимать до нескольких дней. Даже загрузка данных за прошедший день требовала порядка 2-3 часов. Для того чтобы обеспечить соответствие законодательству, компания нуждалась в качественно новом решении.

Решение. Компания перешла на хранилище Greenplum. Аппаратная часть состояла из 4 серверов SunFire x4500, с 20 ядрами и 192 дисками SATA. Полезная емкость — 40 Тбайт без сжатия.

Внедрение Greenplum позволило сократить время выполнения запросов в 5 раз. Кроме решения проблемы извлечения информации, загрузка также ускорилась с 2-3 часов до 10 минут. Развертывание аналитики на платформе хранилища Greenplum улучшило время отклика и снизило риск несоответствия законодательным требованиям.

Внедрение Greenplum принесло настолько впечатляющие плоды и настолько быстро, что компания решила расширить систему, включив в нее другой важный набор данных — информацию о трафике абонентских соединений, так называемые Call Data Record (CDR). Эти записи представляют собой информацию о том, кто, когда, кому звонил. Они используются для мониторинга SLA и биллинга между операторами связи.

Сейчас система поддерживает более 60 Тбайт полезной емкости, 200 SATA дисков и 28 процессорных ядер с приростом — 1 Тбайт данных в неделю.

В нашей стране

Рынок аналитики Больших Данных активно осваивается как молодыми, так и старыми игроками. Например, сейчас наибольшим опытом в интеграции Greenplum с задачами обладает компания World IT Systems, которая была партнером Greenplum еще до приобретения его компанией EMC. Специалистами World IT systems успешно реализовано несколько проектов, а также множество заказчиков уже тестируют Greenplum у себя в ЦОДах при поддержке World IT systems и очень

довольны впечатляющими результатами. В результате тестирования одним из заказчиков конфигурации Greenplum DCA GP10 (1/4 Rack), было получено ускорение реальных задач загрузки — от 33 до 118 раз, а аналитики — от 53 до 113 раз, по сравнению с традиционным информационным хранилищем.

Заключение

Greenplum дает компаниям с разными объемами бизнеса реальную возможность осуществить “золотую мечту” — извлечь реальную выгоду из накопленных данных. Благодаря своей эластичной масштабируемости и скорости работы с данными Greenplum позволяет реализовать мечту о Единой Версии Правды, которая повышает качество и в разы увеличивает скорость принятия бизнес-решений. Greenplum это:

- консолидация аналитических хранилищ без ущерба в скорости загрузки и времени анализа данных;
- отказоустойчивость, необходимая для аналитического хранилища корпоративного класса;
- доступность для скачивания и тестирования разработчиками, партнерами и просто интересующимися;
- возможность приобретения его в виде ПО, не требующего лицензирования для тестовых окружений, что выделяет его среди конкурентов;
- доступность по начальной цене даже для небольшой компании, решившей создать свое первое хранилище данных.

Денис Серов, руководитель направления технического консультирования “EMC Россия и СНГ”

IBM: прорыв в хранении

Июнь 2011 г. — Ученые из исследовательского подразделения IBM Research корпорации IBM впервые продемонстрировали, что относительно новая технология памяти, известная как память с изменением фазового состояния (Phase-Change Memory, PCM), может надежно хранить несколько бит данных в одной ячейке в течение длительного времени. Это позволяет создавать более быструю, надежную и в то же время недорогую память для целого спектра применений — от потребительских устройств до высокопроизводительных центров хранения информации корпоративного уровня.

Благодаря сочетанию быстродействия, износостойкости, энергонезависимости и плотности записи, память PCM способна внести коренные изменения в технологии хранения данных. Ученые давно искали технологию универсальной энергонезависимой памяти с гораздо большей производительностью, чем у флэш-памяти. Преимущества такой технологии памяти могли бы позволить компьютерам и серверам загружаться практически мгновенно и значительно повысить эффективность работы ИТ-систем. Перспективным претендентом на роль такой технологии является память PCM, которая записывает и считывает данные в 100 раз быстрее, чем флэш-память, обеспечивая высокие показатели емкости хранения и не теряя данные при выключении питания. В отличие от флэш-памяти, PCM-память чрез-

вычайно надежна и долговечна в работе — она может выдержать не менее 10 млн циклов перезаписи. Для сравнения: современные флэш-модули корпоративного класса выдерживают около 30 тыс. циклов перезаписи, а флэш-модули потребительского класса — 3 тыс. циклов перезаписи (http://www.flickr.com/photos/ibm_research_zurich/5881252148/in/set-72157626714358795).

LSI вернут в SAS и DAS

Июнь 2011 г. — В конце июня прошел круглый стол компании LSI с Брентом Бланшаром, директором по продажам и маркетингу LSI, и Александром Зейниковым, представителем компании в России и СНГ, посвященный последним достижениям и стратегии компании на ближайшее время.

Большую часть доходов LSI — 80% — приносит обслуживание ИТ-инфраструктуры. За последние 5 лет LSI потратила более \$4 млрд, сделав 9 полезных приобретений для своего бизнеса. В 1 кв. 2011 г. уровень дохода компании достиг \$473 млн. Представители компании отметили, что уже более 28 млн продуктов SAS (IC и адаптеров систем хранения), 750 тыс. адаптеров RAID на 6 Гбит/с, 7 млн MegaRaid было продано по всему миру, а 9 из 10 крупных OEM-производителей используют 6 Гбит/с LSI SAS RoC и ПО MegaRAID. Из последних новинок — контроллеры MegaRAID SAS 9265/85. По словам Брента, эти контроллеры в четыре раза производительнее конкурентов. Также Брент обратил внимание на WarpDrive — первые твердотельные накопители компании. Это серверные PCIe решения с SLC-памятью, обладающие очень высокой производительностью. Кроме того, отдельного упоминания заслужили технологии CacheCade и FastPath. Первая позволяет создавать на контроллерах LSI “гибридные” массивы, в которых твердотельные накопители используются в качестве кэш-памяти. Вторая — FastPath — предназначена для оптимизации работы контроллера при использовании массивов, состоящих исключительно из твердотельных дисков.

Что касается стратегии на ближайшее будущее, то здесь Бланшар выделил несколько основных направлений. Разумеется, компания планирует и далее развивать свой основной бизнес, сделав акцент на совместимости и защите данных и коммуникаций SAS. Также среди целей LSI — продвижение систем хранения данных на базе твердотельных накопителей. Кроме того, компания верит в будущее систем DAS, которые в последнее время становятся все более актуальными. Для поддержки и развития этого направления в прошлом году ею был выпущен SAS-коммутатор.



Брент Бланшард (Brent Blanchard) — директор по продажам и маркетингу компании LSI.