

Оригинальный материал: <http://www.sleepycat.com/products/overview.html>

Перевод: Павел Макаров (pavel.makarov@mail.ru)

Разнообразные потребности, большой выбор СУБД – ответ на специфические задачи управления данными

Официальная публикация компании Sleepycat Software – разработчика Berkeley DB

Berkeley DB представляет собой встраиваемую систему управления базами данных (СУБД) масштаба предприятия, которая значительно отличается от тяжелых клиент-серверных реляционных механизмов баз данных («движков»), что может принести пользу вашему бизнесу.

Поставщики технологий сегодня выбрасывают на рынок ошеломляющее количество товаров для самых различных потребителей. Некоторые компании, такие, как Nokia и Motorola, производят сотовые телефоны и остальную ширпотреб для частных лиц. Телекоммуникационные компании, включая Cisco и British Telecom, разрабатывают системы, обеспечивающие функционирование критически важной инфраструктуры для телефонии и передачи данных. Компьютерные компании, такие, как Compaq, Hewlett-Packard и EMC, строят системы и устройства хранения данных, являющиеся основой корпоративных IT-подразделений. Вэб-услуги от компаний, таких, как Amazon и Ask Jeeves, обеспечивают потребителей услугами Интернет-коммерции и поисковыми услугами.

Каждый из этих поставщиков, как и тысячи подобных им компаний по всему миру, тратят время и деньги на разработку изделий, соответствующих конкретным запросам их заказчиков. Все они сталкиваются с жесткой конкуренцией и постоянным давлением со стороны рынка, требующего совершенствования предлагаемой продукции.

Все программные и аппаратные изделия, создаваемые этими компаниями, так или иначе вынуждены осуществлять управление данными. Требования к базам данных могут быть довольно простыми (управление записной книжкой в сотовом телефоне), равно как и чрезвычайно сложными – слежение за высокоскоростными и постоянно изменяющимися данными о пропускной способности и топографии в коммутаторе передачи данных.

Умные поставщики приобретают СУБД, которые им нужны, а не разрабатывают их самостоятельно. И руководители, которым поручено принять такое решение о приобретении, прежде всего должны выяснить:

- Какие СУБД существуют?
- Какие из них будут лучше соответствовать запросам пользователей на данном рынке?

Структурные и конструктивные отличия

И клиент-серверные реляционные «движки», и Berkeley DB предлагают функции СУБД масштаба предприятия. Оба могут предложить основные функции по хранению и восстановлению данных, необходимые поставщикам технологий для изделий, разрабатываемых ими для своих заказчиков. Berkeley DB, однако, отличается от клиент-серверных реляционных СУБД по нескольким существенным направлениям. Понимание этих отличий является ключевым для наилучшего выбора СУБД для проекта.

Клиент-серверные СУБД часто работают как огромные хранители различных данных. Администратор базы данных для такой системы просто выполняет роль прислужника, помогая руководителям компаний тем, что поддерживает СУБД в рабочем состоянии, предоставляя необходимую информацию лицам, принимающим решения в компаниях.

Berkeley DB разработана для использования профессиональными программистами, создающими

программное обеспечение (ПО), требующее управления данными. Механизм базы данных («движок») достаточно мал для того, чтобы работать в сетевых маршрутизаторах и коммутаторах, но достаточно мощен для того, чтобы управлять критически важными задачами хранения и восстановления. Как программный инструмент, созданный разработчиками ПО для разработчиков ПО, Berkeley DB обеспечивает механизмы, необходимые программистам для создания высокопроизводительных информационных систем.

Все СУБД должны предоставлять заказчикам надёжную, масштабируемую функциональность транзактного типа (см. таблицу «Сравнение характеристик»). Насколько СУБД может обеспечить эти характеристики, зависит от типа работающей СУБД.

SQL: «гвоздь программы» или препятствие?

Большие реляционные СУБД современности выросли из коммерческой задачи управления и анализа громадных объёмов информации. Реляционные СУБД делают возможным написание бизнес-приложений, которые могут управлять большими объёмами сложно структурированной информации без того, чтобы заранее представить, что будет нужно запросам для нахождения хранимой информации.

Эти реляционные СУБД используют интерфейс языка структурированных запросов (structured query language – SQL), позволяющий пользователям, не являющимся программистами, запрашивать и получать информацию из базы данных. Встроенная в SQL гибкость велика в том случае, если вы не можете заранее предсказать тип информации, которую вы пожелаете извлечь из базы данных.

Для таких же задач обработки, когда бизнес может прогнозировать свои потребности в хранении и восстановлении информации, Berkeley DB является, несомненно, наилучшим «движком», который можно выбрать.

Заказчики Berkeley DB используют и опираются на те же самые возможности манипулирования данными, которые используются и в реляционных СУБД. Их приложения, однако, не нуждаются в предельной гибкости (и дополнительной стоимости) уровня SQL. Заказчики Berkeley DB взамен приобретают скорость обработки и улучшенную производительность, поскольку их системы не тормозятся громоздкими процедурами выборки данных, включёнными в каждый SQL-процесс.

Если СУБД может быть не обременена SQL-уровнем, ускорение обработки может быть гигантским. Это не является откровением для заказчика, уже перешедшего с приложения на основе реляционной СУБД на приложение на основе Berkeley DB для более, чем тысячекратного выигрыша в производительности системы.

Администрирование СУБД может не понадобиться

Исключение SQL часто означает и исключение или сокращение затрат на администрирование СУБД. Приложения с Berkeley DB редко нуждаются в администрировании. Во многих случаях применения Berkeley DB пользователь может даже и не осознавать, что СУБД является частью изделия; много ли пользователей сотовых телефонов думают о поиске таблицы данных во время поиска номера в собственной записной книжке? Berkeley DB функционирует без резервирования, установок и настроек со стороны пользователя.

Мировой лидер во встроенных СУБД

Berkeley DB имеет 12-летний опыт адаптации СУБД под требования заказчиков из списка Fortune 100. СУБД используется в самых различных приложениях, как больших, так и маленьких. Можно упомянуть только часть задач, связанных с обработкой данных, которые Berkeley DB выполняет для различных направлений бизнеса: безопасность/идентификация и обработка виртуальных покупок на сайтах электронной коммерции, серверы электронной почты Интернет, управление NAS и SAN накопителями в системах запоминающих устройств, а также управление маршрутизаторами и коммутаторами в системах управления передачей данных.

Начиная со своего первого появления в 1991 году, Berkeley DB обеспечивает превосходное управление базами данных, продолжая эту традицию в коммерческих версиях от Sleepycat Software. Имея более 200

тысяч внедрений, Berkeley DB является лидирующей встраиваемой СУБД в мире.

Требования надёжности, восстанавливаемости и масштабируемости являются обязательными для любой СУБД – как реляционной, так и встраиваемой. Все успешные «движки» должны обеспечивать одинаковые возможности по автоматической классификации и интерпретации данных (automated classification and interpretation of data –ACID) и использовать идентичные, ясные и выверенные десятилетиями технологии.

Berkeley DB обеспечивает управление данными масштаба предприятия во встраиваемых приложениях. Отказоустойчивость и целостность данных гарантируются. Механизм базы данных – «движок» – надёжен и никогда не теряет данные из-за сбоев в приложении, системе или оборудовании. ПО может легко масштабироваться для поддержки тысяч одновременно работающих пользователей и терабайтов информации.

Улучшенная производительность и высочайшие скорости

Простота кода Berkeley DB является главным источником надёжности системы. Berkeley DB мала и элегантна. Она предоставляет те же возможности, что и реляционные СУБД, но она реализует эту функциональность совершенно иначе. Вместо того, чтобы создать низкоуровневый реляционный «движок», затем добавить обработку кучи запросов и сверху добавить поддержку логической структуры данных (schema support), Berkeley DB реализует необходимые транзакции через библиотеку с простым интерфейсом вызова функции (function call API).

Реляционные СУБД включают четыре отдельных компонента: синтаксический анализатор запроса, планировщик, оптимизатор и исполнительный модуль (executor). Задачей этих четырёх компонентов является превращение текстовой строки SQL-запросов в физические операции чтения-записи диска. Berkeley DB позволяет заказчикам обходить всю эту обработку и получать доступ непосредственно к байтам данных, которые им необходимы. Такое архитектурное решение значительно сокращает объём и время исполнения кода и улучшает производительность изделия в целом. Стоимость эксплуатации (runtime costs) и потребность в ресурсах также минимизируются.

Ещё одна причина быстрого действия вытекает из встроенной архитектуры Berkeley DB. «Движки» клиент-серверных реляционных СУБД обычно работают отдельно от связанных с ними прикладных программ; часто СУБД и приложения вообще находятся на различных машинах. Выход из процесса для каждой операции выборки или сохранения записи является по сути своей очень громоздким и ограничивает общую производительность приложения. Berkeley DB, наоборот, работает в адресном пространстве приложения. Это конструктивное отличие устраняет межпроцессное взаимодействие (IPC) и переключение контекста, опираясь вместо этого на простые вызовы функций.

Лучшая адаптируемость

В отличие от клиент-серверных реляционных СУБД, Berkeley DB очень легко настраивается. Berkeley DB предоставляет разработчикам систем и программистам потрясающую гибкость в выборе и конфигурировании опций каждой подсистемы. Любая подсистема, в которой приложение не нуждается – будь то регистрация, транзакция или восстановление – может быть исключена. «Движок» предоставляет программистам контроль над установками, такими, как размер записи, синхронизация операций записи на диск (timing of disk writes) и выделение памяти. Такое «вооружение» людей, пишущих приложения, позволяет создавать компактные и быстрые системы и, как следствие, более успешные внедрения.

Berkeley DB извлекает пользу из своего двойного наследия – как из принципа открытых «исходников» (open source), так и из статуса коммерческого продукта масштаба предприятия. Открытый исходный код Berkeley DB был и продолжает тестироваться и улучшаться гораздо большим числом программистов, чем это может себе позволить любая сопоставимая коммерческая СУБД. Как продукция компании Sleepycat, Berkeley DB обеспечена всеми ресурсами рентабельной компании, включая её деятельность по сопровождению и разработке. Sleepycat содержит штат профессионалов для расширения возможностей продукта, написания тестовых наборов, устранения обнаруженных ошибок и создания документации, что позволяет Berkeley DB оставаться современным и мощным продуктом.

Berkeley DB имеет длинную историю внедрений в изделия для ответственных коммерческих применений.

Как подтвердит вам любой программист, реальные ошибки находятся только в процессе эксплуатации изделия. Имея за плечами годы успешных внедрений и тысячи работающих систем, Sleepycat уже нашла и исправила эти ошибки.

Двойная лицензия предоставляет выбор

Berkeley DB распространяется под лицензией, которая разрешает её бесплатное использование в приложениях с открытым исходным кодом. Коммерческие поставщики могут приобрести пожизненную лицензию (lifetime license) у Sleepycat Software, что гарантирует им неограниченное использование Berkeley DB в их приложениях. Заказчики могут также платить дополнительно ежегодную фиксированную сумму за поддержку. Эта плата даёт им право на получение новых версий того пакета Berkeley DB, который они лицензировали, и включает базовую техническую поддержку.

Редко когда имеет смысл платить за функциональность, которая вам не нужна. И *никогда* не имеет смысла замедлять обработку за счёт компонентов, которые вы не используете.

Sleepycat Software и Berkeley DB доказывают, что для задач обработки данных маленькое всегда может быть лучше и быстрее. Посмотрите и решите сами, что Berkeley DB может предложить для улучшения ваших приложений с базами данных.

Сравнение характеристик

	Berkeley DB	Клиент-серверные реляционные «движки»
<i>Архитектура</i>	Встроенная в приложение библиотека	Отдельный сервер
<i>Взаимодействие</i>	Быстрые локальные вызовы функций	Громоздкие межпроцессные сообщения для каждой выборки, модификации
<i>Интерфейс</i>	Лёгкие в использовании вызовы функций или методов (function or method calls)	Мощные, но сложные SQL-запросы, которые сервер должен сначала проанализировать и интерпретировать
<i>Для кого СУБД предназначена</i>	Профессиональный разработчик систем, создающий высокопроизводительные приложения для последующей продажи	Корпоративный IT-департамент, эксплуатирующий СУБД общего назначения в интересах коммерческих пользователей
<i>Администрирование</i>	Никакого администрирования со стороны пользователя после внедрения не требуется	Требуется администрирование СУБД для резервного копирования, восстановления, мониторинг производительности и настройки
<i>Масштабируемость</i>	Терабайты данных, тысячи одновременно работающих пользователей	Терабайты данных, тысячи одновременно работающих пользователей
<i>Надёжность</i>	Управление данными масштаба предприятия	Управление данными масштаба предприятия
<i>Конфигурируемость</i>	Любая ненужная подсистема может быть удалена	Никакой гибкости в настройке под индивидуальные требования заказчика
<i>Стоимость</i>	Никаких авторских отчислений с продаж для серийных поставок (fixed-cost distribution) вне зависимости от объёма продаж	Цена за копию, за пользователя или за количество процессоров (per-copy, per-user or per-CPU); может быть невыгодна при крупномасштабных продажах
<i>Исходный текст</i>	Всегда включён	Зависит от поставщика

Эта таблица сравнивает характеристики Berkeley DB и клиент-серверных реляционных «движков».