

Системы автоматизированного проектирования

УДК 004.031.43

М.В. Дидковская, канд. техн. наук, В.В. Нарыжный

Особенности построения мультимедийной системы управления базами данных

Проведен анализ мультимедийных СУБД, определены основные требования к ним.

Рассмотрено два подхода к организации мультимедийных СУБД – с использованием файловых серверов и без.

Экспериментальное исследование показало целесообразность использования файловых серверов для хранения мультимедийной информации, так как этот подход обеспечивает большую скорость записи и чтения данных.

The analysis of multimedia databases is carried out, the main requirements were pointed out.

Two approaches to the organization of multimedia databases were considered – with usage of file servers and without.

The experimental study has shown the practicability of file servers' using for storage of multimedia information, because this approach provides higher speed of data read and write.

Введение

На сегодняшний день информация является одним из наиболее ценных ресурсов, поэтому к системам ее хранения предъявляются повышенные требования. Полнота информации, скорость доступа, легкость поиска – одни из главных характеристик таких систем [1]. С появлением высокоскоростных способов передачи данных, традиционная текстовая информация начала дополняться мультимедийной: фото, аудио и видео материалами [2]. Базы данных для хранения мультимедийных файлов были выделены в специальный класс мультимедийных СУБД (ММ СУБД) [1]. На сегодняшний день нет единого подхода к вопросу построения ММ СУБД, работа с мультимедийными данными требует от базы данных более сложного функционала по сравнению с обработкой текстовой информации.

В связи с этим целями данной работы является изучение вопросов организации ММ СУБД и проведение экспериментального исследования для анализа различных подходов к хранению мультимедийных данных.

1. ММ СУБД: требования, функции, анализ подходов

Определим основные требования к ММ СУБД:

- оперирование с мультимедийными объектами (изображения, аудио, видео файлы), поддержка форматов данного вида файлов;
- наличие высокопроизводительного хранилища большой емкости;
- поддержка всех средств реляционных СУБД, таких как транзакции, логирование, версионность, безопасность данных;
- обеспечение средств информационного поиска: ассоциативный поиск, поиск по содержанию, сортировка результатов поиска, извлечение шаблонов для индексации и поиска;
- поддержка многомерного индексирования мультимедиа объектов;
- наличие структуры хранения больших двоичных файлов с возможностью поблочного доступа к данным.

Кроме вышеуказанных характеристик ММ СУБД должна обеспечивать ряд функций для упрощения работы клиентских приложений, а именно [3]:

- наличие иерархического представления ММ объектов;
- поддержка метаданных;
- обеспечение совместного использования данных несколькими пользователями;
- навигационный доступ и интерактивное уточнение запросов.

Большинство ММ СУБД являются реляционными [4,5]. Несмотря на то, что это наиболее распространенный подход, при работе с мультимедиа объектами он имеет ряд недостатков. Процесс нормализации ведет к разбиению объекта – объект представляется в виде большого количества записей в нескольких таблицах; для работы с целым объектом (например, извлечение из БД) каждый раз требуются дорогостоящие операции JOIN. Организация доступа к

мультимедиа объекту осуществляется с помощью хранимых процедур и пользовательских функций, встроенного функционала в реляционных базах не предусмотрено [4].

В свою очередь, объектно-ориентированные (ОО) базы данных обеспечивают явное моделирование отношений «часть-целое» и напрямую позволяют обращаться к объекту. Более того, некоторые ООСУБД поддерживают изъятие объекта из хранилища с блокировкой оригинала и копирование обратно после модифика-

ции (т.е. синхронизацию check-out/check-in) для составных объектов [5].

К сожалению, современные средства не полностью соответствуют всем требованиям ММ СУБД в плане поддержки объектно-ориентированного подхода, алгоритмов поиска и скорости каналов.

2. Организация мультимедийной СУБД

Схематически простейшую ММСУБД можно представить таким образом (Рис. 1):

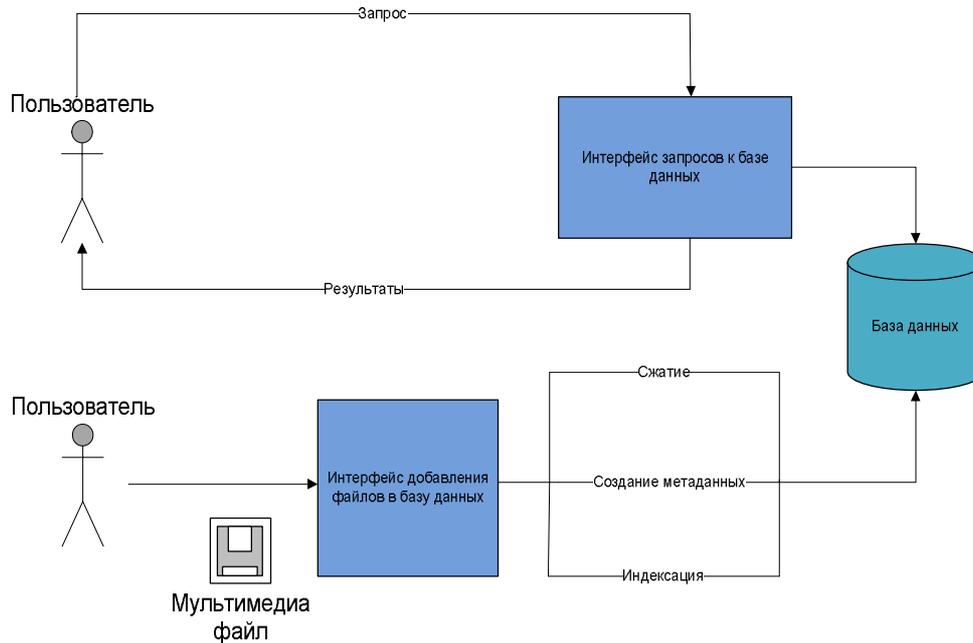


Рис. 1. Схема мультимедийной СУБД

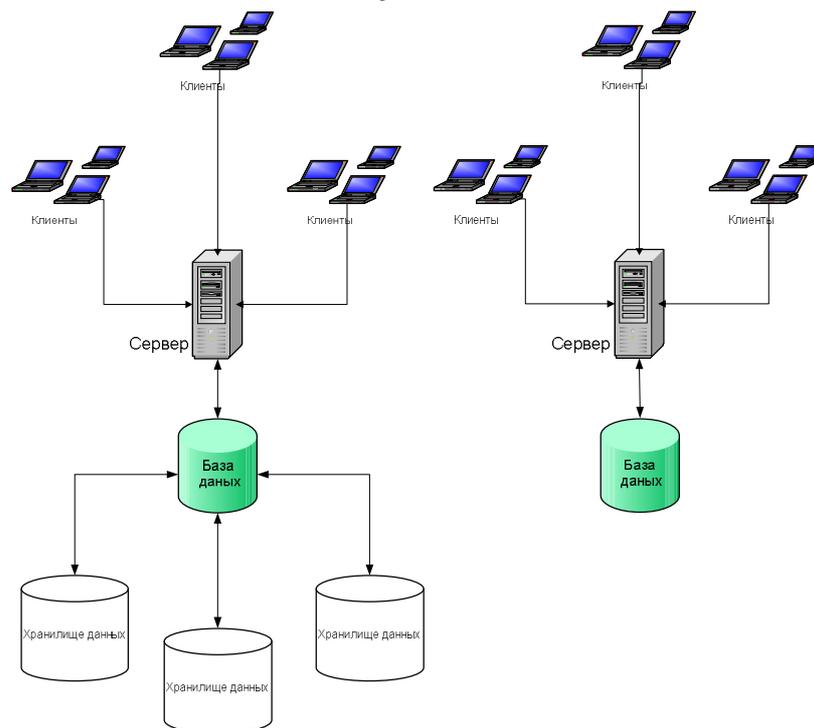


Рис. 2. Подходы к хранению данных в ММ СУБД

Хранение данных может быть организовано следующим образом (рис. 2):

- как поля в мультимедийной базе данных;
- как отдельные файлы на файловых серверах, ссылки на которые (а также их описания) содержатся в базе данных.

Для реализации второго подхода предлагается использовать клиент-серверную систему. Сервер выполняет стандартные функции СУБД, а клиент работает с сервером приложений через пользовательский интерфейс. Сервер выполняет все функции по обработке мультимедиа-файлов: сжатие, создание метаданных, индексирование, запись в базу данных. Сервер принимает файлы по сети, после чего сохраняет их в базу данных.

Сравним два вида организации (рис. 2) мультимедийной базы данных: с использованием отдельного хранилища мультимедийных объектов и без него.

3. Сравнительный анализ подходов к хранению данных в мультимедийной СУБД

В качестве характеристик сравнения подходов предлагается выбрать: удобство хранения информации, безопасность, масштабируемость, конфигурируемость и скорость доступа [6].

1. *Удобство хранения.* При размещении файлов на отдельном файловом сервере, при наличии доступа к нему, файлы могут быть легко открыты, скопированы или удалены. Хранение файлов непосредственно в БД предполагает необходимость использования дополнительного функционала, потому что в базе данных файлы хранятся в двоичном виде. Таким образом, хранение мультимедиа-объектов на файловом сервере является более удобным.

2. *Безопасность.* При наличии нескольких файловых серверов возникает необходимость настройки системы безопасности таким образом, чтобы только приложение и пользователи с соответствующими правами могли получить доступ, что создает дополнительную задачу для администраторов сети и разработчиков приложения. Задача обеспечения безопасности базы данных остается в обоих случаях, но при хранении файлов в базе данных её ценность возрастает по сравнению с вариантом хранения файлов на файловых серверах.

3. *Масштабируемость.* При росте количества запросов к базе данных и ее объемов вопрос производительности может стать критичным, так как запись и передача большого количества файлов создает серьезную нагрузку для

файловой подсистемы сервера базы данных и сетевого канала. В то же время наличие нескольких файловых серверов позволит равномерно распределить эту нагрузку, но требования к сетевым каналам останутся и появятся дополнительные требования к производительности файловых серверов, так как недостаточная производительность одного сервера может существенно замедлить всю систему.

4. *Конфигурируемость.* Конфигурируемость характеризует возможность системы соответствовать своей среде и потребностям приложений. При использовании файловых серверов, кроме стандартных требований к конфигурируемости СУБД возникают дополнительные, так как необходимо настроить права доступа ко всем файловым серверам и их возможное взаимодействие. Кроме того, при необходимости перемещения файлов на другой файловой сервер возникают следующие проблемы: непосредственное перемещение файлов и необходимость перенаправления ссылок на новый файловый сервер.

Таким образом, с точки зрения конфигурируемости, организация СУБД с использованием файловых серверов является более сложной.

5. *Скорость.* Для сравнения скорости записи и чтения данных при использовании рассматриваемых подходов осуществим экспериментальное исследование.

4. Экспериментальное исследование

Проведем сравнение скорости считывания системы с использованием файловых серверов и систем без них.

В качестве СУБД предлагается взять Microsoft SQL Server 2005. Выбор обуславливается тем, что данная СУБД не имеет ограничений на размер таблицы и имеет тип данных Image для записи больших объемов данных. Тип Image поддерживает объекты размером до 2 GB.

Сравним результаты записи и считывания при обоих подходах (рис 3, рис 4, рис 5). При этом кривые 1 (на рисунках обозначаются квадратами) отображают время записи(считывания) для систем с использованием файловых серверов, а кривые 2 (на рисунках обозначается ромбами) – время записи(считывания) для систем без использования файловых серверов.

Таким образом, при использовании архитектуры с файловыми серверами наблюдается значительное сокращение скорости реакции системы.

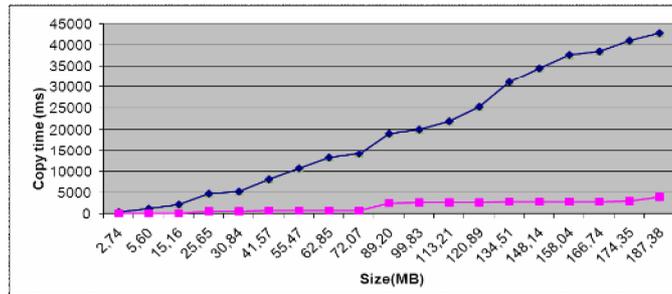


Рис. 3. Графики зависимости времени записи от размера ММ объекта

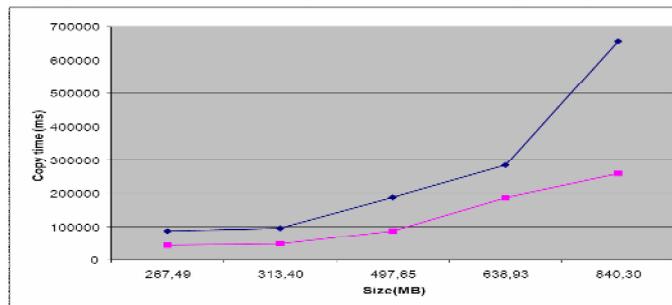


Рис. 4. Графики зависимости времени записи от размера ММ объекта (продолжение)

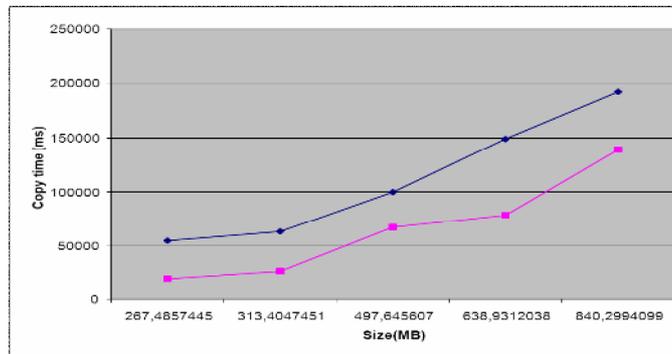


Рис. 5. Графики зависимости времени считывания от размера ММ объекта

Выводы

1. Определены основные требования к мультимедийным СУБД, выделены функции и подходы, используемые для обеспечения работы с мультимедиа объектами.

2. Рассмотрено два варианта хранения мультимедиа объектов в базах данных. Проанализирован подход к организации мультимедийной СУБД с использованием файловых серверов.

3. Проведено экспериментальное исследование ММ СУБД с файловыми серверами и без них. Сравнительный анализ показал, что в общем случае следует выбирать вариант построения системы с файловыми серверами, потому что такой вариант обеспечивает большую скорость записи-чтения данных.

Литература

1. *Коннолли Т., Бегг К., Страчан А.* Базы данных: проектирование, реализация и сопровождение. 2-е издание. – М.: Издательский дом "Вильямс", 2000. – 1120 с.
2. *Дидковский В.С., Продеус А.Н., Дидковская М.В.* Акустическая экспертиза каналов речевой коммуникации. – Киев 2008. – 420 с.
3. *Баженова И.Ю.* Основы проектирования приложений баз данных. – М.: Изд-во "Интернет-университет информационных технологий – ИНТУИТ.ру", 2006. – 320 с.
4. *Харрингтон Дж. Л.* Проектирование реляционных баз данных. – Лори, 2006. – 230 с.
5. *Харрингтон Дж. Л.* Проектирование объектно-ориентированных баз данных. – ДМК пресс, 2001. – 272 с.
6. *Марков А.С., Лисовский К.Ю.* Базы данных. Введение в теорию и методологию. – М.: Финансы и статистика, 2004. – 512 с.