

УДК 004.416.6

И.Г. Боровской, Д.П. Вагнер

Система автоматического восстановления настольных БД

Рассмотрены вопросы, связанные с повреждением файла базы данных MS Access при работе в многопользовательском режиме, проведен анализ методов восстановления. Предложен собственный способ и разработан программный комплекс автоматического восстановления настольных баз данных.

Ключевые слова: настольные БД, восстановление баз данных, MS Access, ODBC, OLE DB, Windows Socket.

Введение

В настоящее время базы данных являются неотъемлемой частью большинства информационных систем. Естественно, масштабы использования баз данных могут варьироваться в широком диапазоне, например, автоматизация промышленного производства или разработка интернет-проекта, рассчитанного на работу с сотнями пользователей, должны основываться на БД соответствующего уровня, таких как Oracle, MS SQL и т.п. Использование подобных БД связано не только с проведением регулярных процедур мониторинга, администрирования и управления, но также и со значительными затратами на приобретение соответствующего программного обеспечения (ПО), а также последующую установку и сопровождение. Однако во многих случаях такая полная функциональность не является востребованной. Для такого рода предприятий, как государственные (муниципальные) учреждения, предприятия сферы образования или сферы обслуживания, организации малого и среднего бизнеса, может оказаться вполне достаточным использование класса БД, именуемых «настольными» [1].

Настольные базы данных, несмотря на своё название, обычно содержат в себе полный комплекс необходимых инструментов для создания достаточно мощных систем хранения и управления данными, с возможностью комфортного использования в сетевой многопользовательской среде. Нередко настольные базы данных используются и как универсальное вспомогательное хранилище данных, управляемое внешней информационной системой. При этом функциональность, связанная с созданием визуального интерфейса приложений, может оставаться не задействованной. С другой стороны, часто может встречаться ситуация, обратная описанной выше, т.е. сама настольная БД выступает в качестве визуального интерфейса, осуществляющего лишь внешнее представление данных, хранение которых возложено на другую, возможно, более мощную СУБД. В качестве примера такого взаимодействия можно привести связку – MS SQL Server и MS Access [2].

Естественно, что такие базы данных имеют и определенные недостатки, которые особенно сильно проявляются при увеличении числа обращений и росте объема данных. В таком случае многие из функций управления конфиденциальностью, целостностью данных, решение проблемы сбоев и восстановления должны чаще всего реализовываться непосредственно при разработке информационной системы. Например, вполне актуальной может стать задача периодического автоматизированного резервного копирования данных при возникновении каких-либо внешних проблем с БД либо проведение различных профилактических работ (сжатие, удаление пустых записей и т.п.).

Постановка задачи

Одной из перечисленных проблем является задача автоматического восстановления базы данных MS Access при работе в сетевом многопользовательском режиме. Особенно, связанные с функционированием баз данных MS Access, основываются на том, что вся работа пользователей происходит непосредственно с файлом MDB, который может быть расположен как на локальной машине, так и на удаленной. Также стоит отметить, что работа MS Access основана не на принципе клиент-серверных приложений, с отправкой запроса серверу и возврату результата клиенту, а на принципе взаимодействия в рамках одноранговой сети, когда одна машина выступает в качестве файл-сервера для БД, а другие – осуществляют к ней подключение. Вся реальная работа, таким образом, происходит в адресном пространстве клиентского приложения, и при выполнении запросов данные, на основании которых выполняется такой запрос, должны быть доставлены в то же самое адресное пространство клиентского приложения. В таком случае при монополь-

ной работе клиента или небольшом количестве одновременных подключений к файлу вероятность возникновения проблемных ситуаций невелика, однако при росте объема данных и числа пользователей проблема повреждения файла базы данных может иметь регулярный характер [3].

Дело в том, что Microsoft Jet – ядро СУБД Microsoft Access – является системой с совместным доступом к файлам. Если Microsoft Jet работает в многопользовательской среде, операции чтения, записи и блокировки совместно используемой базы данных выполняются многими клиентскими процессами. Из-за возможности доступа многих клиентских процессов к одной базе данных, а также из-за того, что в Microsoft Jet не используется журнал транзакций (в отличие от СУБД более высокого класса, таких как MS SQL Server), надежно защитить базы данных от повреждения практически невозможно.

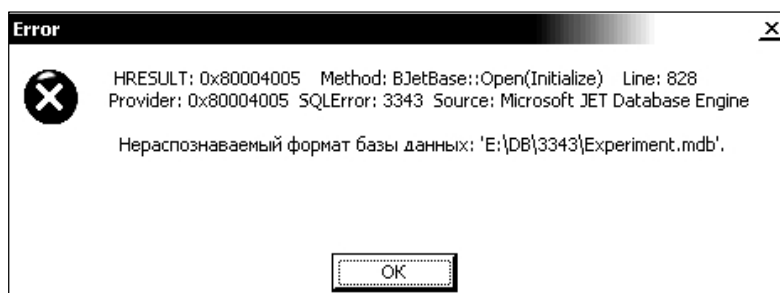
Проблема повреждения файла базы данных MS Access может явиться следствием различных факторов:

– **Повреждение базы данных вследствие прерывания операции записи.** При неправильном завершении работы Access во время записи данных в открытую базу данных ядро СУБД Jet может отметить базу данных как подозрительную или поврежденную. Это может произойти, например, в результате сбоя в энергоснабжении. В других ситуациях завершение работы Access может не производиться, однако при записи данных в файл базы данных на жестком диске все равно возникает ошибка. Это может произойти, например, из-за конфликта данных в сети или сбоя жесткого диска.

– **Неисправное сетевое оборудование.** В данном случае повреждение файла связано с какой-либо внешней причиной. Например, причиной повреждений могут являться связи в цепочке оборудования между компьютером, на котором расположена база данных, и компьютером, на котором она открыта. Неисправными могут оказаться платы сетевого интерфейса, сетевые кабели, маршрутизаторы, концентраторы и другое оборудование.

– **Открытие и сохранение файла MDB в другой программе.** В случае открытия и сохранения файла MDB различными программами, не ориентированными на работу с базами данных, не существует способов восстановления .mdb файла. Единственно возможным решением будет создание резервной копии файла.

Как видно, существует несколько причин, как внутренних, так и внешних, не зависящих от MS Access, вследствие которых файл базы данных может быть поврежден. В большинстве случаев эти причины носят случайный характер, и не могут быть экспериментально описаны. Признаки повреждения базы данных могут быть разные. Для пользователя, уже подключенного к неисправной базе данных, это может вызвать как появление пометок на удаление корректных записей, так и невозможность открыть какой-то объект в базе данных. При попытке нового подключения к базе данных признаком неисправности будет невозможность установить соединение к базе данных вовсе. Например, при использовании утилиты, специально разработанной для тестирования и



отладки БД, в случае открытия поврежденного файла БД появится сообщение, изображенное на рис. 1.

Рис. 1. Пример сообщения о повреждении файла БД

В сообщении указывается, что при открытии указанной БД программным путем провайдер Microsoft Jet возвращает ошибку №3343 «Нераспознаваемый формат базы данных», причины возникновения которой четко не определены даже корпорацией Microsoft [4].

Таким образом, авторами была поставлена задача произвести анализ существующих методов восстановления поврежденных БД, на основе которого разработать систему автоматического восстановления настольных БД при работе в многопользовательском режиме.

Способы восстановления

В настоящее время единственным инструментом, предлагаемым разработчиком для восстановления поврежденных файлов баз данных Microsoft Access, являются стандартные утилиты сжатия и восстановления (Compact, Repair) объекта Microsoft Access Jet Engine.

Описание алгоритма восстановления базы данных, предлагаемое фирмой Microsoft, хотя и носит общий характер, во многих случаях помогает пользователю решить проблему.

- Необходимо сделать резервную копию поврежденного файла базы данных (.mdb).
- Закрывать все имеющиеся соединения к базе данных.
- Удалить вспомогательный файл .ldb, содержащий информацию о всех пользователях, которые установили соединение с базой данных, и позволяющий определить, какие записи заблокированы в совместно используемой базе данных и каким пользователем.
- Выполнить функцию восстановления базы данных. Если функция не была выполнена успешно, будет получено сообщение, говорящее об этом. В этом случае ущерб настолько серьезен, что не может быть исправлен простыми методами и необходимо использование методов ручного импорта из поврежденного файла во вновь созданный [5].

Кроме рекомендаций фирмы Microsoft по использованию стандартных инструментов восстановления, пользователь имеет возможность применить альтернативные программные продукты для решения проблемы:

- MDB Repair;
- Access Recovery;
- Access Fix;
- Advanced Access Repair и др.

Стоит, однако, отметить, что большинство программных продуктов данного направления представляют собой лишь модификацию стандартных функций сжатия или восстановления, предлагая пользователю возможно более удобный в практическом плане инструмент решения проблемы повреждения файла базы данных.

Конечно, использование рекомендаций Microsoft либо альтернативного программного обеспечения в большинстве случаев помогает решить проблему восстановления поврежденного файла БД. Однако при этом пользователю приходится сталкиваться с рядом ограничений и неудобств при использовании тех или иных средств.

Выполнение первых трёх пунктов алгоритма восстановления кажется элементарным в случае, если к базе данных подключены 1–2 пользователя, поскольку у программиста БД на их проведение уйдет минимальное количество времени. Теперь необходимо рассмотреть случай, когда все пользователи базы данных находятся в разных местах здания, не исключено, что и в разных зданиях. Во-первых, большинство из них до определенного момента даже не будут подозревать об имеющихся проблемах, продолжая работать, ведь симптомы повреждения могут проявиться не сразу, да и пользователь, не знакомый детально с процессами, происходящими в БД, может не обратить внимание на некоторые сбои. Далее, после идентификации имеющейся неполадки, администратору будет необходимо произвести отключение всех соединений, на что в данном случае также может уйти значительное количество времени. Как видно, время от момента реального повреждения БД до её полного восстановления может варьироваться от нескольких минут до нескольких часов, а в ситуации, когда будет отсутствовать сам администратор БД, работа базы данных может быть вообще остановлена на неопределенный срок. Не исключен и тот факт, что при наличии неисправного сетевого или локального оборудования повреждение базы данных может носить регулярный характер вплоть до того момента, пока не будет определен источник проблемы. В таком случае все операции придется повторять неоднократно [3].

Система автоматического восстановления настольных БД

Целью данной работы является автоматизация предложенного корпорацией Microsoft алгоритма по восстановлению базы данных.

Основная задача работы состоит в предоставлении центральной компоненте разработанной системы монопольного доступа к файлу MDB на основе алгоритмов своевременного оповещения клиентов о повреждении, а также непосредственное проведение операций восстановления БД.

Ключевым моментом данной работы является замена шага ручного оповещения администратором всех клиентов БД при возникновении повреждения, т.е. автоматизация всех действий администратора, связанных с подготовкой к восстановлению базы данных.

Описание работы программного комплекса

Блок-схема работы программного комплекса «Система автоматического восстановления БД» представлена на рис. 2.

Программный комплекс разработан в среде визуального программирования MS Visual Studio 6.0 и состоит из двух приложений: центральной и клиентской компонент. Центральная компонента устанавливается на машине, на которой физически расположен файл БД. В принципе это условие не является обязательным, и вполне допустим вариант, когда центральная компонента может быть расположена на другой машине, осуществляя все операции по сети, однако при этом вероятность возникновения ошибочных

ситуаций существенно возрастает. Клиентская компонента устанавливается на машинах пользователей, осуществляющих подключение к файлу БД.

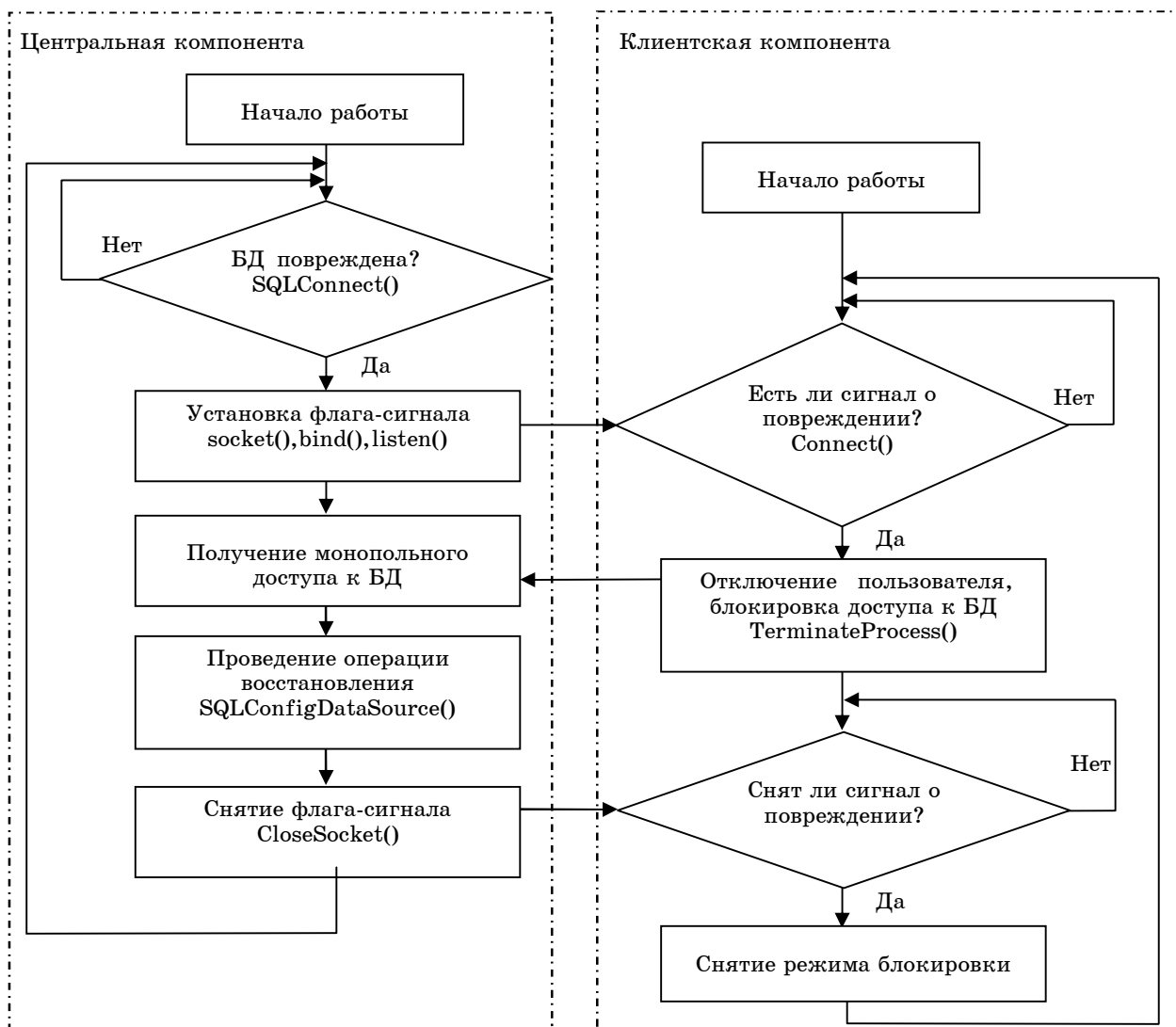


Рис. 2. Блок-схема работы программного комплекса

Центральная компонента не имеет внешнего интерфейса, поскольку функционирует в качестве NT-сервиса, что позволяет использовать её в режиме непрерывной работы, а также осуществлять корректное управление запуском, приостановкой и выключением с помощью стандартных механизмов управления.

Основными функциями центральной компоненты являются:

1. Мониторинг целостности базы данных в режиме реального времени.
2. В случае обнаружения факта повреждения файла БД – информирование об этом остальных участников процесса.
3. Восстановление файла базы данных в монопольном режиме.

Рассмотрим подробнее каждую функцию.

1. Мониторинг целостности состоит в постоянной проверке состояния БД – для этого раз в секунду происходит обращение к БД с помощью функции SQLConnect, в которой одним из параметров является имя необходимого источника данных (DSN). При работе БД в нормальном режиме функцией будет возвращено нулевое значение, после чего, с небольшой задержкой, операция повторится. В том случае, если произошел сбой в БД, новые подключения к ней будут невозможны, а потому функция вернет значение SQL_ERROR.

Поскольку подключение осуществляется только для получения возвращаемого значения функции SQLConnect(), после этого происходит отключение от БД функцией SQLDisconnect().

Наблюдение за файлом БД, а также последующее его восстановление необходимо осуществлять с использованием эффективной технологии доступа к данным. Наиболее распространенными и популярными на сегодняшний день являются следующие: Data Access Objects (DAO), Open Database Connectivity (ODBC), OLEDB и ADO.

Технологии OLEDB и ODBC в данном случае являются наиболее оптимальным вариантом, объединяя в себе свойства универсальности и эффективности, к тому же интерфейс ODBC является ещё и кроссплатформенным и с успехом работает и в Windows, и в UNIX/Linux, и в MacOS. Сравнивая OLEDB и ODBC, стоит отметить тот факт, что изначально компанией Microsoft развивалась и поддерживалась технология ODBC, завоевавшая большую популярность среди разработчиков, однако в настоящее время ей на смену пришла более универсальная и к тому же полноценно объектно-ориентированная OLEDB. Тем не менее в случае использования данных технологий в рамках решаемой задачи выбор одной из них лежит целиком на личных предпочтениях и опыте разработчика, поскольку функции и методы в общем и целом одинаково эффективны и равнозначны при использовании [6].

2. После обнаружения неисправности программный комплекс должен оповестить остальных участников процесса о невозможности работы с БД. Для этой цели чаще всего используются программные средства сетевого взаимодействия (сетевые API), такие как именованные каналы (Named Pipes) или сокететы (Windows Sockets). Использование технологии Named Pipes связано с одним ограничением: создание именованных каналов возможно только в NT-системах, хотя подключение к созданному каналу возможно как в NT-системах, так и в Win9x, тогда как в технологии Windows Socket данная проблема отсутствует. Для организации стабильного процесса обмена информацией между сервером и клиентами системы было решено использовать технологию Windows Sockets с использованием протокола TCP [7].

В данном случае центральная компонента в качестве флагов использует два сокета на локальной машине. Использование функций для работы с сокетами, таких как `socket()`, `bind()`, `listen()`, `closesocket()` и т.п., позволяет управлять состоянием канала связи и сообщать клиентским компонентам о состоянии БД и самой центральной компоненты, открывая и закрывая канал по мере необходимости. Первый сокет является контрольным и сообщает о работоспособности приложения «центральная компонента» всем клиентам, второй – непосредственно информационным, и в этом случае успешное подключение по второму порту (сокету) будет сигналом клиенту о повреждении базы данных.

3. Ещё одной функцией центральной компоненты является проведение операций по восстановлению файла БД, включающих в себя резервное копирование файла MDB, удаление файла-сопровождения ldb, а также непосредственно саму процедуру восстановления.

Перед проведением процедуры восстановления прежде всего обязательно необходимо создать резервную копию БД, которая поможет в тех случаях, когда либо восстановление невозможно стандартными средствами, либо в других критических ситуациях, например на случай сбоя в компьютере во время восстановления.

Восстановление базы данных происходит с помощью ODBC-функции `SQLConfigDataSource`, в которой для описания используемого драйвера задается строка "Microsoft Access Driver (*.mdb)", а в качестве строки атрибутов – "REPAIR_DB=полный_путь_к_файлу_БД.mdb\0\0" [6].

После проведения всех процедур, связанных с восстановлением БД, центральная компонента отключает сигнал о повреждении с помощью функции `closesocket()`, тем самым оповещая клиентов о возможности продолжения работы, а сама возвращается к режиму мониторинга БД.

На клиентскую компоненту возложены функции по обеспечению своевременного отключения пользователей, работающих с базой данных удаленно, а также блокировка новых подключений:

1. Контроль сигналов центральной компоненты.
2. Отключение пользователя от БД при появлении сигнала и блокирование новых подключений.

Клиентская компонента, установленная на пользовательском ПК, также реализована в виде NT-сервиса и запускается автоматически при загрузке операционной системы. Далее организуется периодическая проверка наличия сигнала о повреждении по заданному адресу центральной компоненты. Для этого используется функция `connect()`, параметрами которой являются ip-адрес и порт удаленной машины-сервера. Первоначально происходит подключение по контрольному порту с целью проверки работоспособности самой центральной компоненты, в случае успеха производится попытка подключения по информационному порту. Успешная попытка подключения по второму порту как раз и яв-

ляется сигналом о повреждении БД и необходимости немедленного завершения работы с БД. При обнаружении сигнала приложение автоматически прерывает работу процесса msaccess.exe с использованием функции TerminateProcess() и пользователю выводится сообщение о необходимости закрытия базы данных и повторного подключения через определенный промежуток времени. Тем временем клиентская компонента будет продолжать отслеживать состояние сокетов на сервере, и в том случае, когда контрольный сокет будет продолжать работать, а информационный – закроется, только тогда клиентская компонента снимает режим блокирования работы с БД и переходит к нормальному состоянию.

Таким образом, клиентская часть комплекса, установленная на машинах пользователей, освобождает администратора БД от необходимости самостоятельного оповещения, а также производит отключение от MDB-файла в автоматическом режиме, тем самым обеспечивается монопольный доступ к поврежденному файлу для центральной компоненты.

Заключение

В результате исследования был произведен анализ проблемы повреждения файла БД MS Access и указаны сложности, возникающие в процессе восстановления БД при работе в многопользовательском режиме. Для решения проблемы была разработана автоматизированная система восстановления настольных баз данных, позволяющая обеспечить на основе технологии Windows Socket устойчивое и своевременное выполнение процедуры оповещения и отключения пользователей БД. Кроме того, на основе технологии доступа к данным ODBC реализована процедура восстановления БД.

Также стоит отметить тот факт, что практическая реализация и внедрение разработанного программного комплекса позволило в значительной степени сократить как временные, так и финансовые затраты при работе с настольными БД на предприятиях г. Томска вследствие того, что ручное восстановление файла БД после сбоя было полностью устранено.

Литература

1. Microsoft Corporation Microsoft Access 2000. Шаг за шагом: практ. пособие; пер. с англ. – М.: Изд-во ЭКОМ, 2002. – 352 с.
2. Виллариал Б. Программирование Access 2002 в примерах: пер. с англ. – М.: КУДИЦ-ОБРАЗ, 2003. – 496 с.
3. Вагнер Д.П. Восстановление настольных баз данных / Д.П. Вагнер // Электронные средства и системы управления. Опыт инновационного развития: Докл. междунар. науч.-практ. конф., 31 окт. – 3 нояб. 2007 г. – Томск: В-Спектр, 2007. – Ч. 2. – С. 215–218.
4. Unrecognized database format <filename>. Error (3343) [Электронный ресурс]. – Режим доступа: <http://msdn.microsoft.com/en-us/library/bb223146.aspx>, свободный.
5. Восстановление поврежденной базы данных Access 2002 или более поздней и устранение неполадок в ее работе [Электронный ресурс]. – Режим доступа: <http://support.microsoft.com/kb/283849/>, свободный.
6. Роберт Сигнор, Михаэль О. Стегман. Использование ODBC для доступа к базам данных: пер. с англ. – М.: Бином; Научная книга, 1995. – 384 с.
7. Вагнер Д.П. Координатор распределенных действий для БД в многопользовательской среде / Д.П. Вагнер, И.Г. Боровской // Доклады ТУСУРа. – 2008. – № 1 (17). – С. 74–78.

Боровской Игорь Георгиевич

Д-р физ.-мат. наук, профессор,
зав. каф. экономической математики, информатики и статистики ТУСУРа

Вагнер Дмитрий Петрович

Аспирант каф. экономической математики, информатики и статистики ТУСУРа
Тел.: 8-903-951-52-73
Эл. почта: adios@lenta.ru

I.G. Borovskoi, D.P. Vagner

Automatic desktop DB recovery system

The issues connected with MS Access DB-file corruption during multi-user sessions are considered, the analysis of data recovery methods is given. The author's recovery method and the developed software complex for automatic Desktop DB recovery are described.

Keywords: desktop DB, database recovery, MS Access, ODBC, OLEDB, Windows Socket.