

Утилиты для работы с БД Системы ИРБИС

The utilities for database of IRBIS

Д.Ю. Копытков, А.С. Карауш
Муниципальная информационная библиотечная система г.Томска

Dmitry Yu. Kopytkov, Alexander S. Karaush
Tomsk Municipal Information Library System

При работе автоматизированных библиотечно-информационных систем (АБИС) возникают непредвиденные проблемы, результатом которых является частичная или полная потеря данных. Проведена типизация непредвиденных проблем и рисков при эксплуатации АБИС. Перечислены основные проблемы, возникающие при эксплуатации АБИС. Приведена последовательность действий и перечень мероприятий для уменьшения потерь и восстановления данных, возникающих при внешних непредвиденных воздействиях на Систему «ИРБИС32». Приведены примеры использования утилит, разработанных сторонними поставщиками, для решения возникающих проблем с базами данных и электронными каталогами в Системе «ИРБИС32».

Использование систем автоматизации в повседневной деятельности связано не только с приятными моментами повышения производительности труда, но и с непредвиденными проблемами при работе системы, для которых можно предложить несколько вариантов классификации. Типизация непредвиденных проблем в работе автоматизированной системы может быть следующей:

1. недопустимые простои критически важных сервисов (например, программ определения задолженных документов в филиалах библиотеки для книговыдачи-циркуляции во время рабочего дня, особенно в первых числах сентября);
2. прекращение сопровождения критически важных приложений вследствие распада внутренней команды разработчиков или фирмы-поставщика;
3. масштабное поражение корпоративной сети компьютерным вирусом;
4. необходимость срочной легализации применяемого программного обеспечения (ПО) вследствие правовых проблем и пр.

В рамках данной статьи следует ограничиться раскрытием первого пункта – **ограничением работы сервисов** вследствие нарушения состава данных информационной системы.

К рискам и проблемам внешней среды, которые зачастую невозможно контролировать, относятся:

1. технические проблемы, не предусмотренные в рамках регулярных процедур планирования и управления информационными технологиями (ИТ), например, проблема 2000 года;
2. изменение технической политики производителей оборудования и ПО;
3. изменение лицензионной политики производителей ПО.

Такие проблемы находятся вне зоны контроля отделов информационных технологий библиотек. Особенностью же решения непредвиденных проблем является построение и работа над соответствующими проектами – проектами решения проблем. Подобный проект предполагает срочное изменение инфраструктуры отдела ИТ или всей библиотеки в целом. В таком проекте может возникнуть проект замены (причем срочной) того или иного оборудования или ПО, явившегося причиной проблемы.

По опыту авторов, наиболее часты проблемы, связанные с отказом компьютерного оборудования и системного ПО. Однако нередко случаи, когда вместе с вышеперчисленными проблемами появляются потери данных в АБИС, в том числе Электронного каталога (ЭК) библиотеки. Далее можно предложить следующий перечень

встречающихся проблем с состоянием баз данных (БД), произошедших по причине отказа в работе аппаратного обеспечения или операционных систем:

1. физическое уничтожение данных;
2. потеря целостности структуры данных в среде СУБД;
3. физические нарушения структуры файлов;
4. потеря одного или нескольких значимых файлов СУБД, при которой работа АБИС становится невозможной;

Рассматривая проблему **физического уничтожения данных**, где самым распространенным случаем является отказ жестких дисков или их физическое разрушение, можно сказать, что ничего сопоставимое по эффективности, кроме резервирования данных человечество еще не придумало. По мнению авторов, оптимальным периодом резервного копирования данных информационных систем является период в 24 часа. Это позволяет свести основные затраты от случайной потери файлов баз данных к минимуму, обеспечив при этом потерю данных в ретроспективе одного дня. Сложнее ситуация, когда отказ оборудования происходит «постепенно» и невозможно восстановить данные по копии предыдущего дня. В данном случае для нормальной работы системы резервирования следует применять сложные схемы, где необходимый объем места для хранения копий баз данных должен в 15-50 раз превышать объем самой базы данных. Это необходимо помнить при приобретении оборудования, чтобы обезопасить организацию от полной или глобальной потери данных.

Потеря целостности данных - для локальных АБИС зачастую представляет собой нарушение организации данных в одном или нескольких файлах СУБД, в результате чего становится невозможным поиск или показ результатов поиска, при котором потерь данных ЭК не происходит. В Системе «ИРБИС32» такие ситуации возникают при ошибках уровня таблиц размещения файлов на жестком диске и проявляются в невозможности показа данных словарей или появлении ошибок АРМов Системы «ИРБИС32». Решением данной проблемы является выполнение последовательности команд АРМа «Администратор»: «Диагностика файла документов» и «Создание словарей заново».

При проблемах, связанных с **физическим нарушением в структуре файле**, например потеря данных блока или кластера на жестком диске, возможны случаи потерь диапазона записей базы данных, которые:

1. не поддаются какой-либо диагностике с точки зрения наличия-отсутствия данных и их содержимого;
2. не поддаются анализу по возможной локализации «проблемных» записей;
3. не подлежат восстановлению системными утилитами или утилитами, поставляемыми разработчиками ПО.

При появлении подобных проблем все действия, связанные с доступом к данным следует производить только после резервирования состояния всех файлов системы. Практика показала, что желательно делать резервирование файлов СУБД средствами операционной системы с монопольным доступом к файлам, чтобы уменьшить влияние возможных ошибок на физическом уровне размещения данных в файлах.

Восстановление данных при частичных нарушениях в структуре файла может потребовать проверки на дополнительные условия после восстановления всех записей системы для утверждения, что в процессе восстановления не были потеряны значимые блоки информации. Например, могут возникать случаи при сбое системы, когда потеряны записи, а вместо них в БД присутствует такое же количество хаотично полученных кодов, которые не содержат никакой информации. Простейшим случаем поиска таких проблем является поиск в каждой записи какого-либо ключевого слова (набора слов с условием). Например, создание запроса, который ищет в полях рабочих листов только разрешенные имена.

Потеря файлов является частным случаем потери данных, но далее будет рассмотрен только случай потери файлов, не несущих значимой информации. Для Системы «ИРБИС32» критическим является только один файл с именем, совпадающим с именем конкретной базы данных, имеющий расширение MST (master-file). При потерях других файлов, Система «ИРБИС» полностью сохранит записи и находящуюся в них информацию. Однако для восстановления работоспособности потребуется провести манипуляции, связанные с восстановлением файлов перекрестных ссылок (расширение XRF) и созданием файлов словарей. Наиболее частой причиной потери данных файла перекрестных ссылок (XRF) являются случаи переполнения емкости жесткого диска или попытки сохранения записи, превосходящей физически допустимый разрешенный объем для Системы «ИРБИС32» - 32 Кбайт. Восстановление XRF файла может быть проведено при помощи утилиты MSTRCV (Master Receive), поставляемая UNESCO, которая позволяет восстановить файл XRF по структуре MST файла. После этого для восстановления работоспособности Системы «ИРБИС32» требуется провести заново создание словарей.

В ряде случаев возможны нарушения работоспособности системы «ИРБИС32», которые не поддаются какой-либо систематизации и анализу. Авторами используется следующая последовательность действий в АРМ «Администратор», позволяющая восстановить быстродействие и работоспособность Системы «ИРБИС32» в подавляющем большинстве случаев:

1. Копирование файлов базы данных Системы «ИРБИС32» для задач восстановления в случае неудачного исхода восстановления.
2. Проверка записей в БД («Диагностика файла документов»).
3. Копирование файла документов (для задач дополнительной проверки структуры данных).
4. Экспорт всех записей базы данных в текстовый формат без использования переформатирования. Использование текстового формата передачи позволяет сохранить поля с технологическими метками более 999, которые все чаще используются работниками отделов автоматизации для настройки Системы «ИРБИС».
5. Опустошение БД.
6. Импорт ранее экспортированных записей из текстового формата, при этом следует помнить, что произойдет изменение номеров очередности записей в базе данных, вследствие исчезновения логически и физически удаленных записей, т.е. база данных «сомкнет ряды».
7. Создание заново словарей базы данных.

Данная последовательность действий по восстановлению работоспособности Системы «ИРБИС32» занимает около 1,5-2 часов для базы данных объемом 80000 записей при работе на сервере, но при работе на удаленной машине со скоростью сети 100Mbit данная последовательность занимает уже не менее 4-6 часов. Данная последовательность действий может входить в плановое обслуживание Системы «ИРБИС32» и выполняться не реже 1 раза в 2 месяца.

Следует сказать, что вышеперечисленные операции по восстановлению данных оказались успешными во всех случаях потерь данных, кроме одного, когда потребовалась смена отказавших магнитных головок жесткого диска. Стоимость подобной операции, не дающей 100% гарантии восстановления потерянной информации, составляет от 6000 до 25000 рублей, без стоимости заменяемых деталей. Что существенно выше, чем затраты на ежедневное резервирование данных.

В разработке авторов – ПО «Корректора электронных каталогов» (IsisUtil) имеются функции восстановления данных при потере XRF файла системы «ИРБИС32». Кроме этого IsisUtil позволяет проверять записи на разные условия.

Одним из составляющих проекта «ИРБИС в Томске» (www.irbis.tomsk.ru) является создание и поддержка работы архива утилит и вспомогательного программного обеспечения для обслуживания информационных систем, построенных на базы системы CDS/ISIS (<ftp://library.tomsk.ru/isis/>).

Применение дополнительных утилит и программ восстановления данных, сделанных сторонними поставщиками ПО – весьма сложное и ответственное мероприятие, которое может как повысить устойчивость всей информационной системы библиотеки к внешним деструктивным воздействиям, так и создать дополнительные возможности для расширения возможностей АБИС в целом. При успешном внедрении любой системы автоматизации неизбежно возникает задача изучения сопутствующих программных продуктов, позволяющей также поднять уровень запросов при работе с поставщиками и разработчиками автоматизированных систем.