

УДК 004.056.3

## ПРОГРАММНАЯ СИСТЕМА РЕАЛИЗАЦИИ РЕЗЕРВНОГО КОПИРОВАНИЯ ДАННЫХ

*Федик С.В., Чернышова А.В.*

*Донецкий национальный технический университет*

*Кафедра прикладной математики и информатики*

*svasilich@gmail.com, alla@pmi.dgtu.donetsk.ua*

*Статья посвящена вопросам безопасности корпоративной системы, в частности, резервированию критических данных. В статье рассмотрены существующие методы резервирования данных и предложена концепция построения системы пофайлового резервного копирования.*

Защита корпоративных данных – комплексный многоступенчатый процесс, правильная организация которого является главным условием выживания предприятия в современной среде. Одной из наиболее важных составляющих защиты информации является резервное копирование данных [1].

В условиях современного мира потеря коммерческой информации может привести к серьёзным финансовым убыткам, вплоть до банкротства предприятия.

Для регулярного создания копий важной информации применяются специальные аппаратные и программные средства – системы резервного копирования.

Так как с развитием бизнес-структур объём критических данных, подлежащих долговременному хранению, значительно возрос, повысились и требования к системам резервного копирования (как к программным частям систем, так и к аппаратным).

Резервное копирование данных можно определить как частичное или полное сохранение данных на резервном носителе, в качестве которых могут выступать магнитные ленты, виртуальные библиотеки, дисковые массивы и оптические накопители. Под системами резервного копирования понимаются различные комбинации программного обеспечения для резервного копирования и аппаратного обеспечения для хранения данных [1].

При наличии резервной копии можно достаточно быстро восстановить состояние корпоративной системы на момент последнего резервного копирования [2].

Важными аспектами при разработке системы резервного копирования являются: объём работы с информацией, в частности, объём хранимой информации, структура информационной системы, организация хранения данных, политика компании в области резервного копирования.

Необходимо учитывать совместимость системы резервного копирования с установленным в существующей корпоративной системе оборудованием и программным обеспечением.

Система резервного копирования должна отвечать ряду требований. Рассмотрим наиболее важные.

На первом месте, безусловно, стоит надёжность хранения информации. Обеспечивается применением отказоустойчивого оборудования, дублированием информации, возможностью замены утерянной копии другой.

В связи с огромными объёмами резервируемой информации, важнейшим требованием системы является экономичность используемого в ней алгоритма.

Также важным параметром системы является скорость восстановления зарезервированных данных.

Еще одним важным требованием является целостность хранимых данных. Копирование должно выполняться регулярно. Данные должны быть защищены от несанкционированного изменения. Из этого определяем следующее требование: система резервного копирования должна иметь подсистемы аутентификации, авторизации и шифрования. Данные, резервируемые системой, должны быть доступны только пользователям, имеющим соответствующие права.

Доступность данных: данные, хранящиеся в системе резервного копирования, должны быть

доступны пользователю в любой момент времени.

Простота использования: работа с системой копирования не должна требовать долгого обучения. Интерфейс программной системы должен быть максимально простым и понятным пользователю.

Максимально быстрое и «безболезненное» внедрение в корпоративную систему предприятия.

Экономичность системы в финансовом плане: затраты на поддержку работы системы и её внедрение должны быть минимизированы.

Все системы резервного копирования можно разделить на три класса:

- системы пофайлового копирования;
- системы блочного копирования;
- системы копирования данных на уровне приложений [3].

Системы блочного резервного копирования работают напрямую с носителем данных, игнорируя его файловую структуру. Такие системы позволяют сохранить точную копию всего носителя, включая файлы операционной системы. Некоторые системы позволяют пропускать при копировании неиспользуемые блоки. Основным преимуществом таких систем является высокая скорость копирования данных. Недостаток — сложность при восстановлении отдельных файлов.

Системы пофайлового резервного копирования работают с файловой системой клиента. При таком подходе возрастает нагрузка на операционную систему клиента, однако процесс восстановления отдельных файлов значительно упрощается.

Системы резервного копирования на уровне приложения используют для своей работы API приложения, данные которого, резервируются. При данном подходе могут возникнуть проблемы совместимости системы копирования (формата, в котором хранятся копии) и различных версий приложения [4].

При реализации системы резервного копирования на основе одного из выше перечисленных методов, возникает проблема доступа к копируемым файлам: в момент копирования файл может быть занят другим приложением, в том числе и на изменение, что может привести к нарушению целостности данных.

Эта статья посвящена созданию системы пофайлового резервного копирования.

Существует три базовых метода пофайлового резервного копирования:

- полное резервное копирование;
- инкрементное резервное копирование;
- дифференциальное резервное копирование.

Рассмотрим каждый из них более подробно.

Полное копирование является наиболее простым подходом к резервированию. Этот алгоритм предусматривает полное копирование всех данных при создании каждой новой резервной копии. Реализация такого метода максимально проста, а для восстановления данных достаточно иметь только одну резервную копию. Однако, копирование всех файлов занимает много времени, а хранение избыточных данных требует большой объём дискового пространства.

Другим подходом к резервированию является инкрементный метод. В начале работы программной системы создаётся полная копия данных. После этого система резервирует только те файлы, которые были изменены с момента прошлого резервирования. Такой метод является наиболее экономичным с точки зрения использования дискового пространства. Также увеличивается скорость создания резервной копии. Однако восстановление данных требует много времени. При потере одной из промежуточных копий, есть риск нарушить целостность всей системы. Возможен ещё один вариант работы такого алгоритма. Алгоритм заключается в следующем: каждая создаваемая резервная копия тут же применяется к полной копии (зеркалу), а затем создаётся обратная инкрементная копия. В результате работы такой системы в репозитории всегда находится полная копия текущего состояния системы, и набор инкрементных копий, позволяющий вернуться к предыдущим версиям. [3]

Дифференциальное резервное копирование представляет собой некий компромисс между предыдущими методами. В начале работы системы создаётся полная копия данных. При последующем резервировании копируются те файлы, которые изменились по сравнению с полной копией. Такая

схема позволяет значительно увеличить скорость восстановления данных: для этого нужно применить всего две резервных копии. Также работа алгоритма не требует такого большого объёма дискового пространства для хранения копий, как при реализации полного копирования.

Усовершенствованием инкрементного алгоритма является мультиуровневая схема резервного копирования. Она позволяет значительно увеличить скорость восстановления данных. На нулевом уровне работы алгоритма создаются полные копии резервируемых данных. На последующих уровнях копируются данные, изменённые с момента предыдущего резервирования более низкого уровня. Такой алгоритм обладает достаточно хорошими основными характеристиками, по сравнению с традиционными методами резервного копирования. Для восстановления данных необходимо восстановить последние копии каждого уровня.

На этапе проектирования программной системы резервного копирования было выделено две части: клиентская часть и серверная часть. Серверная часть программной системы должна выполнять следующие задачи: хранение резервных копий, управление процессом копирования файлов, шифрование и расшифрование файлов, контроль целостности данных, аутентификация и авторизация пользователей в системе, хранение служебной информации о состоянии резервных копий, выполненных в соответствии с расписанием, установленным администратором системы или пользователем – владельцем информации. Для хранения служебной информации в системе спроектирована база данных системы.

Клиентская часть программной системы должна выполнять шифрование, расшифрование информации, архивирование данных, установление соединения с сервером и передачу данных на сервер.

При получении сервером уведомления о необходимости создать резервную копию, устанавливается соединение с требуемым клиентом. Клиенту отправляется список файлов, которые содержатся в репозитории и принадлежат данной задаче, и их атрибуты. Получив список файлов, клиентская часть программы выполняет сравнение атрибутов с атрибутами файлов, хранящихся на узле клиента. Затем файлы, атрибуты которых были изменены, архивируются, при необходимости архив шифруется с помощью алгоритма AES и отправляется серверу. Сервер помещает архив в файловое хранилище и записывает изменение в базу данных.

Хранение информации, необходимой для работы системы, реализуется с помощью базы данных. В качестве СУБД предполагается использовать MS SQL Express. Рассмотрим структуру спроектированной базы данных.

База данных будет включать в себя следующие сущности: «задачи», «пользователи», «копии», «расписание», «актуальная копия».

Сущность «задачи» содержит информацию о каждом наборе данных, которые необходимо резервировать. Здесь хранится текстовое представление задачи (отображаемое пользователю), её уникальный идентификатор, информация об алгоритме копирования, дата последнего успешного копирования, информация о владельце задачи и путь к копируемым данным. Также в таблице есть строковое поле хранящее параметры, значение которых зависит от выбранного алгоритма.

Сущность «пользователи» включает в себя информацию о пользователях системы.

Сущность «копии» нужна, в первую очередь, для восстановления данных. Она содержит информацию о каждой копии, созданной системой. Хранит путь к файлу копии, дату создания файла, ссылку на задачу, к которой относится копия, строку с параметрами для корректной работы алгоритмов восстановления.

Сущность «актуальная копия» содержит информацию о файлах, которые включает в себя последняя созданная копия для каждой задачи. Эта информация представляет собой путь к файлу (относительно клиента) и его атрибуты. Также сущность включает информацию о состоянии файла – файл удалён или нет, если удалён, то когда. Набор файлов для каждой задачи зависит от используемого алгоритма копирования.

Для ведения расписания создания копий используется сущность «расписание». Копии могут создаваться как в определённый момент времени, так и по наступлению некоего события (включение

системы и т. п.).

Основными проблемами при создании системы резервного копирования являются необходимость большого объёма дискового пространства для хранения копий, проблемы с доступом к файлам, требование к быстрому восстановлению копии.

В статье описаны классические методы пофайлового резервного копирования. Также в статье описан метод мультиуровневого копирования, который представляет собой сочетание инкрементного и дифференциального методов.

После этапа проектирования программной системы, начата программная реализация системы. В настоящий момент создана база данных, реализован FTP-сервер системы, ведётся работа по реализации подсистемы шифрования данных (посредством библиотеки `сгурторр`) и класса взаимодействия системы с базой данных. На следующих этапах создания системы планируется разработка клиентской части системы, и реализация описанных в статье алгоритмов копирования. Также планируется реализация механизма сжатия данных, аутентификации и авторизации пользователя в системе.

### Литература

- [1] Статья «Резервное копирование: Проблемы решаются просто». Электронный ресурс. Режим доступа: <http://www.cnews.ru/reviews/free/infrastructure2005/articles/reserve.shtml>
- [2] Статья «Технологии резервного копирования». Электронный ресурс. Режим доступа: <http://www.ustar.ua/section/arhive/store/1.html>
- [3] Статья «Технологии и алгоритмы резервного копирования». Электронный ресурс. Режим доступа: <http://www.ict.edu.ru/ft/005653/62330e1-st17.pdf>
- [4] Статья «Резервное копирование и восстановление». Электронный ресурс. Режим доступа: <http://www.trinitygroup.ru/articles/rezcopy>