

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования

**«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ Н.Г. ЧЕРНЫШЕВСКОГО»**

Кафедра теоретических основ
компьютерной безопасности и
криптографии

Система резервирования данных

АВТОРЕФЕРАТ

дипломной работы

студента 6 курса 631 группы

специальности 10.05.01 Компьютерная безопасность

факультета компьютерных наук и информационных технологий

Старичкова Павла Александровича

Научный руководитель

старший преподаватель

А. А. Лобов

22.01.2024 г.

Заведующий кафедрой

д. ф.-м. н., доцент

М. Б. Абросимов

22.01.2024 г.

Саратов 2024

ВВЕДЕНИЕ

В современном мире практически каждый человек взаимодействует с компьютерными системами, которые широко применяются в различных компаниях и предприятиях. Основное предназначение этих систем заключается в обработке и хранении обширных объемов информации.

В процессе своей деятельности компании накапливают огромные массивы данных, часть из которых представляет собой критически важную информацию для пользователей или ключевую для устойчивого функционирования компании. Такая информация включает в себя финансовые данные, коммерческие и производственные тайны, а также личную информацию. Потеря таких данных может стать непредсказуемой проблемой. В этом контексте возникает вопрос: как обеспечить защиту от потери информации?

Один из эффективных методов защиты от потери данных, вызванной различными причинами, такими как сбой оборудования или злонамеренные действия, представляет собой использование систем резервного копирования и восстановления информации. Эти системы нацелены на поддержание целостности и доступности данных. Их ключевая функция заключается в создании дубликатов важной информации в безопасном хранилище. В случае повреждения данных в основной компьютерной системе, эти системы позволяют восстановить информацию из резервной копии.

Для выполнения процедуры резервного копирования часто используются специальные программно-аппаратные системы, известные как системы резервного копирования. Они предназначены для регулярного автоматического копирования системных и пользовательских данных, а также для оперативного восстановления данных.

В рамках данной работы будут рассмотрены различные методы резервного копирования и хранения копий, их преимущества, недостатки и области применения. Также будет проведен обзор существующих систем

резервного копирования данных. В конечном итоге предполагается создание программного продукта, позволяющего эффективно создавать и восстанавливать резервные копии данных.

Дипломная работа состоит из введения, 4 разделов, заключения, списка использованных источников и 1 приложения. Общий объем работы – 97 страниц, из них 42 страниц – основное содержание, включая 42 рисунка, список использованных источников из 21 наименования.

КРАТКОЕ СОДЕРЖАНИЕ

В разделе 1 «Обзор методов резервного копирования» приводится определение резервного копирования и рассматриваются основные методы резервного копирования.

Резервное копирование данных – это процесс сохранения избыточных копий файлов и каталогов, находящихся на локальных дисках, на сменные носители или в защищенное хранилище, которое может быть как в локальной сети, так и вне ее.

В зависимости от важности и частоты использования и изменения данных выбирают один из нескольких основных способов копирования:

1. полное резервное копирование (Full backup);
2. инкрементальное резервное копирование (Incremental backup);
3. дифференциальное резервное копирование (Differential backup).

В подразделе 1 «Полное резервное копирование» подробно рассматривается метод полного резервного копирования, приводятся его усовершенствования, преимущества и недостатки метода.

Метод полного резервного копирования представляет собой создание полной копии всего исходного объема данных. Он выделяется своей простотой и быстротой процесса восстановления, что делает его предпочтительным вариантом для обеспечения безопасности данных. Однако его основными ограничениями являются высокие требования к объему доступной памяти, поскольку каждая копия равна размеру резервируемых данных. Также метод чрезмерно времязатратен при обработке больших объемов данных, и его выполнение создает значительную нагрузку на сеть, что может мешать нормальному функционированию инфраструктуры во время выполнения резервирования.

Полная резервная копия

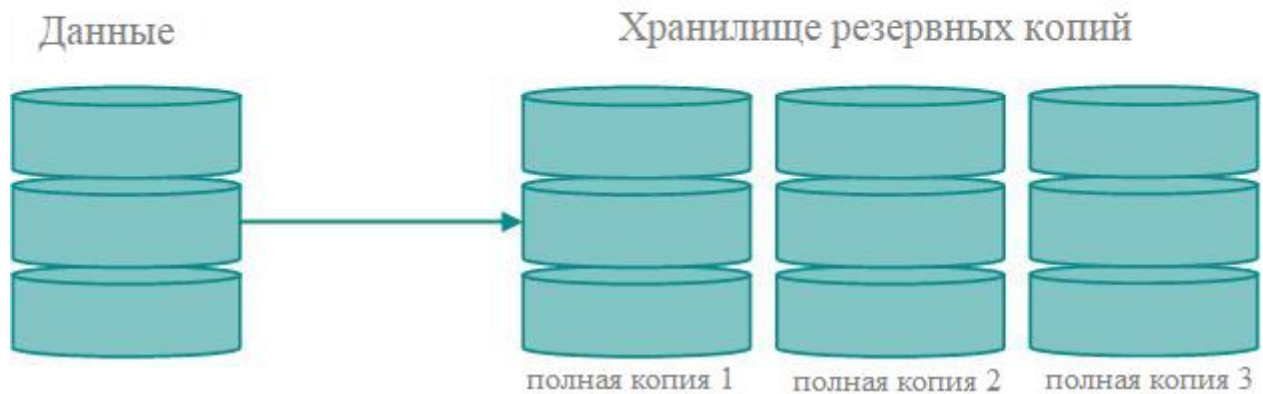


Рисунок 1 – Полное резервное копирование

Существуют усовершенствования этого метода, такие как механизм дедупликации и сжатие данных.

Как правило используется в связке с другими методами.

Быстрое восстановление копии, долгое создание копии, высокие требования к размеру хранилища.

В подразделе 2 «Инкрементальное резервное копирование» подробно рассматривается метод инкрементального резервного копирования, преимущества и недостатки метода.

Инкрементальное резервное копирование исходит из полной копии данных в качестве отправной точки. При последующих операциях копирования сохраняются только измененные данные, произошедшие с момента последнего резервирования. Через определенные большие промежутки времени создается новая полная копия.

Инкрементальный метод резервного копирования

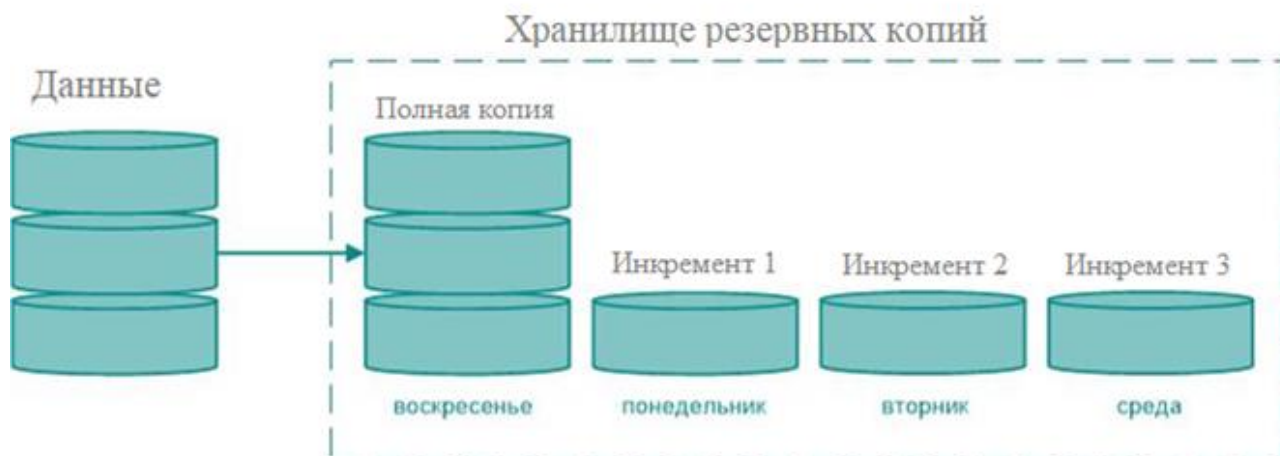


Рисунок 2 – Инкрементальное резервное копирование

При полном восстановлении системы необходимо начать с восстановления из последней полной копии, а затем последовательно восстанавливать данные из инкрементальных копий в порядке их создания.

Быстрое копирование, наименьшие требования к хранилищу, гибкость выбора точки восстановления, медленное полное восстановление.

В подразделе 3 «Дифференциальное резервное копирование» подробно рассматривается метод дифференциального резервного копирования, преимущества и недостатки метода.

Дифференциальное резервное копирование представляет собой средний вариант между полным и инкрементным методами. Стартовой точкой служит полная копия данных, после чего периодически создаются копии только измененных данных, но изменения вычисляются только относительно последней полной копии.

Дифференциальный метод резервного копирования

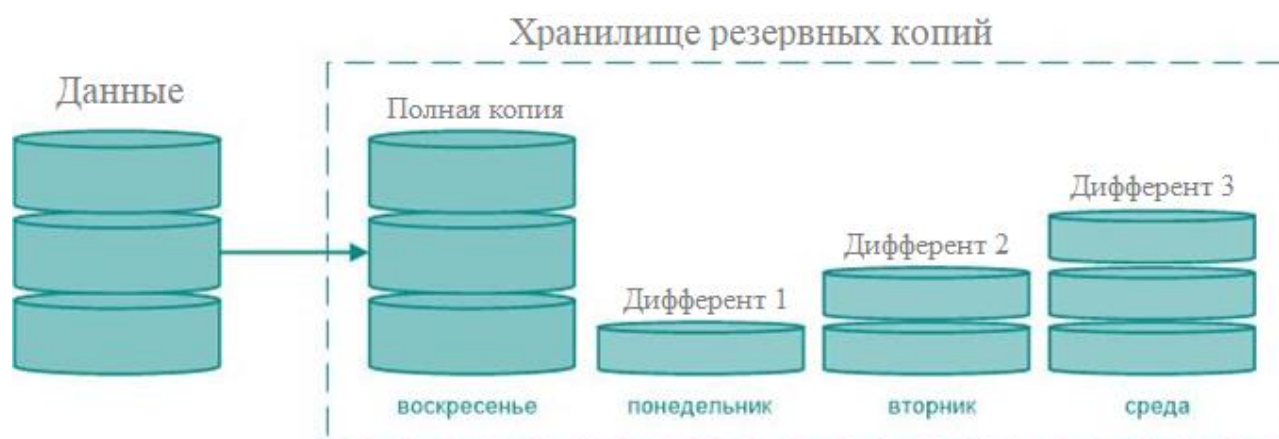


Рисунок 3 – Дифференциальное резервное копирование

Имеет средние показатели по всем параметрам между полным и инкрементальным копированием.

В подразделе 4 «Другие методы» рассматриваются менее распространенные методы резервного копирования и кратко приводятся их принципы работы.

Обратное инкрементальное резервное копирование. Этот метод резервного копирования начинается с создания полной копии данных. В ходе последующих резервных операций все данные из предыдущей полной копии переносятся в новую резервную копию, и предыдущая полная копия заменяется инкрементом. Таким образом, каждая новая резервная копия содержит все изменения, произошедшие с момента предыдущего резервного копирования.

Дельта-копирование. Разновидность инкрементального метода. При таком методе в копию записываются только изменения, происходящие в файлах, а не переписываются полностью изменяемые данные.

Метод бинарных патчей. Этот метод, аналогично дельта-копированию, ориентирован на копирование измененных частей файлов. Однако в отличие от дельта-копирования, в методе бинарных патчей работа ведется на уровне битов информации.

Зеркалирование. Это метод создания точной копии исходных данных, схожий с полным копированием. В отличие от полного копирования, зеркалирование не включает в себя процессы архивирования и систематизации изменений в данных за определенный период. Исходные данные сначала «зеркалируются» в копию. После этого в зеркало копируются только те файлы, которые претерпели изменения, минимизируя объем передаваемых данных. Стоит отметить, что зеркалирование может быть не только физическим, но и географическим. Географическое зеркалирование предполагает хранение зеркальных копий данных в разных географически удаленных местах.

В разделе 2 «Обзор существующих решений для резервного копирования данных» приводятся несколько условных типов продуктов, основные свойства этих типов и их представители, которые можно найти на рынке.

Решения для резервного копирования можно условно разделить на три типа:

- онлайн сервисы резервного копирования;
- программное обеспечение синхронизации файлов;
- программное обеспечение для резервного копирования;

Однако, зачастую продукты можно отнести к двум или всем трем группам.

Онлайн (удаленные, облачные) сервисы резервного копирования – это сервисы, которые предоставляют пользователю систему для резервного копирования, хранения и восстановления данных. Онлайн система резервного копирования обычно собирает, сжимает, шифрует и передает данные на серверы удаленного поставщика услуг резервного копирования.

Программное обеспечение синхронизации файлов. Синхронизация файлов представляет собой процесс обеспечения актуальности компьютерных файлов в двух или более местоположениях согласно определенным правилам. Бывает односторонняя и двусторонняя синхронизация.

Программное обеспечение для резервного копирования. Ориентируется на работу локально или в рабочей сети организации. Некоторые продукты нацелены на работу с томами или физическими дисками. Другие позволяют сохранять отдельные папки и файлы. В большинстве своем поддерживают синхронизацию файлов и создание резервных копий с контролем версий.

Основными возможностями любого продукта являются:

- создание резервных копий с использованием методов резервного копирования или без (синхронизация);
- контроль версий;
- сжатие данных;
- шифрование при передаче и хранении данных;
- поддержка целостности копий и обнаружение ошибок;
- анализ файлов и передача только измененных блоков файлов.

В подразделе 1 «Обзор Яндекс.Диск» приводится обзор на настольную реализацию приложения Яндекс.Диск как средство синхронизации файлов. Выбор был сделан ввиду того, что разрабатываемый в дипломной работе программный продукт взаимодействует с Яндекс Диск.

Приложение удобное и понятное. Имеет внешний вид окна Проводника и позволяет перетаскивать файлы в нужное место назначения. Поддерживает работу с облачным хранилищем Яндекс Диск и двухстороннюю синхронизацию одной локальной папки. Для получения доступа к полному функционалу и синхронизации большего числа папок требуется оплатить подписку.

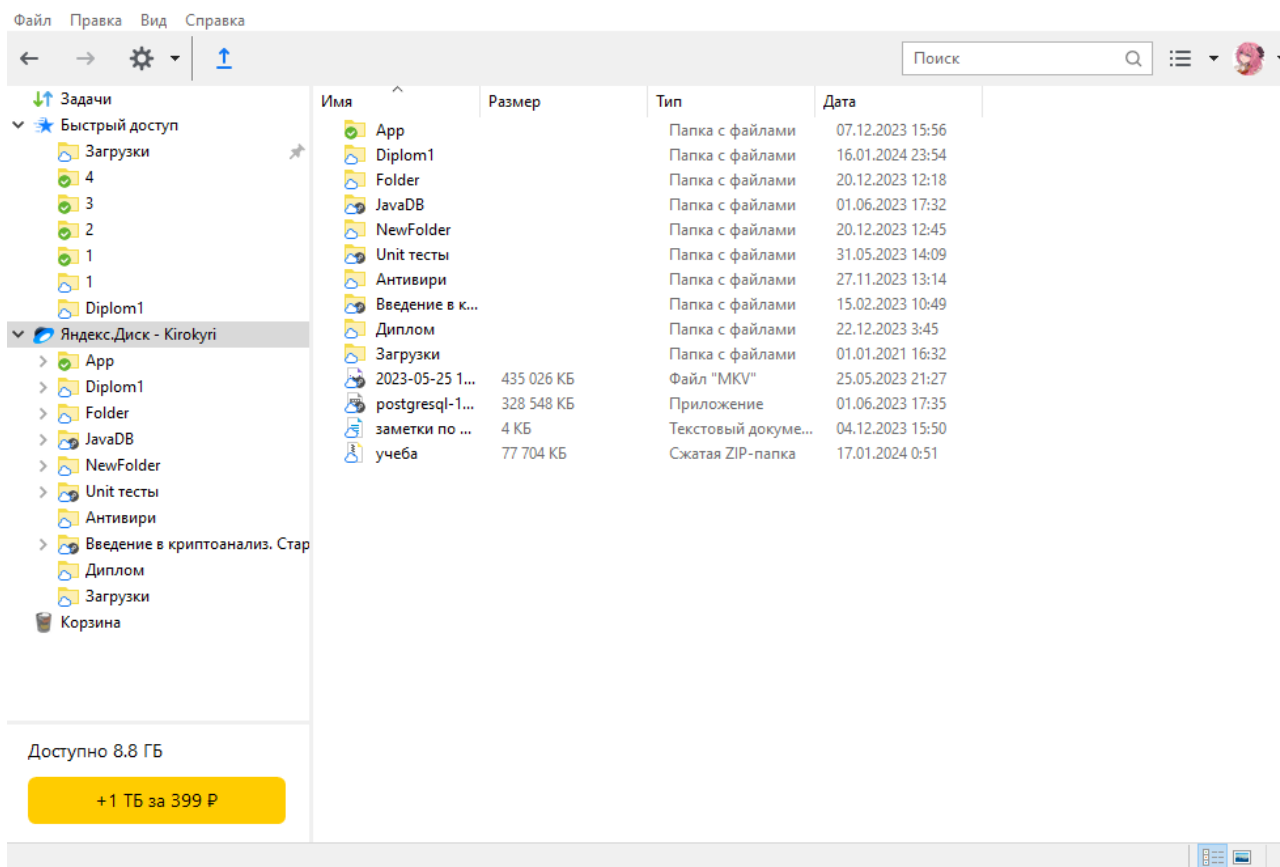


Рисунок 4 – Главное окно программы

В разделе 3 «Программная реализация» описывается программа, разработанная в ходе дипломной работы. Приводятся средства, использовавшиеся для разработки, и библиотеки, используемые в программе. Описываются интерфейс, подсистемы, механизмы и функции программы.

В результате проделанной работы была разработана и реализована программа на языке Python с использованием среды разработки Visual Studio Code.

Программа собрана в исполняемый .exe файл при помощи модуля pyinstaller. Для своей работы требует дополнительные файлы, находящиеся в папке data, в том же каталоге, что и программа. Т.е. пользователю достаточно разархивировать программу и запустить.

Программа предоставляет интерфейс для работы с облачным хранилищем Яндекс.Диск (далее облако) и инструмент создания и восстановления резервных копий данных с автоматическим тихим режимом выполнения копирования.

В качестве основных библиотек используются PyQt5 – для реализации интерфейса и yadisk – для работы с облаком. В качестве вспомогательных библиотек используются sys – для обработки закрытия программы, tempfile – для работы с временными папками, os – для навигации в файловой системе, zipfile – для работы с zip архивами, hashlib – для функции вычисления хэш-суммы sha256 и datetime – для работы с форматом даты и времени.

Программу можно условно поделить на две подсистемы: подсистема работы с облаком и подсистема резервного копирования.

В подразделе 1 «Подсистема работы с облаком» описываются все функции программы, позволяющие работать с облачным хранилищем Яндекс.Диск аналогично рассмотренному ранее одноименному приложению.

В подразделе 2 «Подсистема резервного копирования» описываются все функции и механизмы программы, отвечающие за процесс создания и восстановления резервных копий и соответствующие им элементы интерфейса. Среди них: механизм «рабочих папок», создание резервных копий, выбор и восстановление резервных копий, механизм выбора форматов файлов для автоматической загрузки файлов с выбранными форматами на облако и включения в последующие резервные копии, создание резервных копий в автоматическом режиме по расписанию.

В разделе 4 «Демонстрация работы программы» пошагово проводится демонстрация всех описанных ранее возможностей программы.

В подразделе 1 «Аутентификация» демонстрируется аутентификация в сервисах Яндекс, через которые программа работает с облачным хранилищем, при помощи перехода по ссылке на форму аутентификации Яндекс и последующего ввода полученного кода в поле программы.

В подразделе 2 «Подсистема работы с облаком» демонстрируются возможности программы по работе с облачным хранилищем. Выполняются следующие действия: создание папки в облачном хранилище, загрузка файлов и папок в хранилище, скачивание файлов и папок из хранилища.

В подразделе 3 «Подсистема резервного копирования» демонстрируются возможности программы в области резервного копирования. Описывается принцип определения наличия изменений в файле. Создается новая рабочая папка и в нее загружаются 3 текстовых файла со значением 0. Далее значения в файлах последовательно изменяются и после каждого изменения создаются новые резервные копии. По итогу создаются 5 резервных копий. Затем демонстрируется процесс восстановления резервных копий. После каждого восстановления копии, значения в файлах меняются на те, которые были в них при создании восстанавливаемой копии. И наконец демонстрируется работа автоматического режима программы и механизма односторонней синхронизации папок.

Завершается раздел приведением некоторых особенностей программы:

- не требует установки дополнительных пакетов;
- простой и удобный интерфейс;
- односторонняя синхронизация файлов;
- контроль версий резервных копий.

Также программа совместима с Linux системами. Тестирование выполнялось на дистрибутиве Ubuntu 20.04.6 LTS Focal Fossa.

ЗАКЛЮЧЕНИЕ

Этот исследовательский процесс включал в себя анализ различных методов резервного копирования данных, в их числе исследование преимуществ и недостатков каждого из них.

В ходе работы были достигнуты следующие результаты:

- 1) изучены способы создания резервных копий;
- 2) реализована программа, ориентированная на домашнее использование, которая:
 - 2.1) позволяет работать с облачным хранилищем Яндекс Диск;
 - 2.2) позволяет создавать и восстанавливать резервные копии
 - 2.3) позволяет задать форматы сохраняемых файлов и автоматически загружать их в хранилище;
 - 2.3) имеет возможность создания резервных копий по расписанию;
 - 2.4) имеет графический интерфейс;
 - 2.5) совместима с системами Windows 10 и Linux Ubuntu 22.04.

Следует отметить, что данная работа сконцентрирована на практической реализации и применении изученных методов в реальных сценариях. В ходе исследования мы не только рассмотрели теоретические аспекты, но и успешно применили полученные знания для создания и тестирования программного продукта.

Таким образом, основные поставленные задачи были решены, и цель данного исследования успешно достигнута.