

Министерство образования Республики Беларусь  
Учреждение образования  
Белорусский государственный университет  
информатики и радиоэлектроники

УДК004.056

Гавриленко  
Денис Вадимович

Методы и алгоритмы управления распределением данных между  
облачными хранилищами

**АВТОРЕФЕРАТ**

на соискание степени магистра технических наук  
по специальности 1-40 80 05

«Математическое и программное обеспечение вычислительных машин,  
комплексов и компьютерных сетей»

Научный руководитель  
Куликов С.С.

Минск 2016

## КРАТКОЕ ВВЕДЕНИЕ

Современный этап развития общества подразумевает эффективное управление данными, и в данном контексте особую важность приобретают такие системы, в которых данные хранятся на многочисленных распределённых в сети серверах – так называемых облачных хранилищах данных.

В противовес модели хранения данных на собственных (выделенных) серверах, приобретаемых или арендуемых специально для нужд функционирования системы, внутренняя инфраструктура облачного хранилища может быть неизвестна потребителю услуг: данные хранятся и обрабатываются в так называемом «облаке», которое представляет собой с точки зрения клиента один большой виртуальный сервер. Физически же такие сервера могут располагаться весьма удалённо друг от друга географически, вплоть до расположения на разных континентах.

Практически все сервисы для облачного хранения данных предоставляют пользователям бесплатное пространство в облаке для хранения своих данных. Пользователь может купить дополнительное место в хранилище, тем самым увеличив объём своего облачного хранилища. Так же пользователь может воспользоваться другим облачным сервисом и сохранить часть своих данных.

Одним из важных вопросов является безопасность передаваемых и хранимых данных, поэтому важно, чтобы облачный сервис использовал надёжные протоколы шифрования, что на сегодняшний день в полной мере поддерживается большинством стандартных решений в области облачного хранения данных.

Другим важным показателем качества облачного хранилища является надёжность хранения и доступность данных в облаке: здесь стоит отметить, что основную проблему для отечественных потребителей услуг облачного хранения данных представляют некачественные и медленные каналы передачи данных, т.к. сами по себе облачные хранилища обеспечивают в среднем более высокую надёжность хранения данных, чем локальные решения, что достигается за счёт многократного нагруженного резервирования, распределения данных между несколькими узлами хранения и иных технологических решений. Несмотря на все преимущества использования облачных хранилищ существует вероятность отказа, вызванная временным или постоянным прекращением доступа к облаку.

Кроме того, не стоит полностью доверять хранению важных документов какому-либо одному сервису, так как известны случаи удаления файлов пользователей по инициативе владельцев сервисов, например если появились основания полагать, что файл нарушает авторские права. Так же известны случаи взлома облачных хранилищ с целью получения пользовательскими данными.

Все это наводит на мысли некоторого объединения нескольких облачных сервисов в вычислительную сеть, представленную пользователям в виде однологоического устройства большой емкости.

# ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

## Цель и задачи исследования

*Целью* диссертационной работы является разработка алгоритмов и программного обеспечения для увеличения размера облачного хранилища доступного для хранения пользовательских данных, повышение надежности хранения и степени защиты.

Для достижения поставленной цели необходимо решить следующие задачи:

1. Разработать архитектуру системы объединения нескольких облачных сервисов в одну вычислительную сеть.
2. Разработать методы и алгоритмы защиты пользовательских данных.
3. Разработать методы и алгоритмы для минимизации потери данных при отсутствии доступа к какому-либо облачному сервису.
4. Реализовать возможность добавлять или высвобождать сервисы по мере необходимости.
5. Реализовать ПО в виде локального диска на компьютере, для работы с данными.
6. Провести экспериментальные исследования разработанной системы.

*Объектом* исследования являются технологии облачного хранения данных.

*Предметом* исследования является математическое и программное обеспечение компьютерных систем для решения задач защиты данных находящихся в облачных хранилищах.

Основной *гипотезой*, положенной в основу диссертационной работы, является шифрование и разделение данных при помощи помехоустойчивого кодирования, для повышения уровня защиты данных и отказозащищенности всей системы.

## Связь работы с приоритетными направлениями научных исследований и запросами реального сектора экономики

Работа выполнялась в соответствии с научно-техническим заданием и планом работ кафедры «Программное обеспечение информационных технологий» по теме «Разработать модели, методы, алгоритмы для оценки параметров, повышения надежности и качества функционирования аппаратно-программных средств систем и сетей сложной конфигурации и внедрить в современные обучающие комплексы» (ГБ № 11-2004, № ГР 20111065, научный руководитель НИР – В. В. Бахтизин).

## Личный вклад соискателя

Результаты, приведенные в диссертации, получены соискателем лично. Вклад научного руководителя С.С. Куликова, заключается в формулировке целей и задач исследования.

### **Опубликованность результатов диссертации**

По теме диссертации опубликовано 2 печатные работы, из них 1 работа в материалах конференции аспирантов, магистрантов и студентов БГУИР и 1 публикация в научном журнале.

### **Структура и объем диссертации**

Диссертация состоит из введения, общей характеристики работы, трех глав, заключения, списка использованных источников, списка публикаций автора и приложений.

Во введение была обоснована актуальность диссертационной работы, показана обоснованность использования облачных хранилищ. Представлена практическая значимость эффективного управления данными.

В первом разделе проводится анализ предметной области. Были найдены источники информации, проведена оценка текущего этапа развития облачных технологий, в частности алгоритмов, методов и протоколов облачного хранения данных. В частности дается определения основным составляющим элементам облачного хранения данных. В рамках настоящей главы проводится рассмотрения помехоустойчивого кодирования, а также возможность его применения для повышения надежности хранения данных в облачных хранилищах. Производится обзор существующих облачных сервисов, а также сервисов для взаимодействия с несколькими облачными сервисами. На основе проведенного анализа определены и сформулированы задачи к диссертационной работе.

Во втором разделе производится разработки архитектуры системы, описываются требования, предъявляемые к системе. Выделены задачи и требования предъявляемые к разрабатываемой системе. Разрабатываемая система должна обеспечивать выполнение следующих задач:

- сохранение файла в облаке;
- восстановление исходного файла из облака;

Этап сохранения предполагает разбиение файла на части и генерирование дополнительных частей. После пересылку всех частей на отдельные облачные хранилища.

На этапе восстановления исходного файла должна осуществляться загрузка необходимого количества частей с разных облачных сервисов, объединение в единый файл и сохранение результата в файловой системе.

Третий раздел описывает эксперименты, проведенные над системой распределения данных между облачными хранилищами. Проводится сравне-

ние производительности при разных входных значениях системы. Экспериментально определяются оптимальные конфигурации системы, которая способна распределять файлы между облачными сервисами таким образом, чтобы отказ нескольких из них не приводил к отказу всей системы.

Общий объем работы составляет 57 страниц, из которых основного текста – 40 страниц, 22 рисунка на 20 страницах, 5 таблиц на 14 страницах и список использованных источников из 22 наименований на 2 страницах.

## ОСНОВНОЕ СОДЕРЖАНИЕ

Пояснительная записка состоит из 4 основных разделов.

Во **введение** была обоснована актуальность диссертационной работы, показана обоснованность использования облачных хранилищ. Представлена практическая значимость эффективного управления данными.

В **первой главе** проводится анализ предметной области. В разделе 1.1 представлены предпосылки к появлению облачных технологий. Были сделаны выводы о преимуществах, недостатках и областях применения.

В разделе 1.2 проведен аналитический обзор основных понятий. Проведен анализ существующих методов, их недостатки, преимущества и области применения.

В разделе 1.3 рассматривается протокол установления безопасного соединения для передачи данных в сети интернет.

В разделе 1.4 проведен аналитический обзор открытого протокола авторизации.

В разделах 1.5, 1.6, 1.7, 1.8 выполнен подробный анемический обзор помехоустойчивого кодирования. Рассматриваются возможности и целесообразность применения в системе облачного хранения данных. Определены преимущества и недостатки использования.

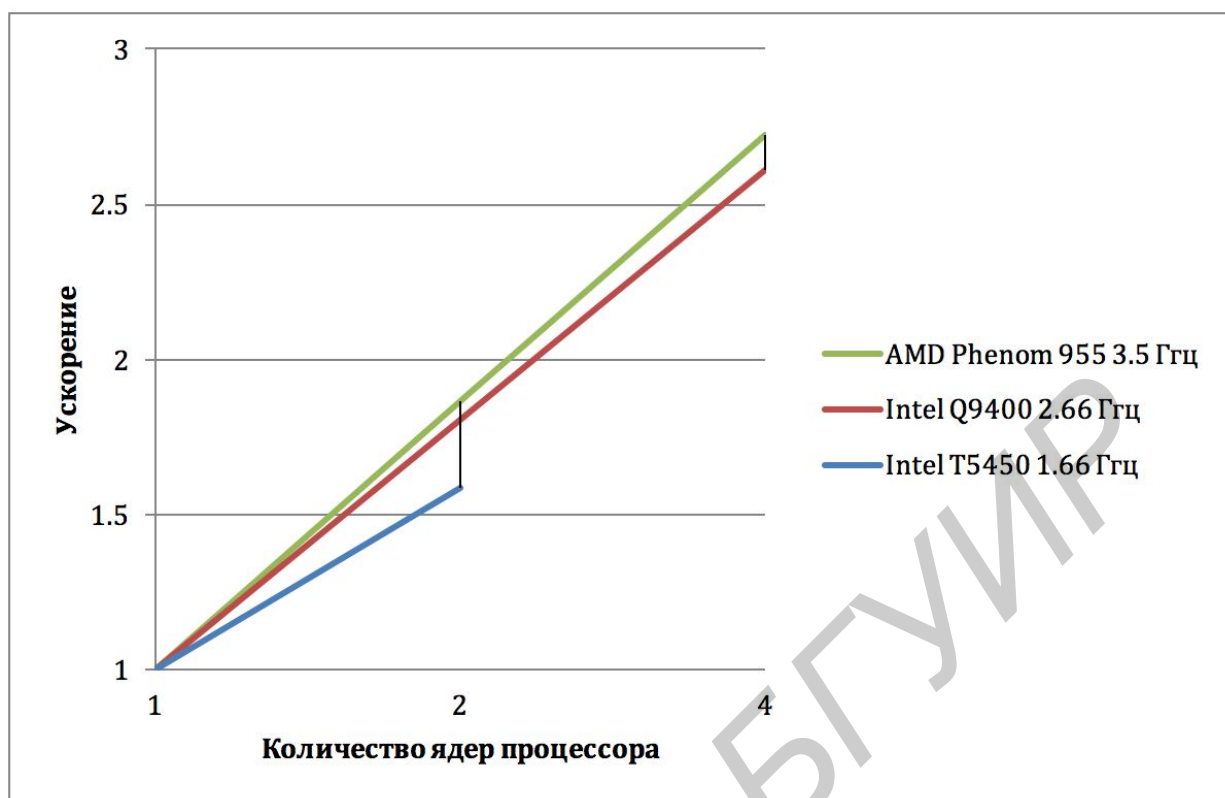
В разделе 1.9 рассматривается возможность масштабирования облачного хранилища. Дается определение миграции данных.

В разделе 1.13 производится обзор существующих облачных сервисов, а также сервисов для взаимодействия с несколькими облачными сервисами.

В разделе 1.14 приводятся результаты обзора существующих аналогов.

В разделе 1.15 приведена укрупненная спецификация требований к разрабатываемой системе.

В **третьей главе** описываются эксперименты, проведенные для определения оптимальной конфигурации системы. Также проведены экспериментальные исследования направленные на определение эффективности использования систем распределения данных. Предложены способы оптимизации кода для устройств, а также архитектурные решения, позволяющие максимально извлечь преимущество от использования параллельных вычислительных систем. Рисунок 1 доказывает, что система работает быстрее при использовании параллельных вычислений.



**Рисунок 1 – График эффективности использования параллельного алгоритма в зависимости от количества используемых ядер процессора.**

## ЗАКЛЮЧЕНИЕ

### Основные научные результаты диссертации

Облачные хранилища предоставляют пользователям много возможностей, недоступных при использовании традиционных хранилищ данных. Поэтому объем данных хранящихся на облачных хранилищах данных растет с каждым днем. Однако в ходе обзора существующих облачных сервисов были выявлены недостатки, связанные защитой и доступностью данных. Поэтому мы создали ПС, которое учло все минусы предыдущих и предотвратила появление последующих критических моментов.

Были изучены основные технологии и методы работы с облачными хранилищами. В ходе рассмотрения существующих облачных сервисов стало ясно, что, для увеличения уровня защиты данных, нельзя доверять хранение целого файла какому-либо облачному сервису. Также традиционное использование облачных сервисов приводило к полной потере данных при отказе облачного сервиса.

Исходя из этого была разработана программа, которая способна распределять файлы между облачными сервисами таким образом, чтобы отказ нескольких из них не приводил к отказу всей системы.

Основные научные результаты диссертации:

1. Предложена архитектура системы для сохранения и восстановления данных, использующая несколько облачных сервисов.

2. Разработаны схемы кодирования и декодирования файлов с использованием МДР-кодов.

3. Экспериментально определены оптимальные значения для кодирования и декодирования файлов.

Достоинствами разработанной системы перед существующими аналогами являются:

1. Высокая отказоустойчивость, так как отказ нескольких используемых облачных сервисов не приводит к отказу созданного хранилища.

2. Высокая защищенность хранимой информации от владельцев облачных сервисов, так как ни один из них не может восстановить исходные файлы.

3. Высокая защищенность от третьих лиц, так как получение доступа к нескольким различным аккаунтам (не более K-1) не позволяет восстановить исходный файл.

4. Возможность бесплатного хранения большего объема информации, так как в итоге пользователю становится доступно пространство на всех облачных сервисах.

5. Возможность хранить файлы большего объема.

Подводя итоги, можно сказать, что в ходе работы было создано приложение, представляющее собой облачный шлюз к большому числу уже существующих облачных сервисов, которая обладает вышеперечисленными преимуществами по сравнению с использованием одиночного облачного сервиса.

### **Рекомендации по практическому использованию результатов**

1. Полученные результаты формируют теоретическую и практическую базу для разработки ПО компьютерных систем для решения задач безопасного хранения данных в облачных хранилищах. Они могут быть использованы для модернизации и дальнейшего развития существующих систем.

2. Разработанные методы и алгоритмы позволяют восстановить данные даже при отсутствии доступа к нескольким облачным хранилищам, тем самым повысить отказоустойчивость системы хранения данных.

3. Результаты работы могут использоваться при подготовке персонала для разработки и обслуживания компьютерных систем, решающих задачи хранения данных в облачных хранилищах.

## **СПИСОК ОПУБЛИКОВАННЫХ РАБОТ**

1. Гавриленко, Д.В. Создание защищенного облачного хранилища/ Д.В. Гавриленко // Научно-методический журнал “Проблемы науки” – 2016. – № 5(6). – с. 31–33.

2. Гавриленко, Д.В. Программное средство управления распределением данных между облачными хранилищами/ Д.В. Гавриленко // 51-я научная конференция аспирантов, магистрантов и студентов. Секция: Программное обеспечение информационных технологий. Тезисы док. – Минск : БГУИР, 2015. – с. 78 - 79.

Библиотека БГУИР