



Сколковский институт науки и технологий

Автономная некоммерческая образовательная организация высшего образования

**«Сколковский институт науки и технологий»**

143025, Московская область, Одинцовский район, Сколково, ул. Новая, д. 100

ОГРН 1115000005922 ИНН/КПП 5032998454/503201001

Тел.: +7 (495) 280-14-81

---

Утверждаю

Ректор, д. т. н., академик РАН

А. П. Кулешов

6 апреля 2018

## **ОТЗЫВ ВЕДУЩЕЙ ОРГАНИЗАЦИИ**

на диссертацию

Иваничкиной Людмилы Владимировны

«Математические модели надежности и методы ее повышения в современных

распределенных отказоустойчивых системах хранения данных»,

представленную на соискание ученой степени кандидата технических наук по специальности

05.13.11 – «Математическое и программное обеспечение вычислительных машин, комплексов и

компьютерных сетей»

Диссертационная работа посвящена исследованию математических моделей надежности хранения данных в современных многодисковых системах хранения данных (СХД). Особое внимание уделяется методам повышения надежности подобных систем с учетом опыта их практической реализации.

Объем накопленных человечеством данных непрерывно растет, что делает проблему их надежного хранения особенно актуальной. Дисковые носители информации имеют ограниченную емкость и подвержены отказам. Поэтому, надежное хранение больших объемов данных можно обеспечить лишь объединением ресурсов множества дисков и организацией автоматического восстановления после дисковых отказов за счет хранения избыточной информации.

В то время как технология хранения данных претерпевали существенные изменения, известные математические модели надежности как правило не менялись и основывались на представлении системы в виде Марковской цепи с дискретным набором состояний. Подобная модель адекватно описывает дисковый массив, но в случае современных СХД с разбиением данных на фрагменты, она способна корректно описать эволюцию лишь одного фрагмента данных, что в значительной степени сужает границы ее применимости.

В диссертационном исследовании представлен новый класс математических моделей, описывающих эволюцию хранилища с разбиением данных на фрагменты, как целостной системы, что делает их более адекватным современным системам хранения данных, чем Марковская модель.

## **Содержание работы**

Работа состоит из введения, четырех глав, заключения и списка литературы.

Во введении обосновывается актуальность темы исследования, научная и практическая ценность работы, её научная новизна, кратко излагается содержание и структура диссертации, представлены положения, выносимые на защиту.

В первой главе определены и проанализированы основные требования к современным системам хранения данных, исследованы преимущества и недостатки существующих решений.

Во второй главе изучены математические модели надежности систем хранения данных. Описана классическая Марковская модель надежности хранения данных, приведено аналитическое решение задачи о среднем времени наработки до отказа. Рассмотрен оригинальный метод решения задачи нахождения среднего времени наработки до отказа в приближении малости интенсивности отказов по сравнению со скоростью восстановления для произвольного графа состояний системы.

Рассмотрены причины, делающие результаты, полученные в Марковской модели, которые лишь качественно применимы для описания реальной СХД. Так, Марковская модель описывает эволюцию единственного фрагмента данных, а в реальной СХД существует множество фрагментов данных, которые эволюционируют отнюдь не независимо друг от друга. В частности, отказ диска приводит к утрате всех хранившихся на нем дисковых блоков, а восстановление утраченных блоков происходит не независимо, а последовательно со скоростью, ограниченной суммарной производительностью дисков.

Учет вышеперечисленных факторов одновременно с отказом от Марковской модели, описывающей эволюцию отдельного фрагмента данных, позволили автору создать оригинальную математическую модель, описывающую эволюцию СХД в целом. Для проверки предсказаний разработанной математической модели автором представлены 2 имитационные модели, имитирующие реальные СХД с различной степенью детальности.

В третьей главе рассмотрены факторы, влияющие на надежность хранения данных в реальных СХД, такие как различные политики размещения дисковых блоков (группы размещения),

наличие скрытых повреждений дисковых блоков, а также учет особенностей аппаратной инфраструктуры (области отказов). Математическая модель надежности дополнена с целью адекватного учета групп размещения и скрытых повреждений. Доказана важная для практической реализации СХД теорема, дающая нижнюю оценку для вероятности потери данных в единицу времени при наличии скрытых повреждений.

В четвертой главе приведено описание внедрения результатов исследования надежности СХД, которые нашли непосредственное применение при создании кластерного хранилища данных, вошедшего в состав программных комплексов Acronis Storage, Virtuozzo Storage и P-платформа.

В заключении приведены основные результаты работы.

### **Новизна исследований и полученных результатов**

В качестве научных результатов следует отметить следующие положения:

- Предложена математическая модель надежности СХД с разбиением данных на блоки, более адекватно описывающих реально существующие системы хранения данных, чем Марковская модель.
- Разработан комплекс имитационных моделей надежности СХД, отражающий реальную архитектуру подобных систем.
- Исследовано влияние различных факторов на надежность и масштабируемость СХД. Впервые изучено теоретически и проверено с помощью имитационных моделей влияние различных политик размещения дисковых блоков на надежность хранилища. Получены количественные оценки влияния скрытых повреждений на надежность СХД.
- На базе полученных теоретических результатов построена распределенная система хранения данных, гарантирующая высокую надежность за счет использования различных схем обеспечения избыточности и оптимизации надежности.

### **Теоретическая и практическая значимость**

Теоретическая значимость диссертационной работы состоит в разработке нового класса математических моделей СХД, описывающих эволюцию хранилища как целостной системы. Результаты теоретических исследований проверяются на имитационных моделях и соотносятся с опытом практической реализации СХД. Полученные в ходе исследования результаты нашли применение при создании кластерного хранилища Acronis Storage, Virtuozzo Storage и продукта «Р-хранилище».

## **Степень обоснованности научных положений, выводов и рекомендаций, сформулированных в диссертации**

Все выводы, полученные в результате проведенной научной работы в диссертации, обоснованы, при этом теоретические результаты доказаны.

## **Достоверность и новизна исследования, полученных результатов, выводов и рекомендаций, сформулированных в диссертации и их достоверность.**

Достоверность научных исследований и результаты работ подтверждаются корректным математическим анализом и проверкой полученных теоретических результатов на имитационных моделях надежности СХД, отражающих реальную архитектуру и особенности реализации подобных систем.

## **Замечания к работе**

Хотя в работе и приведено сравнение классических дисковых массивов с современными СХД с разбиением данных на блоки, хотелось бы иметь более полную информацию об их принципиальных отличиях, которые делают Марковскую модель неприменимой во втором случае.

## **Заключение**

Диссертация Л.В. Иваничкиной на соискание ученой степени кандидата технических наук является научно-квалификационной работой, в которой содержатся исследования и решение задачи повышения надежности систем хранения данных.

Полученные автором результаты достоверны, выводы и заключения обоснованы. Основные научные результаты работы отражены в восьми публикациях. При этом, в списке изданий присутствуют рекомендованные ВАК, индексируемые системами Web Of Science и SCOPUS, два свидетельства о регистрации программы и заявка на патент.

Автореферат полностью отражает содержание работы.

На основании вышеизложенного можно заключить, что диссертация Иваничкиной Л.В. «Математические модели надежности и методы ее повышения в современных распределенных отказоустойчивых системах хранения данных» является законченной научно-квалификационной работой и удовлетворяет требованиям «Положения о порядке присуждения ученых степеней», утвержденного постановлением Правительства Российской Федерации от 24.09.2013 г. №842, предъявляемым на соискание ученой степени кандидата наук а ее автор заслуживает присуждения ученой степени кандидата технических наук по специальности 05.13.11 – «Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей».

Отзыв на диссертацию обсужден и одобрен на заседании ЦНИО “Центр Сколтеха по научным и инженерным вычислительным технологиям для задач с большими массивами данных”, протокол №21-08-18/кд-01 от 21 марта 2018.

доцент ЦНИО “Центр Сколтеха  
по научным и инженерным вычислительным  
технологиям для задач с  
большими массивами данных”, к.ф.-м.н.

Е. В. Бурнаев

директор ЦНИО “Центр Сколтеха по научным  
и инженерным вычислительным технологиям  
для задач с большими массивами данных”  
д.х.н., профессор

М. В. Федоров

Автономная некоммерческая образовательная организация высшего образования “Сколковский институт науки и технологий”

Адрес: 143025, Московская область, Одинцовский район, деревня

Сколково, ул. Новая, д. 100

Телефон: +7 (495) 280 14 81

E-mail: [info@skoltech.ru](mailto:info@skoltech.ru)

Web-site: <http://skoltech.ru>