

## **ОТЗЫВ ОФИЦИАЛЬНОГО ОППОНЕНТА**

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 вычисли-  
тельных машин, комплексов и компьютерных сетей»

### **1. Актуальность темы диссертации**

Надежное хранение данных – проблема, актуальная для всех современных предприятий и пользователей. Актуальность развития теории и практики надежного хранения данных постоянно растет в связи с опережением возрастания сложности систем хранения относительно роста надежности элементной базы систем хранения. Постоянно растущий объем данных требует новых подходов к их хранению. На современном этапе развития информационных технологий значительный интерес вызывает развитие распределенных, в том числе, облачных хранилищ, обеспечивающих высокую надежность хранения данных. Теоретические основы обеспечения и анализа надежности хранения данных в таких системах к настоящему времени разработаны явно недостаточно. В существующих исследованиях, направленных на оценку надежности систем хранения, доминирует использование Марковских моделей в силу простоты их построения и анализа. Простота Марковской модели в ряде случаев приводит к упрощенному представлению процессов функционирования реальных системам хранения данных (СХД) и не позволяет исследовать влияния некоторых существенных факторов организации и функционирования СХД на их надежность и производительность СХД. Автор оппонируемого диссертационного исследования взял на себя смелость выйти за рамки традиционных марковских моделей, и на основе предлагаемых новых решений, в том числе эвристических, разработать модели надежности определенного класса СХД, позволяющие проводить уточненный анализ их надежности.

### **2. Содержание работы**

Работа содержит введение, четыре тематические главы, заключение и список литературы.

Во введении обосновываются актуальность темы исследования, научная и практическая ценности работы, её научная новизна, кратко излагаются содержание и структура диссертации, перечисляются положения, выносимые на защиту.

В первой главе содержится исторический обзор подходов к построению СХД под влиянием требования к теоретически неограниченной масштабируемости хранения с сохранением приемлемого уровня надежности, кратко рассматриваются наиболее известные модели надеж-

ности хранения, используемые для расчета надежности в RAID-массивных дисках и современных облачных хранилищах. Показана ограниченность моделей на основе одномерных цепей Маркова применительно к анализу надежности программных СХД и обоснована необходимость создания модели надежности многодискового хранилища, явно учитывающей разбиение данных на фрагменты.

Вторая глава содержит тщательный анализ наиболее используемых математических моделей систем хранения данных, а также аналитический расчет надежности хранения данных в Марковской модели и оценку среднего времени до отказа. Описан оригинальный метод решения задачи нахождения среднего времени перехода в состояние отказа в приближении малой вероятности одиночного отказа и высокой скорости восстановления для произвольного графа состояний модели.

Утверждение о том, что результаты, полученные для Марковской модели, ограничено применимы к оценке надежности реальной СХД, обосновано тем, что данная модель описывает эволюцию хранилища как единственного фрагмента данных, в то время как в реальной СХД данные разбиты на множество связанных друг с другом фрагментов. Например, выход из строя одного диска приводит к потере сразу всех дисковых блоков, хранившихся на нем, а восстановление утраченных блоков происходит не параллельно, а последовательно с ограничением по скорости, определяемым производительностью оставшихся работоспособных дисков.

Расширение модели для учёта дополнительных условий работы СХД, описанных выше и переход к модели с непрерывным спектром состояний привели автора к созданию оригинальной математической модели, описывающей эволюцию СХД в целом. Проверка корректности построенной математической модели достигается автором путем анализа двух имитационных моделей разной степени детализации.

Полученные в рамках разработанной математической модели результаты использованы для получения аналитических оценок зависимости ее надежности от количества дисков в системе (масштабируемости), позволяя также оценить известные существующие технологические решения по единой методологии.

Третья глава содержит изложение методов оптимизации надежности хранения данных с учетом свойств реальных СХД, таких как настраиваемые политики размещения дисковых блоков (группы размещения) и учет скрытых дисковых повреждений, которые входят в разработанную модель СХД в виде параметров, а также особенностей аппаратной инфраструктуры (областей отказов). В данной главе доказана теорема о нижней границе вероятности потери данных в единицу времени при наличии скрытых повреждений. Основным выводом, имеющим важное практическое значение при построении будущих СХД, является преимущественное влияние скрытых повреждений данных на потерю данных в случае использования типичных

дисковых накопителей, используемых в реальных СХД. В таком случае вероятность потери данных оказывается гораздо больше алгебраической суммы вероятностей потери данных вследствие одних только дисковых отказов и скрытых повреждений. Дополнительно доказан тезис о том, что приближение независимости дисковых отказов и скрытых повреждений приводит к неверным оценкам надежности, накладывая требование на их совместное рассмотрение.

Четвертая глава посвящена описанию внедрения результатов исследования надежности СХД в процессе создания программного комплекса «Acronis Storage», конкретно в части кластерного отказоустойчивого хранилища данных.

### **3. Новизна исследований и полученных результатов**

Анализ результатов исследования, представленных в диссертации, и опубликованных в научных журналах и сборниках научных докладов, позволяет сделать вывод о том, что результаты диссертационного исследования, обладают научной новизной.

Научная новизна диссертации заключается в том, что в ней предложены модели надежности многодискового хранилища данных с разбиением данных на фрагменты, отражающие особенности систем СХД рассматриваемого класса, что обуславливает эффективность использования этих моделей как при обосновании проектных решений, реализуемых при построении СХД, так и при прогнозировании их надежности.

К результатам, обладающим новизной, следует отнести следующие положения:

- Предложена математическая модель надежности СХД с разбиением данных на блоки, в которой преодолены ограничения, свойственные известным Марковским моделям систем хранения данных. Адекватность предложенной аналитической модели подтверждена результатами имитационного моделирования.

- Исследовано влияние различных факторов на надежность и масштабируемость СХД, рассматриваемого класса. Впервые изучено теоретически и проверено с помощью имитационных моделей влияние различных вариантов размещения дисковых блоков на надежность хранилища, рассматриваемой организации. Получены количественные оценки влияния скрытых повреждений на надежность СХД, рассматриваемого класса. Предложены и обоснованы методы борьбы со скрытыми повреждениями и даны оценки их эффективности.

- На основе, предложенных моделей проведено обоснование проектных решений построения распределенной высоконадежной системы хранения данных рассматриваемой организации, при этом проведено обоснование использования различных схем обеспечения надежности СХД рассматриваемого класса.

### **4. Теоретическая и практическая значимость.**

Несомненная теоретическая значимость результатов диссертационных исследований состоит в разработке новых математических моделей надежности, в которой преодолены ограни-

чения свойственные Марковским моделям, что позволило отразить ключевые особенности реальных СХД, рассматриваемого класса.

Предложенный подход к построению аналитической модели надежности многодисковых СХД рассматриваемого класса, позволил осуществить переход от рассмотрения дискретного набора состояний отдельного фрагмента данных к описанию эволюции состояния всей системы.

Результаты теоретических исследований применены при создании кластерного хранилища «Acronis Storage». При этом следует подчеркнуть, что система хранения «Acronis Storage» является одним из ключевых участников международного рынка СХД.

Теоретическая и практическая значимость, полученных результатов подтверждается их использованием в рамках проведения прикладных научных исследований, выполняемых по Соглашению с Министерством науки и образования о предоставлении субсидии №14.579.21.0010 от "05"июня 2014 года по теме "Технология и программное обеспечение распределенных и высокопроизводительных вычислительных систем хранения и обработки больших данных". Уникальный идентификатор Соглашения RFMEFI57914X0010.

Практическая значимость и реализуемость исследований подтверждается двумя свидетельствами о государственной регистрации программы для ЭВМ № 2015618800. 2015 и № 2016613180. 2016

**5. Степень обоснованности научных положений, выводов и рекомендаций, сформулированных в диссертации их достоверность**

Выводы, полученные в диссертационном исследовании, достаточно обоснованы, при этом теоретические результаты доказаны.

**6. Степень обоснованности научных положений, выводов и рекомендаций, сформулированных в диссертации, их достоверность**

Проведенные автором исследования основываются на использовании математического аппарата теорий надежности, теории вероятностей, теории случайных процессов, являющегося общепризнанной в настоящее время методологической базой анализа процессов обработки, хранения и передачи данных.

Достоверность научных исследований и полученных результатов подтверждаются корректным использованием математического аппарата и проверкой полученных теоретических результатов на имитационных моделях СХД. Достоверность результатов подтверждается их непротиворечивостью с результатами других исследований других авторов и результатами эксплуатации СХД отражающих реальную архитектуру подобных систем, в том числе моделирующих поведение коммерческой СХД «Acronis Storage», реализованной при участии автора в составе коллектива разработчиков.

Достоверность полученных результатов, судя по представленным материалам, диссертации подтверждается их использованием в рамках проведения прикладных научных исследований ООО «Проект ИКС»,

Достоверность основных положений, выводов и рекомендаций, сформулированных в диссертации, подтверждается их представлением и обсуждением на международных и всероссийских конференциях, а также публикацией в изданиях из перечня ВАК и изданиях, индексируемых системами WebOfScience и SCOPUS.

Анализ публикаций автора позволяет сделать вывод, что они отражают основные научные результаты диссертационной работы.

Автореферат соответствует установленным стандартам, достаточно полно отражает содержание диссертации.

## **7. Замечания и недостатки**

Диссертация и автореферат, как и всякая новое исследование, несвободны от недостатков, к числу которых могут быть отнесены следующие:

- 1) В разделе «Современные системы хранения данных и проблема моделирования их надежности» упомянуты только подходы к построению СХД на основе RAID – массивов и СХД СЕРН, но не освещены принципы построения распределенных СХД корпораций EMC, DELL-EMC, NetApp и др., при этом в диссертации отсутствуют сведения о моделировании надежности упомянутых СХД. О недостаточности анализа состояния вопроса принципов построения, моделирования и обеспечения надежности СХД косвенно можно судить по незначительности списка литературы в котором имеется всего 47 наименований при отсутствии классических работ по теории надежности.
- 2) В качестве показателей надежности СХД в диссертации вполне обоснованно оценивается среднее время до отказа (т.е. состояния, когда данные не могут быть восстановлены), однако отсутствие оценки таких показателей надежности как стационарный и нестационарный коэффициенты готовности, коэффициент оперативной готовности, а также коэффициент сохранения эффективности ограничивает применимость предлагаемых моделей.
- 3) В названии и цели диссертации в качестве объекта исследования декларируются распределенные отказоустойчивые СХД, однако предложенные модели не позволяют отразить влияние на надежность СХД сетевой составляющей (наличие сетевого оборудования, каналов связи, используемых протоколов), хотя механизмы обеспечения надежности и отказоустойчивости описаны в разделе 4.
- 4) В диссертации не рассмотрены вопросы оценки отказоустойчивости СХД, например, минимального числа отказов, гарантированно выдерживаемых системой или условную

вероятность сохранения функционирования при возникновении определенного числа отказов и др. показателей отказоустойчивости, отличающихся от показателей надежности.

- 5) В работе недостаточно подробно описаны комплекс имитационных моделей надежности СХД, а также сценарии проверки адекватности предлагаемых аналитических моделей, что затрудняет восприятие результатов работы.
- 6) В названии диссертации слово «современных» представляется излишним, так как зачем исследовать не современные системы.
- 7) В диссертации и автореферате имеются опечатки и неточности оформления: объем автореферата (28 с) превышает рекомендованные значения, после формул в автореферате и диссертации не проставлены знаки препинания, в автореферате в конце подрисуночных подписей проставлены точки, в списке литературы диссертации и автореферата не везде указаны страницы источников, например, п.п. 1, 2.

Отмеченные замечания носят частный характер и не влияют на общую положительную оценку диссертационной работы.

Соискатель продемонстрировал хорошую теоретическую, в том числе математическую, подготовку, знание современных информационных и компьютерных технологий, знакомство с основными достижениями в выбранной области исследования, умение ставить и решать сложные задачи в области анализа надежности и проектирования распределенных СХД. Автор показал знания и умения применять математический аппарат теории надежности и теории случайных процессов к построению к анализу и поиску принципов построения СХД, что позволяет сделать вывод о высокой квалификации соискателя.

Материал диссертации изложен ясно, диссертация четко структурирована и оформлена в соответствии с требованиями нормативной документации по оформлению научных работ.

#### 8. Заключение.

Диссертация Иваничкиной Людмилы Владимировны на соискание ученой степени кандидата технических наук является законченной научно-квалификационной работой, в которой содержится новое решение задачи повышения надежности систем хранения данных, имеющей существенное значение для науки и практики.

Полученные автором результаты достоверны, выводы и заключения обоснованы. Основные научные результаты работы отражены в 8 публикациях. При этом, в списке изданий присутствуют работы, рекомендованные ВАК и индексируемые системами Web Of Science и SCOPUS, имеются два свидетельства о регистрации программы. Автореферат полностью и достоверно 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, Иваничкина Людмила Владимировна, заслуживает присуждения ученой степени кандидата технических наук по специальности 05.13.11 – «Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей».

Официальный оппонент,

Профессор кафедры Вычислительной техники Университета ИТМО,

доктор технических наук

В.А.Богатырев

«18» апреля 2018 г.

Федеральное государственное автономное образовательное учреждение высшего образования «Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики».

Университет ИТМО, Кронверкский пр., д.49, Санкт-Петербург, Российская Федерация, 197101.  
Телефоны: общий отдел +7 (812) 232-97-04; ректорат: +7 (812) 233-00-89. Факс: +7 (812) 232-23-07

E-mail: [od@mail.ifmo.ru](mailto:od@mail.ifmo.ru), [org@mail.ifmo.ru](mailto:org@mail.ifmo.ru)

<http://www.ifmo.ru>