

ДУНАЕВ ВАЛЕРИЙ АЛЕКСАНДРОВИЧ

**МОДЕЛЬ И АЛГОРИТМЫ УПРАВЛЕНИЯ ПАРАМЕТРАМИ
РЕПЛИКАЦИИ В РАСПРЕДЕЛЕННОЙ БАЗЕ ДАННЫХ
ПРЕДПРИЯТИЯ ГОРНОПРОМЫШЛЕННОГО КОМПЛЕКСА**

Специальность 05.13.06 – "Автоматизация и управление технологическими
процессами и производствами (промышленность)"

ДИССЕРТАЦИЯ

на соискание ученой степени кандидата технических наук

Научный руководитель: кандидат технических наук, доцент Тараканов О.В.

Орел 2014

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ	7
ГЛАВА 1. ОБСЛЕДОВАНИЕ МОДЕЛЕЙ ОБРАБОТКИ ИНФОРМАЦИИ В РАСПРЕДЕЛЕННЫХ БАЗАХ ДАННЫХ ПРИ РЕПЛИКАЦИИ.....	18
1.1 Описание технологий репликации в распределенных баз данных.....	18
1.2 Описание информационного обеспечения управления предприятием ГПК	22
1.3 Описание подходов к моделированию процессов, протекающих в распределённых базах данных.....	30
1.4 Описание процесса репликации в РБД предприятия ГПК "ШахтИнвестКузбасс"	36
1.5 Постановка задачи исследования	37
ГЛАВА 2. РАЗРАБОТКА МОДЕЛИ ОТКЛИКА РБД НА ЗАПРОСЫ ПРИ РЕПЛИКАЦИИ.....	43
2.1 Выбор математического аппарата для разработки модели	43
2.2 Модель отклика РБД на запросы при репликации	47
2.2.1 Обоснование выбора схемы владения данными.....	47
2.2.2 Проверка гипотезы о согласовании законов распределения потоков заявок с распределением Пуассона.....	50
2.2.3 Общий вид модели.....	54
2.2.4 Модель обработки запросов на резервном сервере.....	58
2.2.5 Модель обработки запросов на главном сервере	58
2.2.6 Модель обработки запросов на участке сети от главного сервера до резервного	59
2.2.7 Модель обработки запросов на участке сети от резервного сервера до главного.....	60
2.3 Проверка адекватности модели отклика РБД на запросы при репликации.....	60
2.4 Проверка чувствительности модели отклика РБД на запросы при репликации.....	62

ГЛАВА 3. РАЗРАБОТКА АЛГОРИТМОВ ВЫЧИСЛЕНИЯ ЗНАЧЕНИЙ ПАРАМЕТРОВ РБД ПРИ РЕПЛИКАЦИИ	67
3.1 Задача минимизации среднего времени отклика РБД на запросы.....	67
3.2 Обоснование математического метода решения задачи минимизации среднего времени отклика РБД на запросы.....	69
3.3 Алгоритм вычисления оптимальной загруженности резервного узла распределенной базы данных при репликации	74
3.4 Свойства алгоритма вычисления оптимальной загруженности резервного узла распределенной базы данных при репликации	82
3.4.1 Оценка корректности алгоритма	82
3.4.2 Оценка сложности алгоритма.....	84
3.4.3 Оценка точности алгоритма.....	85
3.4.4 Оценка вычислительной устойчивости алгоритма	86
3.5 Алгоритм выбора фрагментов данных для немедленной репликации... ..	88
3.6 Свойства алгоритма выбора фрагментов данных для немедленной репликации.....	95
3.6.1 Оценка корректности алгоритма	95
3.6.2 Оценка вычислительной сложности алгоритма.....	96
3.6.3 Оценка точности алгоритма.....	96
3.6.4 Оценка вычислительной устойчивости	97
ГЛАВА 4. РАЗРАБОТКА АЛГОРИТМА АВТОМАТИЗАЦИИ ПРОЦЕССА КОНФИГУРИРОВАНИЯ РЕПЛИКАЦИИ В РБД ПРЕДПРИЯТИЯ ГПК....	100
4.1. Алгоритм автоматизации процесса конфигурирования репликации в РБД предприятия ГПК.....	100
4.2. Этап вычисления значений характеристик РБД	101
4.2.1 Оценка применимости этапа вычисления значений характеристик РБД предприятия ГПК "ШахтИнвестКузбасс"	105
4.3. Этап вычисления оптимальной загруженности резервного узла и выбора фрагментов данных для немедленной репликации.....	113
4.3.1 Пример выбора фрагментов данных для немедленной репликации	116

4.4 Этап формирования рекомендаций для администратора РБД предприятия ГПК по репликации.....	119
ЗАКЛЮЧЕНИЕ	123
Список использованных источников	127
Приложение А. Характеристики фрагментов РБД предприятия ГПК "ШахтИнвестКузбасс"	138

Определения, обозначения и сокращения

В настоящей работе применяют следующие термины с соответствующими определениями:

Область адекватности – некоторая окрестность значений показателей существенных свойств [54].

Репликация – процесс копирования информации из одного хранилища данных в другое с последующей синхронизацией их содержимого [16].

Фрагментация – процесс, посредством которого файлы (или эквивалентные объекты базы данных, такие как отношения) разбиваются на части и распределяются по нескольким локальным узлам [76].

База данных – совместно используемый набор логически связанных данных, предназначенный для удовлетворения информационных потребностей пользователей [48].

Параметр – количественное отражение некоторого свойства элемента системы [54].

БД – база данных.

РБД – распределенная база данных.

ГПК – горнопромышленный комплекс.

СМО – система массового обслуживания.

СеМО – сеть массового обслуживания.

МО – математическое ожидание.

FCFS - First-Come, First-Served.

СКО – среднее квадратическое отклонение.

Gflops – Billions of Floating-point Operations Per Second.

СУБД – система управления базами данных.

SQL – Structured Query Language.

АС – автоматизированная система.

АСУ – автоматизированная система управления.

ТКС – телекоммуникационная подсистема.

СУ ГПК – система управления горнопромышленным комплексом.

КИВС – корпоративная информационно-вычислительная сеть

ВВЕДЕНИЕ

Актуальность темы. Горнопромышленные комплексы (ГПК) по добыче угля являются территориально распределенными большими организационно-техническими системами. Производственный процесс угледобычи относится к классу опасных производств, что обуславливает высокие требования к системе управления ГПК [75, 38, 96].

Множество служб ГПК, отличающихся по уровню и выполняемым задачам, порождает высокую интенсивность потоков информации, передаваемых между различными объектами шахты. Эффективное управление сложным объектом обеспечивается за счет применения комплекса средств автоматизации, на основе которого построена АСУ предприятием ГПК [89]. АСУ предприятием ГПК "ШахтИнвестКузбасс" сформирована по классическим принципам и территориально распределена по службам шахты.

Повышенные требования к безопасности в угольных шахтах делают необходимым применение оперативных и достоверных систем сбора и обработки информации [71]. Необходима организация информационного обеспечения, при которой пользователи на различных участках ГПК могли бы оперативно получать актуальную информацию для принятия управленческих решений. В системе управления ГПК "ШахтИнвестКузбасс" информационное обеспечение построено на основе распределенной базы данных (РБД), интегрирующей на уровне информации функциональные подсистемы АСУ.

В общем случае СУ ГПК содержит следующие службы:

- аэрогазового контроля;
- вентиляционного проветривания;
- теплоэнергетики шахты;
- пожарного водоснабжения;
- водоотлива;
- электроснабжения;
- конвейерного транспорта;

-технологического оборудования очистных и подготовительных участков шахты;

- геосейсмического мониторинга;
- мониторинга параметров шахтного пространства;
- виброакустического мониторинга горного массива;
- радиологического мониторинга горного массива;
- мониторинга и контроля гидрогеологической обстановки;
- радиозондирования горного массива;
- системы наблюдения и оповещения персонала.

Так, например, система наблюдения и оповещения персонала обслуживает большое количество запросов по сбору и обработке данных о местоположении каждого персонального датчика, собирает данные о перемещениях и по запросу передает информацию горному диспетчеру. Современное оборудование позволяет устанавливать местонахождение рабочего персонала со средней точностью 0,34 метра, а при необходимости производить аварийное отключение угледобывающих установок. Территориальное распределение компонентов АСУ предприятием ГПК и высокие вычислительные нагрузки в купе с требованиями по отказоустойчивости компонентов АСУ предприятием ГПК создают предпосылки к эффективной организации распределенной обработки данных [61].

Построение распределенных систем обработки информации реализуется в рамках трех архитектур [85]:

1. *Архитектура с разделением файлов*, включающая несколько клиентов, связанных сетью с файловым сервером. При этом файловый сервер хранит все разделяемые клиентами файлы. Данная архитектура самая простая с точки зрения построения. Однако при большом количестве запросов от клиентов и их территориальном удалении использование архитектуры с разделением файлов приводит к высоким нагрузкам на телекоммуникационную подсистему и файловый сервер. С другой стороны, при выходе из строя файлового сервера или перебое связи с ним происходит

отказ системы в целом. В условиях повышенных требований к безопасности в угольных шахтах, применение архитектуры с разделением файлов для эффективной организации информационного обеспечения АСУ предприятием ГПК недопустимо.

2. *Сервер баз данных*, передающий лишь ответы на запросы, тем самым уменьшает нагрузку на телекоммуникационную подсистему. Основной недостаток такой архитектуры заключается в ее низкой масштабируемости. Так, при увеличении количества клиентов и их территориальном удалении задержки, возникающие при обслуживании удаленных запросов, выходят за рамки допустимых пределов. Преодоление данного недостатка возможно построением АСУ на основе архитектуры РБД.

3. *Распределенная база данных* позволяет равномернее распределять нагрузку между удаленными серверами, собирающими и обрабатывающими данные от различных служб ГПК, повышать отказоустойчивость системы. При выходе из строя одного из серверов остальные продолжают работу. После восстановления вышедший из строя узел синхронизируется с остальными и продолжает работу в штатном режиме.

Применение в рамках специализированного информационного обеспечения управления предприятием ГПК технологий РБД предполагает решение дополнительных задач по поддержанию копий данных в непротиворечивом состоянии, разрешению конфликтов при блокировках наборов записей, восстановлению при сбоях. Это достигается за счет репликации данных. Расчет значений её параметров, при которых достигается высокая эффективность функционирования системы, является сложной задачей, которая на настоящий момент, применительно к условиям функционирования ГПК, не имеет однозначного решения.

Так как АСУ предприятием ГПК функционирует в условиях большого количества случайных факторов: повышенная влажность, запыленность, сейсмоактивность, низкие и высокие температуры, вибрации, то используемое оборудование оснащается дополнительными элементами, снижающими влияние данных факторов. Например, в помещениях шахты с

взрывоопасной средой используются понижающие разделительные трансформаторы, оборудование взрывозащищенного и искробезопасного исполнения [77]. С другой стороны, наличие специализированного оборудования повышает стоимость и, как следствие, требует эффективного распределения ресурсов при настройке специализированного информационного обеспечения. При этом нерациональный выбор параметров репликации данных влечет за собой превышение допустимых временных задержек на различных этапах обработки запросов и, как следствие, снижает оперативность отклика РБД [6].

Так как основное назначение РБД – удовлетворение информационных потребностей пользователей (приложений), то оценка эффективности функционирования РБД должна производиться с точки зрения эффективности выполнения запросов и транзакций, отражающих информационные потребности конечных пользователей, приложений и действий над РБД [55].

Согласно теории эффективности целенаправленных процессов эффективность процесса выполнения запросов и транзакций оценивается по трем показателям: оперативности, результативности и ресурсоемкости. При этом под оперативностью выполнения запросов и транзакций, в общем случае, понимают расход времени, потребного для отклика РБД на запросы. Результативность выполнения запросов и транзакций определяется получаемым целевым эффектом – результатом, ради которого функционирует система. Ресурсоемкость характеризуется ресурсами всех видов, используемыми для получения целевого эффекта [68, 54, 10].

На практике ограничиваются оценкой отклика РБД на запросы при репликации по частным показателям оперативности, а показатели результативности и ресурсоемкости вводят в ограничения. Это связано с тем, что, как правило, вычислительные и сетевые ресурсы заданы априорно, а результативность характеризуется единственно возможным значением показателя – завершенностью операции отклика.

При синхронной репликации транзакции не фиксируются до тех пор, пока на всех узлах, содержащих копии данных, не будут внесены изменения. Это позволяет добиться максимальной актуальности данных, но вносит дополнительные задержки, вызванные блокировкой обновляемых данных. Реализация синхронной репликации возможна только при наличии надежных высокоскоростных каналов связи.

При асинхронной репликации изменения в копии данных вносятся независимо, что позволяет избавиться от задержек блокировок. Однако при таком подходе снижается текущая актуальность данных в РБД. Режим асинхронной репликации предъявляет более гибкие требования к ресурсам сети.

Помимо выбора типа репликации для достижения высокой оперативности отклика на запросы необходим правильный выбор фрагментов немедленно реплицируемых данных на узлах РБД. Принятие решения о выборе того или иного фрагмента для немедленной репликации должно производиться с точки зрения оценки интенсивности поисковых запросов и запросов на обновление к этому фрагменту, а также времени их обработки в соответствии с доступными вычислительными и сетевыми ресурсами.

В результате возникает противоречие, состоящее в том, что уменьшение количества реплицируемых данных порождает увеличение числа удаленных заявок на выборку данных и снижение числа заявок на выборку, обрабатываемых на резервных серверах.

Как следствие, среднее время отклика РБД на запросы увеличивается за счет удаленного обслуживания запросов, рисунок 1. Где $I'u$ – интенсивность потока заявок на обновление; $\overline{I'q}$ – интенсивность потока заявок на выборку данных, обрабатываемых удаленно; $I'q$ – интенсивность потока заявок на выборку, обрабатываемых локально; n – количество резервных серверов.

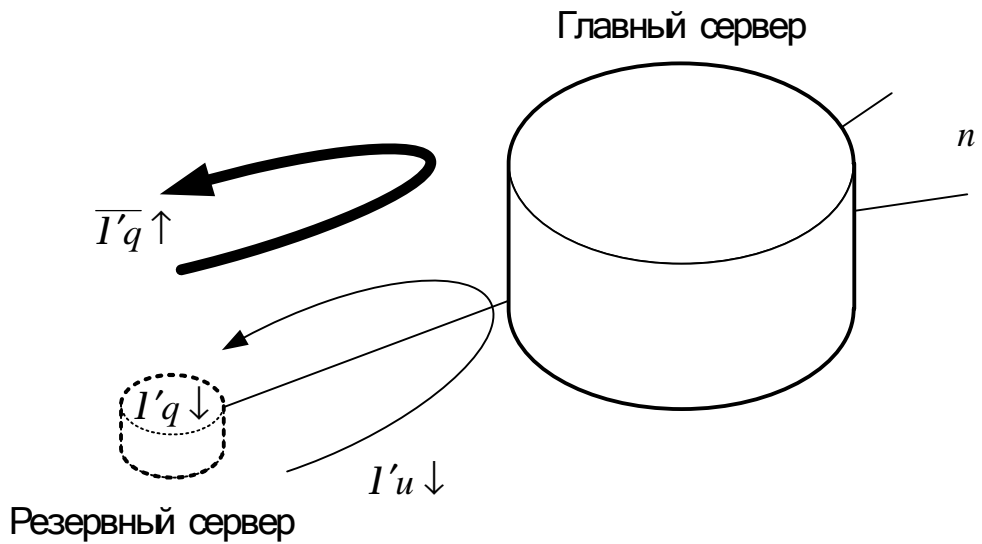


Рисунок 1 – Случай, когда доля реплицированных данных мала

С другой стороны, увеличение количества реплицируемых данных приводит к росту числа заявок на репликацию, что создает дефицит вычислительных и сетевых ресурсов. Так, среднее время отклика РБД на запросы увеличивается за счет роста временных задержек на различных этапах обработки запросов, рисунок 2.

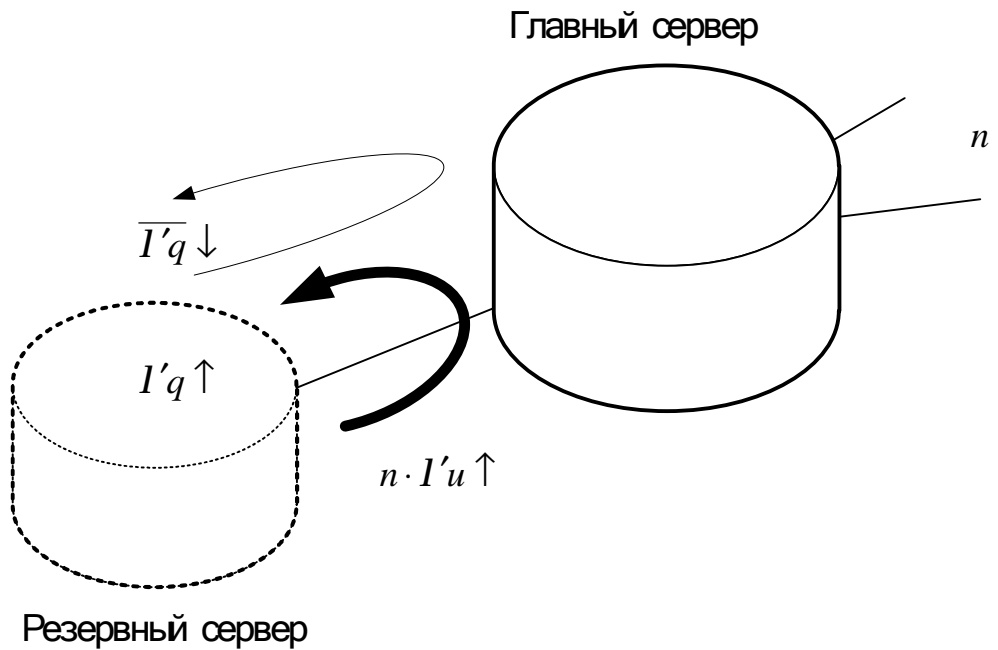


Рисунок 2 – Случай, когда доля реплицированных данных велика

Таким образом, в условиях заданных ограничений на временные задержки обработки запросов в различных элементах РБД необоснованный выбор немедленно реплицируемых фрагментов данных на узлах РБД снижает оперативность отклика на запросы.

Вопросам управления репликацией в РБД посвящено множество работ зарубежных ученых: Э. Таненбаума, Т. Коннолли, К. Луни, Б. Брила, С. Рига, Х. Кросинга, Ж. Йохансона, С. Марча, Ж. Науманна и др. [105, 48, 76, 85, 74, 57]. Среди отечественных исследователей вопросами, посвященными разработке моделей и алгоритмов управления репликацией в РБД, занимались: Д. А. Апанасевич, А. Ю. Иванов, В. Н. Кухарев, Л. И. Мейкшан, К. А. Карельская, И. В. Сергеев, В. Е. Белоусова, С. Д. Кузнецов, В. В. Кульба и др. [3, 8, 43, 56, 79, 41, 60, 45, 53, 55].

Несмотря на продолжительный период активного изучения данной тематики, вопрос разработки моделей и алгоритмов, обеспечивающих эффективное функционирование РБД, не потерял своей актуальности [85].

Исходя из данных положений тема, посвященная **модели и алгоритмам управления параметрами репликации в распределенной базе данных предприятия горнопромышленного комплекса**, является актуальной и обуславливает выбор объекта и предмета исследования.

Степень разработанности проблемы. Проблема уменьшения времени отклика РБД на запросы поднималась в работах Д. А. Апанасевича, А. Ю. Иванова, В. Н. Кухарева, Л. И. Мейкшан, С. Д. Кузнецова, В. В. Кульбы и др. При исследовании процесса репликации в РБД авторами учитывается ограниченный набор параметров, влияющих на время отклика РБД на запросы. При этом перспективным направлением является расширение существующих моделей на предмет учета более широкого спектра управляющих параметров, разработка алгоритмов, позволяющих учитывать не только обобщенные характеристики обслуживания запросов, но и их детальные представления.

При работе над диссертацией были изучены коллективные труды и отдельные монографии российских и зарубежных авторов, посвященные вопросам функционирования РБД.

Цель исследования: уменьшение времени отклика РБД предприятия ГПК на запросы при заданных ограничениях на временные задержки путем управления параметрами репликации.

Объект исследования: система управления репликацией в распределенной базе данных предприятия горнопромышленного комплекса.

Предмет исследования: способы, алгоритмы и методы управления репликацией в процессе обработки информации в РБД предприятия ГПК.

В диссертации поставлены следующие частные исследовательские задачи:

1. Провести сравнение известных моделей функционирования РБД при репликации, способов управления параметрами репликации, используемых в АСУ предприятием ГПК.

2. Разработать математическую модель отклика РБД на запросы при репликации, обеспечивающую управление совокупностью параметров репликации на уровне физической интерпретации при ограничениях на временные задержки обработки запросов в различных её элементах.

3. Разработать алгоритм вычисления оптимальной загруженности резервного узла РБД при репликации, позволяющий определять значения параметров репликации и обеспечивающий снижение среднего времени отклика РБД на запросы.

4. Разработать алгоритм выбора фрагментов данных для немедленной репликации, позволяющий по установленным значениям параметров репликации определять наборы фрагментов данных для немедленной репликации, при которых достигается снижение среднего времени отклика РБД на запросы.

5. Разработать алгоритм автоматизации процесса конфигурирования репликации в РБД предприятия ГПК, позволяющий вычислять параметры

репликации в РБД предприятия ГПК и формировать решения для администратора по её конфигурированию.

При написании работы в методологическом плане применялись положения теории вероятностей и математической статистики, теории массового обслуживания, дифференциального исчисления, теории эффективности целенаправленных процессов, статистического планирования экспериментов.

Гипотеза исследования заключается в предположении о том, что среднее время отклика распределенной базы данных на запросы при репликации может быть уменьшено за счет обоснованного выбора фрагментов данных для немедленной репликации с учетом полученных оптимальных значений интенсивностей обработки запросов на резервных серверах.

Диссертационная работа соответствует паспорту специальности 05.13.06 – «Автоматизация и управление технологическими процессами и производствами (промышленность)» по пункту №9: «Методы эффективной организации и ведения специализированного информационного и программного обеспечения АСУТП, АСУП, АСТПП и др., включая базы и банки данных и методы их оптимизации».

Научная новизна:

1. Математическая модель отклика РБД на запросы при репликации, базирующаяся на модели двухуровневой информационной системы с репликацией данных, отличающаяся учетом совокупности параметров: интенсивности запросов на обновление ($I'u$) и интенсивности поисковых запросов ($I'q$), обрабатываемых на резервных серверах, на уровне физической интерпретации.

2. Алгоритм вычисления оптимальной загруженности резервного узла при репликации в РБД, описываемой математической моделью отклика на запросы, основанный на модифицированном методе линейных комбинаций, отличающийся формированием ограничений, обеспечивающих режим функционирования РБД предприятия ГПК без блокировки.

3. Алгоритм выбора фрагментов данных для немедленной репликации, основанный на оптимизированном методе частично-целочисленного линейного программирования с аддитивным алгоритмом для задач с двоичными переменными, отличающийся процедурой принятия решения по критерию минимума объема пересылаемых реплик.

4. Способ управления репликацией в РБД, основанный на гибридном методе репликации, отличающийся автоматизацией подготовки принятия решения по управлению репликацией, защищенный патентом на изобретение.

Положения, выносимые на защиту:

1. Математическая модель отклика РБД на запросы при репликации.
2. Алгоритм вычисления оптимальной загруженности резервного узла РБД при репликации.
3. Алгоритм выбора фрагментов данных для немедленной репликации.
4. Способ управления репликацией в РБД.

Теоретическая значимость полученных решений заключается в разработке нового гибридного метода репликации, позволяющего за счет управления параметрами репликации в РБД предприятия ГПК подстраиваться под имеющиеся вычислительные и сетевые ресурсы с целью повышения её реактивности.

Практическая значимость заключается в разработке совокупности алгоритмов и доведении их до программной реализации, что подтверждается свидетельствами о государственной регистрации программ для ЭВМ № 2013611771 от 4 февраля 2013 года и № 2013616315 от 19 июня 2013 года, патентом на полезную модель № 126161 от 20 марта 2013 года и изобретением (положительное решение от 25.10.2013 о выдаче патента на изобретение "Способ репликации информации в распределенных базах данных с конкурентным распределением потоков" по заявке № 2012116021).

Полученные результаты могут использоваться на предприятиях ГПК с целью эффективной организации специализированного информационного обеспечения, создающей условия для снижения среднего времени отклика на

запросы при заданных ограничениях на временные задержки обработки запросов в различных элементах РБД.

Апробация. Основные положения и результаты работы были доложены и обсуждены на 17-ой Международной открытой научной конференции "Современные проблемы информатизации в моделировании и социальных технологиях" (г. Воронеж, 2012 г.), Международной молодежной научно-практической конференции СКФ МТУСИ "ИНФОКОМ-2012" (г. Ростов-на-Дону, 2012 г.), Всероссийской научно-технической конференции студентов, аспирантов и молодых ученых "Научная сессия ТУСУР-2013" (г. Томск, 2013 г.), Международной молодежной научно-практической конференции СКФ МТУСИ "ИНФОКОМ-2013" (г. Ростов-на-Дону, 2013 г.), Всероссийской научно-практической конференции "Многоядерные процессоры, параллельное программирование, ПЛИС, системы обработки сигналов" (г. Барнаул, Алтайский государственный университет, 2013 г.); 19-ой Международной открытой научной конференции "Современные проблемы информатизации" (г. Воронеж, 2014 г.)

Научные результаты изложены в 6 печатных статьях [30, 31, 32, 33, 34, 36], 6 тезисах докладов. По материалам исследования получен патент на полезную модель [80], получены 2 свидетельства о государственной регистрации программ для ЭВМ [100, 73], получено положительное решение от 25.10.2013 по заявке № 2012116021 о выдаче патента на изобретение "Способ репликации информации в распределенных базах данных с конкурентным распределением потоков".

Структура и объем работы. Диссертация состоит из введения, четырех глав и заключения. Диссертация содержит 142 страницы, 31 рисунок, 20 таблиц, 1 приложение. Список литературы содержит 111 наименований.

ГЛАВА 1. ОБСЛЕДОВАНИЕ МОДЕЛЕЙ ОБРАБОТКИ ИНФОРМАЦИИ В РАСПРЕДЕЛЕННЫХ БАЗАХ ДАННЫХ ПРИ РЕПЛИКАЦИИ

1.1 Описание технологий репликации в распределенных баз данных

Сфера информационных технологий характеризуется быстрым развитием используемых средств и методов обработки данных. Бурный рост вычислительных систем и высокие требования к техническим и информационным ресурсам привели к крупномасштабному развитию распределенных систем. Одним из элементов таких систем являются РБД, основная особенность функционирования которых состоит в том, что для пользователя не имеет значения, где пространственно расположены необходимые ему данные. Это достигается за счет автоматической репликации данных, которая позволяет поддерживать данные в актуальном состоянии.

РБД, по сравнению с централизованными базами данных, обладают дополнительными возможностями [94, 92]:

1. Независимость от транспортных систем при установке соединений, позволяющая передавать запросы и данные между распределенными узлами.
2. Ведение каталога, позволяющее сохранять сведения о распределении данных в сети.
3. Обработка распределенных запросов, позволяющая задействовать ресурсы различных узлов РБД.
4. Механизмы оптимизации распределенных запросов, позволяющие эффективно использовать ресурсы РБД при обработке распределенных запросов.
5. Управление защитой распределенных данных.
6. Поддержание целостности копируемых данных, предотвращающее сбои, вызванные потерей целостности данных при их копировании.

7. Восстановление, учитывающее возможность отказов в работе отдельных узлов и элементов ТКС.

Рекомендуемая архитектура РБД представлена на рисунке 3 [48].

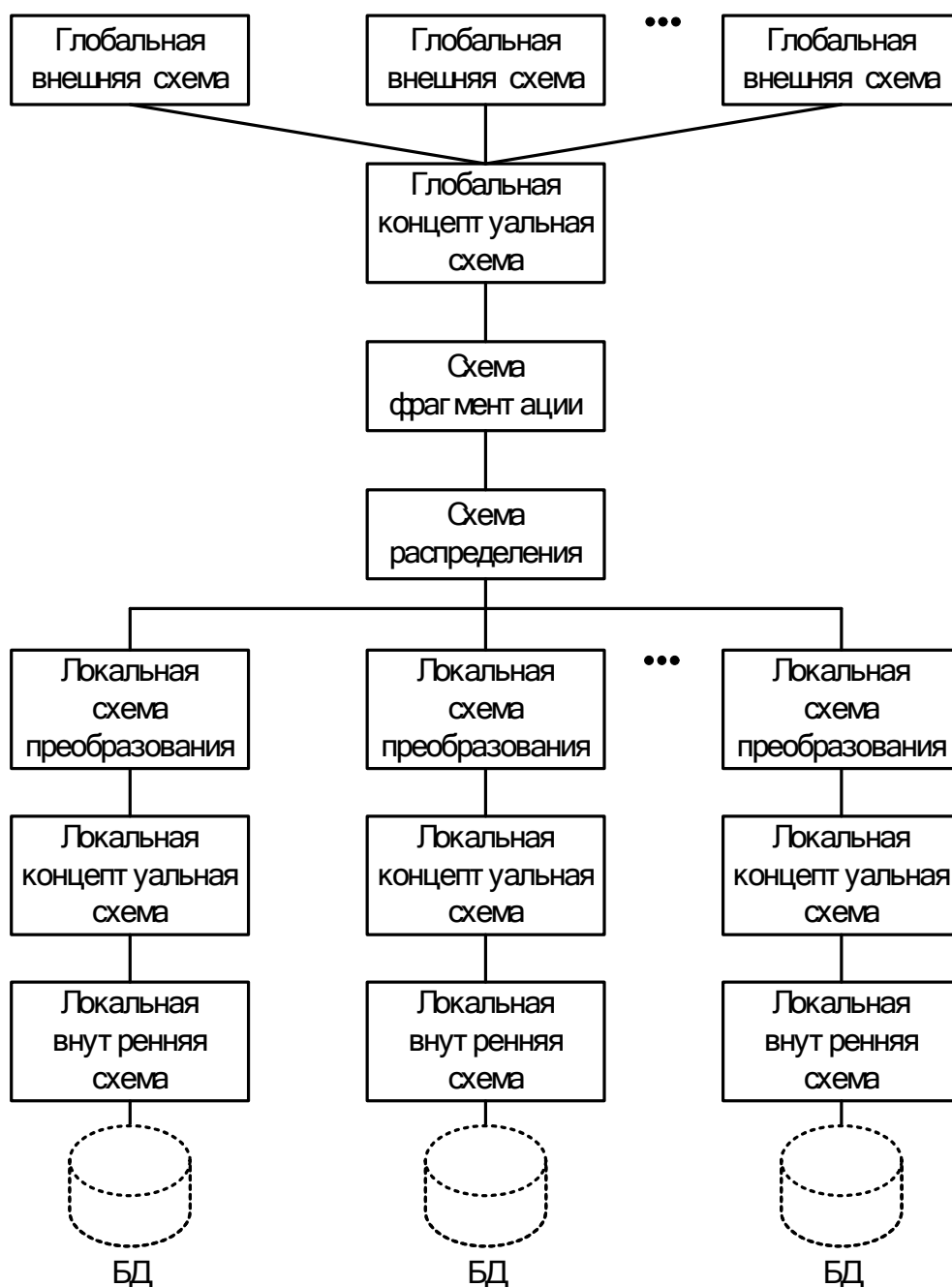


Рисунок 3 – Рекомендуемая архитектура РБД

При этом глобальная концептуальная схема – это высокоуровневое логическое описание всей совокупности информационных объектов, учтенных в базе данных.

Схема фрагментации описывает, как данные должны логически распределяться по разделам.

Схема размещения показывает, где расположены имеющиеся данные с учетом необходимости репликации.

Каждая локальная СУБД имеет свой собственный набор схем.

В настоящее время существуют четыре стратегии размещения данных в РБД:

1. Централизованное. На одном из узлов создается единственный экземпляр базы данных под управлением СУБД, доступ к которому имеют все пользователи сети.

2. Раздельное. Вся база данных разбивается на непересекающиеся фрагменты, каждый из которых размещается на одном из узлов РБД.

3. С полной репликацией. Каждый узел РБД содержит полную копию всей базы данных.

4. Избирательное. Наиболее общий и гибкий вариант размещения, представляющий собой комбинацию методов фрагментации, репликации и централизации.

Сравнительные характеристики различных стратегий размещения представлены в таблице 1.

Таблица 1 – Характеристики стратегий размещения данных в РБД

	Локализация ссылок	Надежность и доступность	Стоимость хранения	Затраты на передачу данных
Централизованное размещение	Самая низкая	Самая низкая	Самая низкая	Самые высокие
Раздельное размещение	Высокая	Низкая для элементов, высокая для системы в целом	Самая низкая	Низкие
Полная репликация	Самая высокая	Сама высокая	Самая высокая	Высокие при обновлении, низкие при чтении
Избирательная репликация	Высокая	Низкая для элементов, высокая для системы в целом	Средняя	Низкие

Различают два режима работы РБД: режим обслуживания запросов и режим репликации. К задачам, решаемым в рамках репликации, относят [8]:

- поддержание узлов данных в актуальном состоянии;
- обеспечение резервирования данных;
- объединение информации из нескольких массивов данных;
- поддержание в работоспособном состоянии узлов с непостоянным соединением с ядром системы.

Классификация режимов репликации РБД, в зависимости от характера решаемых задач, представлена на рисунке 4 [8].

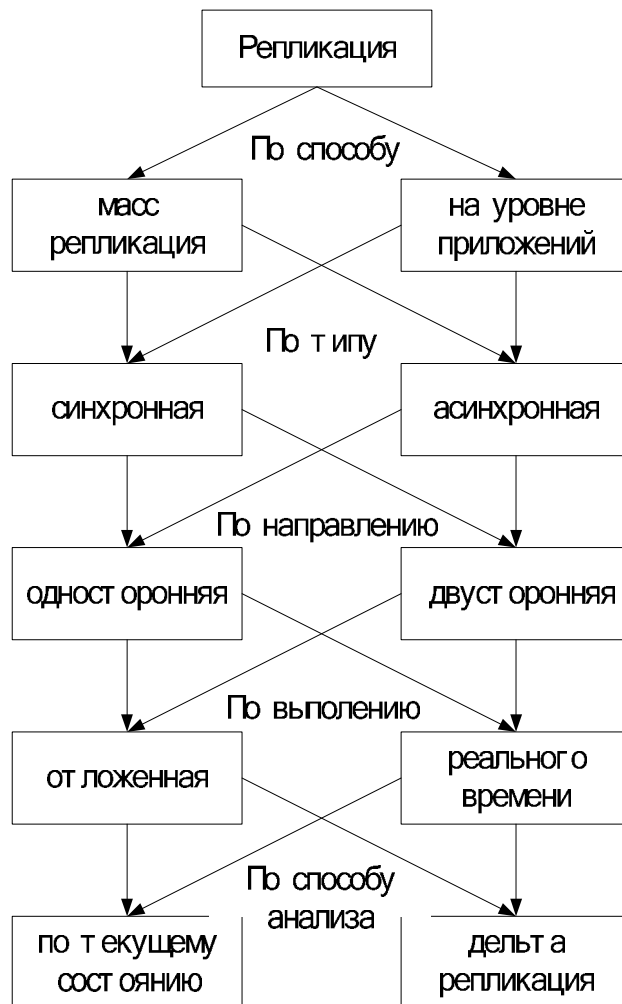


Рисунок 4 – Классификация режимов репликации

Большинство из представленных режимов репликации реализуется в рамках функционирования существующих РБД. При этом различные режимы

репликации предъявляют неодинаковые требования к ресурсам РБД. Поэтому вопросы обоснованного выбора параметров репликации являются актуальными.

Каждый раз при изменении копии она начинает отличаться от всех прочих. Соответственно, для сохранения непротиворечивости эти изменения должны быть перенесены и на остальные копии. При этом обновления должны распространяться по узлам РБД как можно быстрее [85].

При синхронной репликации транзакции не фиксируются до тех пор, пока на всех узлах, содержащих копии данных, не будут внесены изменения. Это позволяет добиться максимальной актуальности данных, но вносит дополнительные задержки, вызванные блокировкой обновляемых данных. Реализация синхронной репликации возможна только при наличии надежных высокоскоростных каналов связи.

При асинхронной репликации изменения в копии данных вносятся независимо, что позволяет избавиться от задержек блокировок. Однако при таком подходе снижается текущая актуальность данных. Режим асинхронной репликации предъявляет более гибкие требования к доступным ресурсам.

1.2 Описание информационного обеспечения управления предприятием ГПК

Информационное обеспечение управления предприятием ГПК предназначено для оптимизации задач диспетчерского, производственно-технологического, организационно-экономического управления технологическими процессами шахты, а также повышения безопасности работ проводимых в шахте [1].

Информационное обеспечение управления шахтой представляет собой многоуровневую систему сбора данных и управления с определенным кругом задач, решаемых на каждом уровне, рисунок 5:

- нижний уровень – датчики и исполнительные механизмы для поверхностных и подземных объектов;

- средний уровень – контроллерное и сетевое оборудование локальных систем контроля и управления поверхностными и подземными объектами;

- верхний уровень – серверы, операторские станции диспетчеров и руководства шахты, коммуникационное оборудование, локальные вычислительные сети и программное обеспечение.

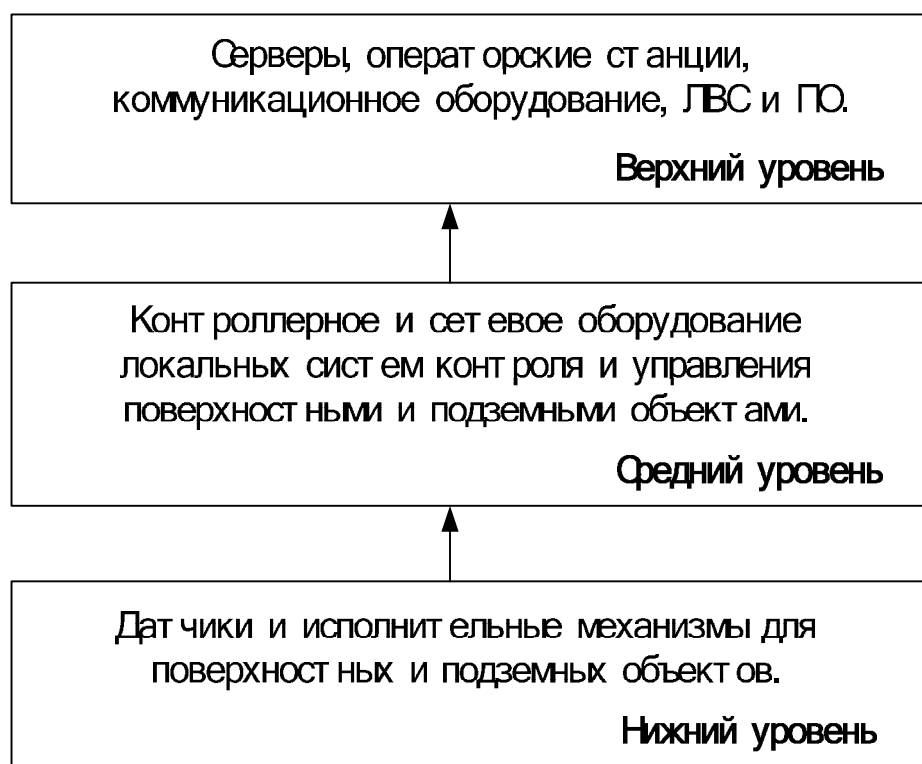


Рисунок 5 – Уровни системы сбора данных и управления шахтой

В соответствии с федеральной целевой программой "Национальная технологическая база" на 2013-2016 годы одним из перспективных направлений исследований по повышению безопасности жизнедеятельности по состоянию горного массива угольной шахты является разработка интегрированной системы управления её безопасностью, включающей:

– систему сбора, учета, анализа и прогноза информации по совокупности всех факторов (горного массива, техногенных, подземного персонала) для обоснования допустимых рисков и выработки эффективных упреждающих воздействий в масштабе реального времени;

– систему обоснования допустимых рисков и выработки эффективных упреждающих воздействий и рекомендаций по устранению сложившейся нештатной или чрезвычайной ситуаций;

– систему двунаправленной передачи информации и управляющих команд между персоналом угольных шахт, оборудованием и диспетчерской, со способностью сохранения работоспособности в аварийной ситуации и до окончания воздействия аварийных факторов;

– типовые технологии управления рисками (идентификации рисков, удаленного контроля, мониторинга и моделирования процессов, прогноза рисков, выработки и обоснования упреждающих воздействий, реализации мер управления и анализа их эффективности) для различных сценариев развития угроз (моделей угроз), позволяющие обеспечить интеграцию с системами аэрогазового контроля и АСУ;

– типовую систему менеджмента риска по требованиям международных стандартов (ISO 9001, ISO/IEC 15288, 31000 и др.) и документы системы сертификации средств и систем жизнеобеспечения подземного персонала угольных шахт для подтверждения их безопасности, обеспечивающую интеграцию с нормативными документами и Российским законодательством;

– моделирующие тренажеры и программы обучения и повышения квалификации персонала для системного управления рисками на угольной шахте;

– стандартные унифицированные программно-аппаратные информационные интерфейсы сопряжения подсистем управления безопасностью угольной шахты по состоянию горного массива, техногенных факторов и подземного персонала.

При этом реализация интегрированной системы управления безопасностью угольной шахты невозможна без комплексного подхода к сбору и обработке данных от следующих элементов:

- программно-аппаратного комплекса автоматизированного моделирования характеристик горного массива в масштабе реального времени по информации датчиков контроля и предупреждения об опасности;
- автоматизированного комплекса геосейсмического мониторинга горного массива, позволяющего обеспечить непрерывное получение данных сейсморазведки перед забоем и получение заблаговременной информации о геосейсмических изменениях окружающего массива;
- автоматизированного комплекса мониторинга параметров шахтного пространства путем дистанционного измерения температуры, деформации растяжения-сжатия, ускорения, вибрации, подвижки грунтов;
- автоматизированного комплекса мониторинга геодинамических явлений, позволяющего обеспечить непрерывное получение и обработку данных о напряженно-деформированном состоянии окружающего массива;
- автоматизированного комплекса виброакустического мониторинга горного массива, обеспечивающего обмен данными по стандартным каналам, позволяющим реализовать непрерывное получение и обработку данных о возникающих процессах в окружающем шахту массиве, получение возможности организации противоаварийных мероприятий на ранней стадии возникновения опасности;
- автоматизированного комплекса радиологического мониторинга горного массива, включающего контроль наличия и концентрации радона в рудничной атмосфере, позволяющего на ранней стадии определить зарождение эндогенных пожаров в выработанных пространствах угольных шахт, микросейсмическую активность пластов горного массива, а также обеспечить непрерывный контроль за эквивалентной равновесной объемной активностью аэрозолей дочерних продуктов радона в непроветриваемых или слабопроветриваемых объемах шахты;
- автоматизированного комплекса радиозондирования горного массива, позволяющего обеспечить непрерывное получение и обработку данных о неоднородностях и возникающих изменениях в структуре горного массива для точной локализации аномалий;

– автоматизированного комплекса мониторинга и контроля гидрогеологической обстановки, позволяющего определить устойчивость горного массива, временной и постоянной крепи, а также давления грунтовых вод, в том числе оборудования автоматизированной системы интеллектуального адаптивного управления водоотливными установками шахты;

– испытательного комплекса для изучения и мониторинга физико-механических свойств горных пород, позволяющего обеспечить повышение безопасности горных работ за счет заблаговременного предупреждения об опасности, профилактики геодинамических явлений, квазистатических движений в массиве горных пород и других опасных факторов.

Интеграция перечисленных комплексов в автоматизированной системе достигается в результате организации эффективной системы управления данными. При этом распределенный характер и нестабильность условий функционирования шахты создает необходимость использовать для управления данными распределенные базы данных.

Так, РБД шахты "ШахтИнвестКузбасс", состоящая из одного главного сервера и множества резервных серверов, обрабатывает данные от следующих служб, рисунок 6:

- аэрогазового контроля;
- вентиляционного проветривания;
- теплоэнергетики шахты;
- пожарного водоснабжения;
- водоотлива;
- электроснабжения;
- конвейерного транспорта;
- технологического оборудования очистных и подготовительных участков шахты;
- геосейсмического мониторинга;
- мониторинга параметров шахтного пространства;
- виброакустического мониторинга горного массива;

- радиологического мониторинга горного массива;
- мониторинга и контроля гидрогеологической обстановки;
- радиозондирования горного массива;
- системы наблюдения и оповещения персонала.

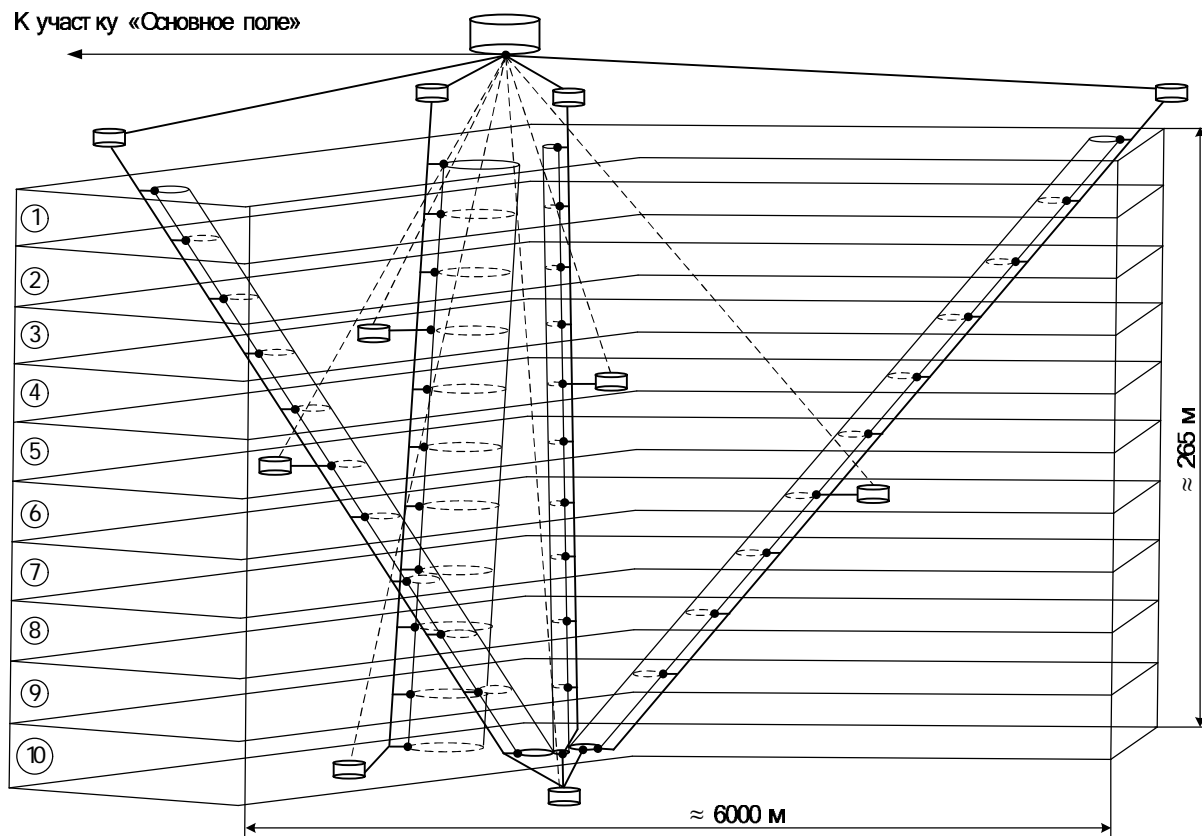


Рисунок 6 – Обобщенная структура РБД предприятия ГПК "ШахтИнвестКузбасс"

На рисунке 6 выделены угольные пласты 1-10, соответственно: Выклинившийся, Надартельный 2, Артельный, Абрамовский, Лыжинский, Кумпановский, Верхний, Двойной-Промежуточный (1-ая пачка), Двойной-Промежуточный (2-ая пачка).

Общая площадь территории, на которой размещается РБД предприятия ГПК «ШахтИнвестКузбасс» составляет более 30 квадратных километров (рисунок 7).

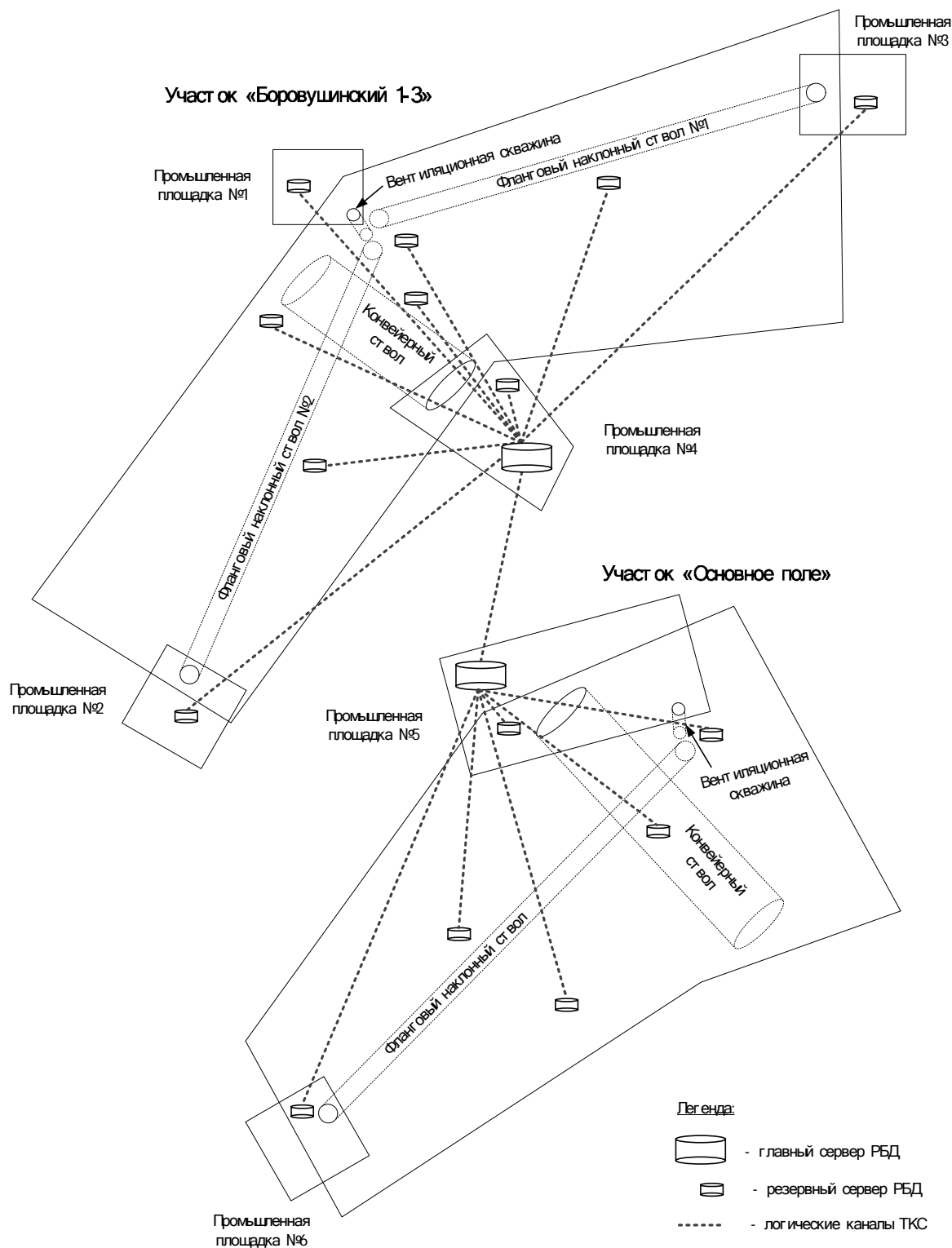


Рисунок 7 – План размещения главного и резервных серверов РБД предприятия ГПК

АСУ предприятием ГПК создаются в виде совокупности обеспечивающих подсистем (информационная, математическая, программная, техническая, организационная) и функциональных (техническая подготовка угледобычи, технико-экономическое планирование,

материально-техническое снабжение, оперативное управление, бухгалтерский учет) [15, 89, 90]. Их взаимодействие строится на основе общего информационного фонда, обеспечивающего единый общесистемный подход на всех этапах сбора, обработки и выдачи информации. При этом к подсистеме информационного обеспечения выдвигается ряд требований со стороны других подсистем по обеспечению оптимальным объемом информации в требуемые сроки, рисунок 8.

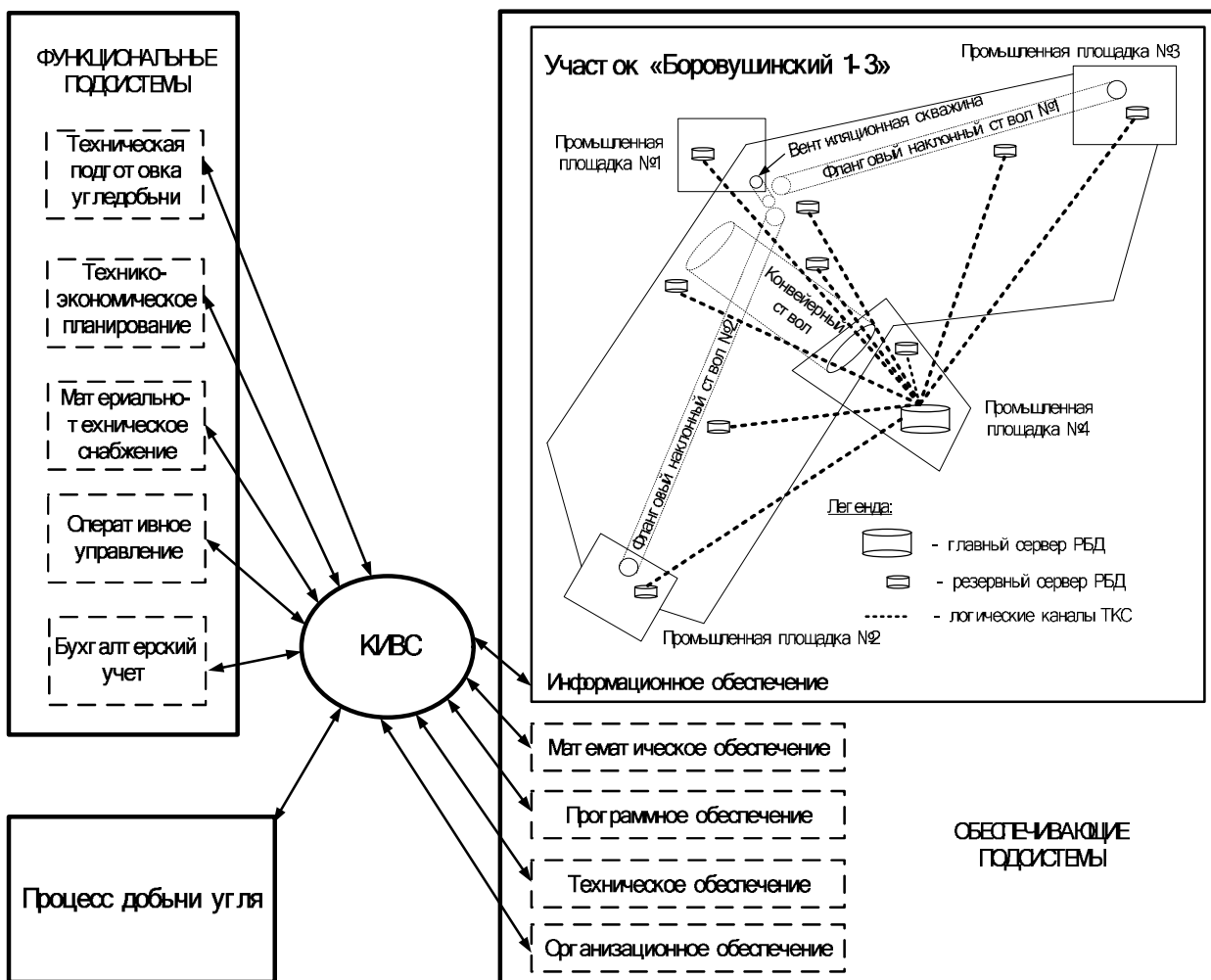


Рисунок 8 – Место РБД в информационном обеспечении управления предприятием ГПК

1.3 Описание подходов к моделированию процессов, протекающих в распределённых базах данных

РБД можно представить совокупностью N серверов, объединенных телекоммуникационной подсистемой, и совокупностью групп рабочих станций, соединенных с серверами средствами подсистемы доступа. При этом каждая группа рабочих станций имеет прямое соединение только с одним сервером РБД (резервным сервером РБД).

Телекоммуникационная подсистема, выполняющая функции узлов коммутации, связанных каналами передачи данных, предназначена для организации логических каналов между точками логического подключения серверов РБД.

В качестве подсистемы доступа выступает совокупность коммуникационного оборудования и каналов передачи данных, отвечающих за обмен данными между серверами РБД и рабочими станциями [39].

Существует множество подходов к моделированию процессов в РБД, однако все они имеют недостатки, сужающие область их применения. Так, в исследованиях Д. А. Апанасевича для построения математических моделей информационных процессов применяется математический аппарат теории конечных автоматов [3]. При этом модель процесса информационного обмена представляется конечным автоматом, который в любой момент времени t находится в некотором состоянии $S(t) = \{s_1, s_2, s_3\}$, где:

- состояние s_1 – данные отправляются;
- состояние s_2 – данные принимаются;
- состояние s_3 – система ожидает сеанса связи.

При наступлении определенных событий (к примеру, поступление запроса на обновление данных) состояние РБД изменяется. Свойства информационного процесса в зависимости от состояний конечного автомата определяются системой:

$$P(s) = \begin{cases} D, & \text{если } S(t) = s_1 \\ R, & \text{если } S(t) = s_2 \\ O, & \text{если } S(t) = s_3 \end{cases}, \quad (1)$$

где D – автомат является отправителем, R – автомат является получателем, O – автомат не является ни отправителем, ни получателем.

Основное преимущество модели – простота описания, но при этом отсутствует возможность учета вероятностно-временных характеристик процессов, протекающих в РБД.

Другой подход для математического описания процесса функционирования РБД предложен в работе Т. Коннолли на основе теории вероятностей и теории массового обслуживания. РБД представляется как совокупность некоторого множества независимых файлов с заданными на них подмножествами запросов на обновление и получение данных. При этом объемы порождаемых в ТКС данных зависят от узлов-источников. В течение каждой единицы времени по ТКС пересылается некоторый объем данных, связанный с распределением копий файлов по РБД.

$$V = \frac{1}{\lambda} \sum_{i=1}^m \sum_{j=1}^n \lambda_{ij} V_{ij}, \quad (2)$$

где n – количество узлов РБД, m – количество файлов РБД, λ_{ij} – интенсивность запросов к i -му файлу в j -ом узле; V_{ij} – объем пересылаемых данных при запросе к i -му файлу в j -ом узле; λ – суммарная интенсивность запросов по всем узлам и файлам РБД (аддитивная свертка).

Данная модель описывает транзакционные особенности запросов и не учитывает влияние характеристик ТКС и параметров репликации данных.

Наиболее полно процесс обработки запросов в РБД исследован в работе А. Ю. Иванова [41]. Обработка запросов в РБД описывается в рамках обслуживания заявок в стохастической сети массового обслуживания (СеМО), рисунок 9. При этом системы массового обслуживания (СМО), составляющие СеМО, интерпретируются типовыми элементами РБД: фрагментами сети доступа (СМО 1-го типа), серверами РБД (СМО 2-го типа), каналами передачи данных (СМО 3-го типа), коммуникационным оборудованием (СМО 4-го типа). В качестве источников заявок выступают поисковые запросы к локальным серверам.

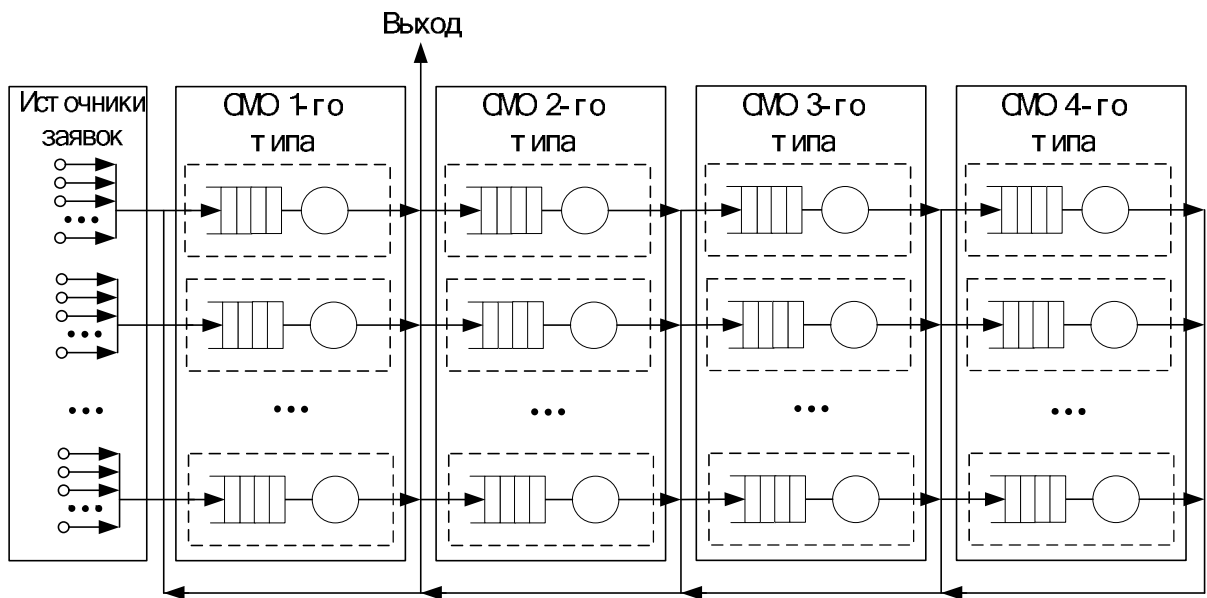


Рисунок 9 – Обобщенная структура модели обработки запросов к РБД

Такая модель обработки запросов к РБД описывается функционалом:

$$t = F(V_x, T, S, L), \quad (3)$$

где t – время выполнения запроса в РБД; V_x – объемные характеристики РБД (размеры запросов, размеры откликов, количество копий данных, размеры копий данных и др.); T – характеристики технических средств в РБД (производительность серверов, доступная пропускная

способность каналов передачи данных, размеры буферов в коммуникационном оборудовании, доступное количество памяти в серверах и др.); S – характеристики структуры РБД (топология сети, количество серверов РБД, распределение заявок по серверам РБД и др.); L – характеристики информационных потоков в сети (интенсивности поступления поисковых запросов и запросов на обновление и др.).

Детально СМО каждого типа можно представить сетью Петри – графа особого вида, состоящего из вершин двух типов: позиций и переходов, соединенных ориентированными дугами, причем каждая дуга может связывать лишь разнотипные вершины. Вершины-позиции обозначаются кружками, вершины-переходы – прямоугольниками [52]. Так, сеть Петри, описывающая один из этапов обработки запроса в РБД, представлена на рисунке 10 [70].

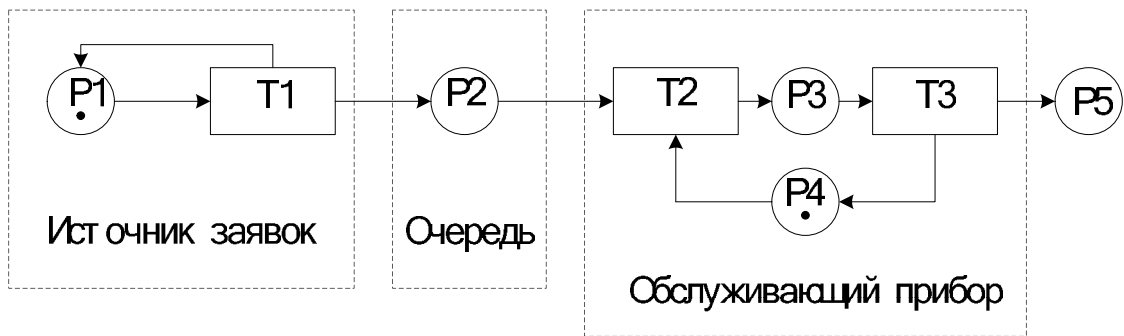


Рисунок 10 – Сеть Петри, детализирующая этап обработки запроса в РБД

Применительно к РБД в качестве источника заявок могут выступать: потоки запросов на обновление и поисковых, потоки заявок на передачу удаленных запросов и обновлений.

Очередь может быть ассоциирована с ожиданием обслуживания на главных и резервных серверах, а также с ожиданием освобождения участков ТКС при передаче между различными узлами РБД.

Обслуживающий прибор может описывать работу главных и резервных серверов, а также коммуникационное оборудование.

Маркер в позиции $P1$ соответствует готовности источника к выдаче заявок. Обратная связь перехода $T1$ с позицией $P1$ необходима для генерации

последующих заявок. Позиция Р2 моделирует очередь заявок на обслуживание. Маркер в позиции Р4 моделирует свободное состояние обслуживающего прибора, а Р3, когда обслуживающий прибор занят. При этом переходы Т2 и Т3 определяют распределение заявок по позициям Р3 и Р4.

Существенным недостатком такой модели является отсутствие учета параметров репликации РБД.

В исследовании Ж. М. Йохансона, Ж. Д. Науманна, С. Т. Марча рассматривается подход к определению времени обновления данных в РБД. В общем виде расчет времени обновления описывается выражением [105]:

$$T_u = 3 \cdot [\text{MAX}_{j=1}^n [(\sum_{i=1}^j S_i) + R_j]] + \sum_{i=1}^n S_i, \quad (4)$$

где S_i – время передачи запроса от источника транзакции до i -го узла, R_j – время передачи ответа от j -го узла, n – число узлов с фрагментами данных.

Другим вариантом, совмещающим достоинства описанных моделей, является математическое описание процесса отклика РБД на запросы при репликации представленное в работе Л. И. Мейкшан [60]. Так, процессы обслуживания заявок разного вида определяют следующие величины:

- среднее время обработки поискового запроса t_q (на удаленном сервере) и t'_q (на локальном сервере);
- среднее время обработки запроса на обновление t_u (на удаленном сервере) и t'_u (на локальном сервере);
- среднее время, требуемое удаленному серверу на отправку одного сообщения для обновления фрагмента данных на локальном сервере t_r .

Среднее время отклика РБД на запросы при репликации определяется на основе выражения:

$$R = h[W_w + t'_q] + (1 - h)[2t_r + W_c + t_q], \quad (5)$$

где W_w и W_c – средние значения времени пребывания запроса в очереди на обработку для локального и удаленного серверов, соответственно. Коэффициент h определяет долю запросов, обрабатываемых локальным сервером. Время доставки сообщения между удаленным и локальным серверами рассматривается как постоянная величина t_r .

Основным недостатком данной модели является допущение о том, что величина t_r постоянна для всех участков ТКС и типов запросов. Тогда как в условиях ограниченных ресурсов величина t_r является переменной и существенно зависит от множества факторов: протяженности участка сети; количества и характеристик коммуникационного оборудования, задействованного в передаче; ресурсов пропускной способности сети; технологий передачи; схемы владения данными; типа репликации и др. С другой стороны, выбор в качестве управляющего параметра коэффициента h , характеризующего долю реплицированных данных, лишает модель гибкости с точки зрения выбора фрагментов данных для немедленной репликации.

В связи с этим в настоящем диссертационном исследовании выполнена доработка модели, представленной в работе Л. И. Мейкшан [60], на предмет снятия ограничения по управляющим параметрам и детализации времени ожидания и передачи данных по ТКС.

1.4 Описание процесса репликации в РБД предприятия ГПК

"ШахтИнвестКузбасс"

В исследованиях А. И. Иванова, В. В. Кульбы, Э. Таненбаума [55, 41, 85], посвященных вопросам функционирования РБД, обосновывается выбор системы её значимых свойств:

1. Целостность РБД – это способность данных в РБД сохранять однозначность и информационное содержание.

2. Полнота РБД – это возможность выборки из РБД необходимого количества информации.

3. Достоверность – это степень соответствия данных из РБД реальному состоянию объектов предметной области.

4. Реактивность – это способность в среднем выполнять запросы за время, не превышающее декларативное.

5. Защищенность – это способность РБД сохранять информацию в неизменном виде при наличии воздействий различного характера.

Поскольку РБД разрабатываются и внедряются с целью повышения оперативности решения задач в информационных системах, то в качестве значимого свойства РБД признают реактивность [41].

В настоящее время в РБД предприятия ГПК "ШахтИнвестКузбасс" репликация проводится по расписанию, что порождает высокую интенсивность удаленных запросов. При этом нерациональная нагрузка на вычислительные и сетевые ресурсы негативно сказывается на временных задержках обслуживания запросов на различных этапах их обработки, что, в свою очередь, влечет за собой снижение реактивности РБД.

Рациональный выбор фрагментов немедленно реплицируемых данных необходимо проводить на основе параметров, существенно влияющих на реактивность РБД, а именно: интенсивности поступления запросов на выборку, обрабатываемых на резервном сервере $I'q$, и интенсивности поступления запросов на обновление к резервному серверу $I'u$. При этом в

качестве показателя интенсивности реактивности РБД целесообразно использовать среднее время её отклика на запросы при репликации [60].

В результате возникает противоречие, состоящее в том, что уменьшение объема реплицируемых данных (снижение интенсивности поступления запросов на обновление к резервному серверу) порождает увеличение числа удаленных заявок на обслуживание (уменьшение интенсивности поступления запросов на выборку, обрабатываемых на резервном сервере). Как следствие, среднее время отклика РБД на запросы увеличивается за счет удаленного обслуживания запросов. С другой стороны, увеличение объема реплицируемых данных (увеличение интенсивности поступления запросов на обновление к резервному серверу) приводит к росту числа заявок на репликацию, что приводит к дефициту вычислительных и сетевых ресурсов и, как следствие, среднее время отклика РБД на запросы увеличивается за счет роста временных задержек на различных этапах обслуживания запросов.

Таким образом, в условиях заданных ограничений на временные задержки обработки запросов в различных элементах РБД необоснованный выбор фрагментов данных для немедленной репликации снижает реактивность РБД.

1.5 Постановка задачи исследования

Дано:

1. Распределенная база данных в виде совокупности отдельных ее элементов: главного сервера, резервных серверов, участков ТКС между серверами.

2. Множество значений интенсивностей поступления запросов различных типов $\Lambda = \{I_q, I'_q, I_u, I'_u\}$.

3. Множество значений, характеризующих среднее время обслуживания запросов (\bar{T}) на различных участках РБД.

4. Множество фрагментов данных $\{X\}$, участвующих в процессе репликации: немедленной $\{X'\}$, отложенной $\{X''\}$.

Требуется:

Используя научно-обоснованные методы разработать адекватную модель отклика РБД на запросы при репликации, учитывающую совокупность параметров: интенсивность запросов на обновление ($I'u$) и интенсивность поисковых запросов ($I'q$), обрабатываемых на резервных серверах, на уровне физической интерпретации. На основе полученной модели разработать алгоритм вычисления оптимальной загруженности резервного узла РБД при репликации и алгоритм выбора фрагментов данных для немедленной репликации, использование которых в совокупности позволит построить алгоритм автоматизации процесса конфигурирования репликации в РБД предприятия ГПК, позволяющий вычислять параметры репликации в РБД предприятия ГПК и формировать решения для администратора по её конфигурированию.

Ограничения:

Учитывая, что в исследовании рассматривается РБД, которая должна обладать высокой способностью к масштабированию и функционирует в условиях воздействия множества случайных факторов, то возможны ситуации дефицита вычислительных и сетевых ресурсов. Следовательно, полученное решение необходимо рассматривать в рамках заданных ограничений на временные задержки обработки запросов в различных элементах РБД.

Целевой функционал задачи уменьшения среднего времени отклика РБД на запросы при заданных ограничениях на временные задержки обработки запросов в различных её элементах за счет обоснованного выбора значений интенсивностей обработки запросов на резервных серверах:

$$\bar{T}(I'q(\{X'\}, type), I'u(\{X'\}, type)) \rightarrow \min_{I'q, I'u} \left| \begin{array}{l} (\bar{T}) \\ "A", \bar{T} \leq T\delta \end{array} \right. , \quad (6)$$

где "А" – совокупность ограничений по свойствам: результативность и ресурсоемкость [68]; $I'u$ – интенсивность запросов на обновление, обрабатываемых на резервном сервере; $I'q$ – интенсивность поисковых запросов, обрабатываемых на резервном сервере; Td – допустимое время отклика на запросы, $type$ – тип реплицируемых фрагментов данных (немедленно, отложено).

Согласно приказу Министерства связи и массовых коммуникаций Российской Федерации от 2 сентября 2011 года № 221 допустимое время доступа к данным в распределенных системах не должно превышать трех секунд [72]. В исследованиях А. И. Губинского [27] приводится обоснование наиболее удобного для пользователей, с эргономической точки зрения, временного интервала отклика информационных систем: две – четыре секунды. Данные рекомендации можно отнести и к РБД. Таким образом, допустимое время отклика РБД на запросы находится в диапазоне от двух до трех секунд.

Требуется решить следующие частные задачи:

1. Разработать математическую модель отклика РБД на запросы при репликации, обеспечивающую управление совокупностью параметров репликации на уровне физической интерпретации при ограничениях на временные задержки обработки запросов в различных её элементах.

2. Разработать алгоритм вычисления оптимальной загруженности резервного узла РБД при репликации, позволяющий определять значения параметров репликации и обеспечивающий снижение среднего времени отклика РБД на запросы.

3. Разработать алгоритм выбора фрагментов данных для немедленной репликации, позволяющий по установленным значениям параметров репликации определять наборы фрагментов данных для немедленной репликации, при которых достигается снижение среднего времени отклика РБД на запросы.

4. Разработать алгоритм автоматизации процесса конфигурирования репликации в РБД предприятия ГПК, позволяющий вычислять параметры

репликации в РБД предприятия ГПК и формировать решения для администратора по её конфигурированию.

Для решения задачи исследования предлагается использование модели двухуровневой информационной системы с репликацией данных, представленной в работе Л. И. Мейкшан [60], с учетом ее доработки на предмет снятия ограничения по управляющим параметрам и детализации времени ожидания и передачи данных по ТКС.

Предполагаемый эффект от решения поставленной задачи состоит в снижении среднего времени отклика РБД на запросы при репликации более чем на 6%. При этом самую минимизацию задержек репликации можно представить как новый гибридный механизм репликации, позволяющий за счет управления параметрами репликации в РБД предприятия ГПК подстраиваться под имеющиеся вычислительные и сетевые ресурсы с целью повышения её реактивности.

Выводы по главе 1

1. Различные технологии репликации данных обладают рядом достоинств и недостатков, вопросы выбора наиболее рационального варианта организации внутрисистемных средств тиражирования данных в РБД требуют изучения с использованием соответствующих математических моделей. При этом нерациональный выбор фрагментов данных для немедленной репликации увеличивает среднее время отклика РБД на запросы.

2. Существующие в настоящий момент модели функционирования РБД имеют множество ограничений, не позволяющих адекватно описывать процессы отклика на запросы при репликации. Вследствие этого возникает противоречие между необходимостью выбора фрагментов данных для немедленной репликации, создающего условия для оптимизации среднего времени отклика РБД на запросы, и адекватностью существующих моделей.

3. Система управления шахтой представляет собой многоуровневую систему сбора данных и управления с определенным кругом задач, эффективная интеграция которых в одной автоматизированной системе достигается в результате правильной организации системы управления данными. При этом распределенный характер и нестабильность условий функционирования шахты делает целесообразным использование для управления данными технологий РБД.

4. Тема диссертационного исследования актуальна. Научность задачи исследования состоит в разработке алгоритма автоматизации процесса конфигурирования репликации в РБД предприятия ГПК, позволяющего за счет управления параметрами репликации в РБД предприятия ГПК подстраиваться под имеющиеся вычислительные и сетевые ресурсы, для повышения её реактивности.

5. Предполагаемый эффект от решения поставленной задачи состоит в снижении среднего времени отклика РБД предприятия ГПК на запросы при репликации более чем на 6%.

6. Задача исследования относится к классу задач автоматизации процесса конфигурирования репликации в РБД предприятия ГПК для эффективной организации и ведения специализированного информационного и программного обеспечения, включая базы и банки данных и методы их оптимизации.

ГЛАВА 2. РАЗРАБОТКА МОДЕЛИ ОТКЛИКА РБД НА ЗАПРОСЫ ПРИ РЕПЛИКАЦИИ

Данная глава посвящена разработке математической модели отклика РБД на запросы при репликации, базирующейся на модели двухуровневой информационной системы с репликацией данных, отличающейся учетом совокупности параметров: интенсивности запросов на обновление ($I'u$) и интенсивности поисковых запросов ($I'q$), обрабатываемых на резервных серверах, на уровне физической интерпретации.

Новизна и возможность применения представленной модели подтверждается патентом на полезную модель № 126161 "Система децентрализованного управления структурой распределенной базы данных" [80] и изобретением (положительное решение от 25.10.2013 о выдаче патента на изобретение "Способ репликации информации в распределенных базах данных с конкурентным распределением потоков" по заявке № 2012116021).

2.1 Выбор математического аппарата для разработки модели

Эффективность отклика РБД определяется результативностью, ресурсоемкостью и оперативностью исхода операции [18].

РБД автоматизированных систем разрабатываются и внедряются, в первую очередь, с целью повышения оперативности автоматизированного управления предприятием [55, 41]. При этом под оперативностью автоматизированного управления предприятием, в общем случае, понимают не превышение допустимого лимита расхода времени, потребного для исполнения управления [54]. Данное положение транспонируется на составляющую управления – отклик РБД на запрос. Распределенный характер БД обуславливает необходимость рассмотрения отклика на запросы в совокупности с процессом репликации. В зависимости от типа систем и внешних воздействий операции могут быть детерминированными, вероятностными и неопределенными [54].

Выбор конкретного типа критерия основывается на степени детализации условий функционирования системы: все условия определены, система функционирует в условиях риска и система рассматривается в условиях неопределенности. Так как модель отклика РБД на запросы при репликации рассматривается, когда все условия функционирования определены в виде средних характеристик, полученных в рамках теории математической статистики, то используют детерминированный критерий.

Для оценки детерминированных операций применяют три типа критериев [54]:

1. **Критерий пригодности**, согласно которому процесс отклика РБД на запросы при репликации считается эффективным, если все частные показатели исхода операции принадлежат области адекватности:

$$K^{npuz} : (\forall i)(y_i^j \in d \mid d_i \rightarrow y_i^{don}, i \in \langle \mathcal{E}, R, O \rangle), \quad (7)$$

где i – индекс показателя отдельного свойства (результативности, ресурсоемкости и оперативности) j -го процесса, y – значение показателя эффективности, d – область адекватности.

2. **Критерий оптимальности**, определяющий правило, по которому процесс отклика РБД на запросы при репликации считается эффективным, если существует хотя бы один частный показатель исхода операции, значение которого принадлежит области адекватности, а её радиус по этому показателю оптимален:

$$K^{onm} : (\exists i)(y_i^j \in d \mid d_i \rightarrow y_i^{opt}, i \in \langle \mathcal{E}, R, O \rangle). \quad (8)$$

3. **Критерий превосходства**, определяющий правило, по которому процесс отклика РБД на запросы при репликации считается эффективным, если все частные показатели исхода операции принадлежат области

адекватности, а радиус области адекватности по этим показателям оптимален:

$$K^{прев} : (\forall i)(y_i^j \in d \mid d_i \rightarrow y^{opt}, i \in \langle \mathcal{E}, R, O \rangle). \quad (9)$$

Применение критериев оптимальности или превосходства предполагает получение глобальных оптимумов по показателям исхода операции: результативности, ресурсоемкости и оперативности, а их использование является задачей отыскания глобального экстремума функции. Система управления ГПК является сложной организационно-технической системой, целевая функция для которой не формализована, следовательно, достижение глобального оптимума затруднительно.

Так как доминирующее свойство процесса отклика запросов в РБД при репликации – оперативность [55, 41], то эффективность процесса необходимо рассматривать с точки зрения нахождения локального минимума среднего времени её отклика на запросы при репликации с учетом ограничений на результативность и ресурсоемкость.

Особенность РБД состоит в том, что все её узлы территориально рассредоточены и, как правило, функционируют в условиях воздействия случайных факторов. Применительно к РБД горнопромышленного комплекса по добыче угля дополнительными случайными факторами, влияющими на эффективность её функционирования, являются: повышенная влажность, запыленность, сейсмоактивность, низкие и высокие температуры и вибрации. В связи с этим устойчивость значений выбранных показателей эффективности процесса обработки информации в РБД при репликации должна рассматриваться в статистическом смысле.

На рисунке 11 представлен общий вид зависимости среднего времени отклика РБД на запросы при репликации от интенсивности обрабатываемых резервным узлом запросов: поисковых и на обновление. Видно, что существуют значения интенсивностей, при которых среднее время отклика минимально.

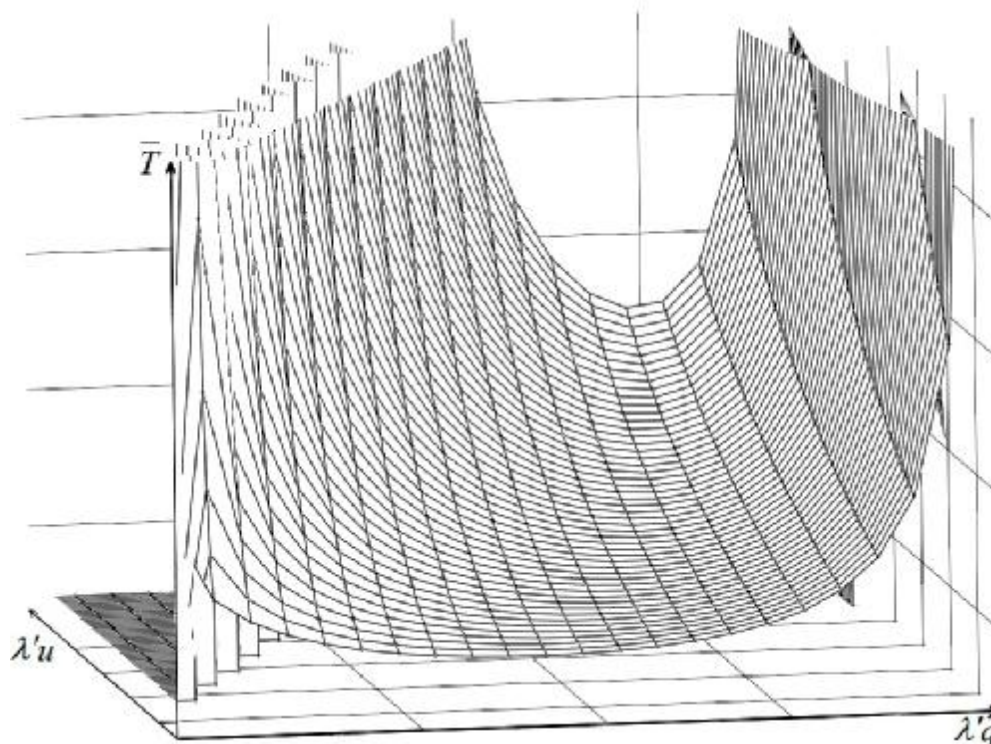


Рисунок 11 – Общий вид зависимости среднего времени отклика РБД от интенсивности обрабатываемых узлом запросов: поисковых и на обновление

Сложность описания процесса отклика РБД на запросы на основе вероятностной модели делает целесообразным сведение модели к детерминированной средней. Для этого случайные величины заменяются их математическими ожиданиями [39].

Необходимо разработать математическую модель отклика РБД на запросы при репликации, базирующуюся на модели двухуровневой информационной системы с репликацией данных, отличающуюся учетом совокупности параметров: интенсивности запросов на обновление ($\lambda'u$) и интенсивности поисковых запросов ($\lambda'q$), обрабатываемых на резервных серверах, на уровне физической интерпретации. Полученная модель позволит разработать алгоритмы для решения задач по повышению оперативности отклика РБД на запросы при репликации.

В настоящем исследовании на основании обследования предметной области, приведенного в первой главе, за основу взята математическая модель двухуровневой информационной системы с репликацией данных,

предложенная Л. И. Мейкшан [60]. Модель доработана с учетом особенностей рассматриваемых РБД, функционирующих в условиях воздействия случайных факторов: снято ограничение на временные задержки при передаче информации по ТКС, обрабатываемой в рамках запроса, введены дополнительные управляющие параметры, уточнены законы поступления заявок [42, 84].

2.2 Модель отклика РБД на запросы при репликации

2.2.1 Обоснование выбора схемы владения данными

При построении модели необходимо учитывать особенности схемы владения данными, определяющей какому из узлов РБД предоставлена привилегия обновления данных. Выделяют три типа схем владения [76]:

1. Схема владения "обновление любой копии". Создается одноранговая среда, в которой множество узлов РБД имеют одинаковые права на обновление копируемых данных. Локальные узлы работают автономно, когда другие узлы недоступны.

2. Схема владения "рабочий поток". Право на обновление копируемых данных передается от одного узла РБД другому. При этом в каждый конкретный момент времени существует только один узел, имеющий право обновлять данные.

3. Схема владения "ведущий/ведомый". В распределенной системе выделяется ведущий узел, на котором разрешается обновление данных. Другие узлы определяют как ведомые. Ведомые узлы получают копии данных от ведущего узла без самостоятельной возможности вносить изменения в данные.

Графики общего вида зависимости среднего времени обработки запросов в РБД от числа узлов для различных схем владения данными представлены на рисунке 12 [48].

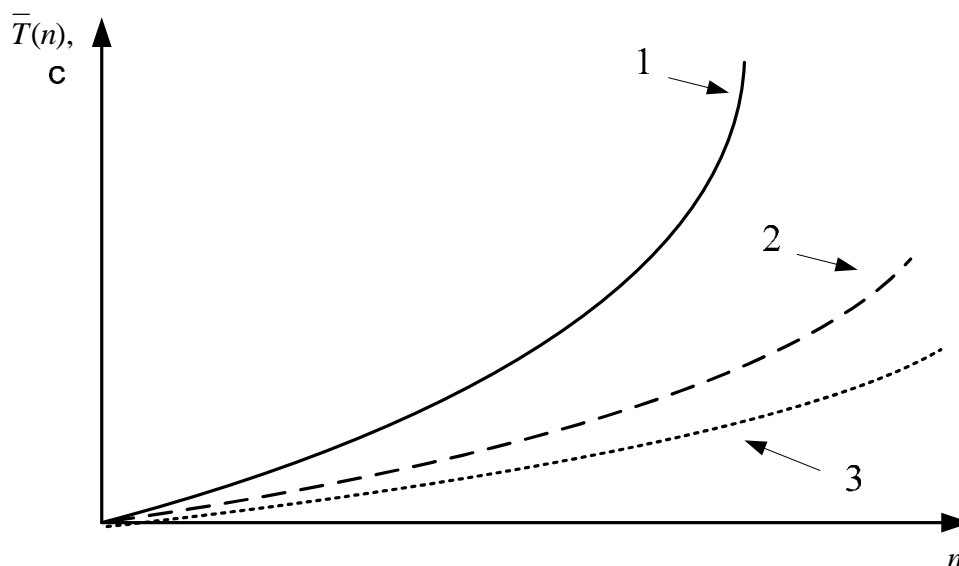


Рисунок 12 – Общий вид зависимости среднего времени обработки запросов в РБД от числа узлов для схемы владения данными: 1 – "обновление любой копии", 2 – "рабочий поток", 3 – "ведущий / ведомый"

Качественные характеристики сравнительных признаков различных схем владения данными РБД представлены в таблице 2.

Таблица 2 – Качественные характеристики различных схем владения данными в РБД [48]

Признаки	Схема владения данными		
	"Ведущий / ведомый"	"Рабочий поток"	"Обновление любой копии"
Способность к масштабированию	высокая	средняя	низкая
Сложность реализации	низкая	средняя	высокая
Наличие однотипного оборудования	необязательно	обязательно	обязательно
Требовательность к ресурсам	низкая	средняя	высокая
Способность к восстановлению	низкая	средняя	высокая

Реализация РБД со схемой владения данными "обновление любой копии" с увеличением количества узлов порождает экспоненциальный рост среднего времени обработки запросов за счет возрастания числа блокировок при обновлении данных [74, 48, 76].

РБД со схемой владения данными "рабочий поток" масштабируется лучше, чем со схемой "обновление любой копии", но требует наличия однотипного оборудования, способного к смене ролей в процессе функционирования.

Единственно допустимой, применительно к ГПК по добыче угля, является схема владения данными "ведущий/ведомый". Это объясняется наличием большого количества разнородного оборудования от различных служб ГПК, в рамках которого реализовать смену ролей в ходе эксплуатации РБД технически сложно. С другой стороны, существует необходимость иметь высокую способность к масштабированию, так как разработка новых участков ГПК требует постоянного расширения РБД. При этом низкую способность к восстановлению при сбоях в системе необходимо компенсировать дополнительным резервированием главного сервера РБД, расположенного в наиболее благоприятном участке, с точки зрения влияния случайных факторов на стабильность его работы.

Обобщенная структура РБД со схемой владения данными – "ведущий/ведомый" представлена на рисунке 13.

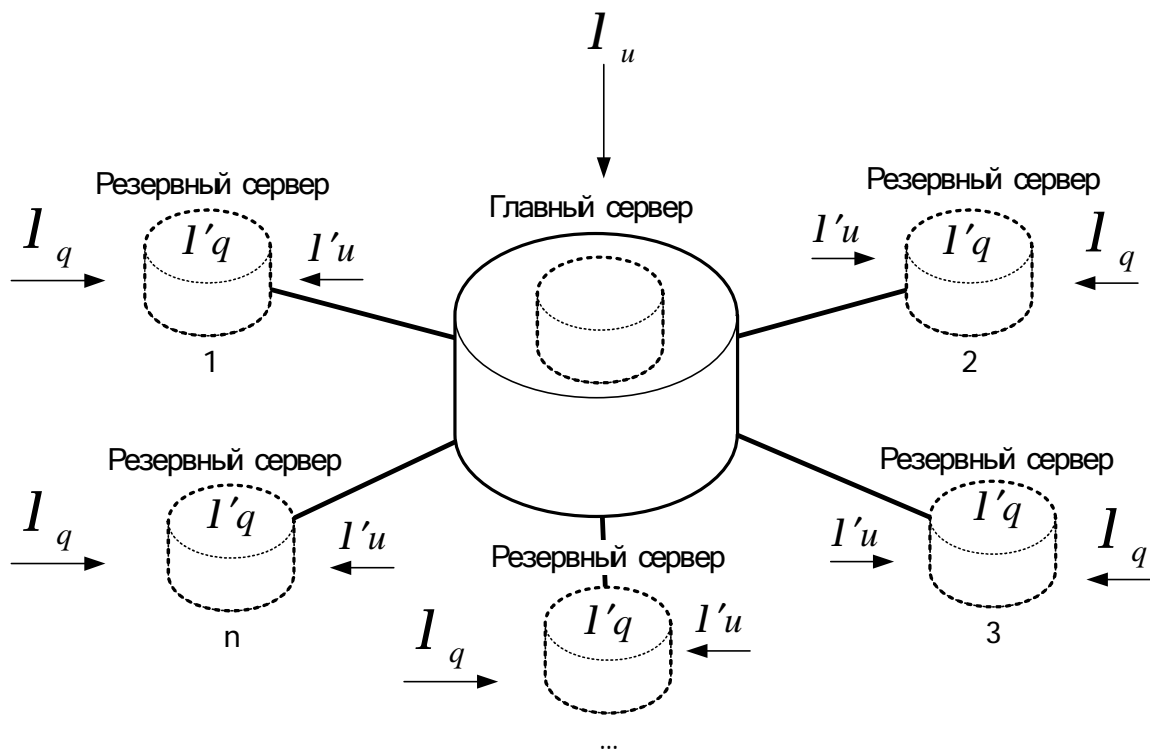


Рисунок 13 – Обобщенная структура РБД со схемой владения данными – "ведущий/ведомый"

2.2.2 Проверка гипотезы о согласовании законов распределения потоков заявок с распределением Пуассона

Главный сервер содержит основной экземпляр базы данных и получает все ее обновления, при этом запросы на обновление образуют поток заявок со средней интенсивностью I_u . Резервные узлы содержат локальные БД, где хранится копия некоторой доли БД главного сервера, при этом I'_u запросов на обновление обрабатывается резервными узлами. С другой стороны, на резервные серверы поступает поток поисковых запросов с интенсивностью I_q , часть из которых – I'_q обрабатывается локально.

Определение закона распределения времени между поступлением заявок в узлы РБД является важной задачей для построения адекватной модели её отклика на запросы при репликации. В исследованиях А. Ю. Иванова высказывается предположение о том, что закон распределения времени между поступлением запросов в узлы РБД согласуется с законом Пуассона [41].

Для проверки согласованности законов распределения статистических данных и гипотетического распределения Пуассона необходимо проверить гипотезу H_0 – закон распределения поступления запросов в узлах РБД предприятия ГПК "ШахтИнвестКузбасс" согласуется с законом Пуассона. В ходе обследования предметной области была получена статистика поступления поисковых запросов на резервный узел в случайные моменты времени для 1000 наблюдений.

Проверка гипотезы о виде распределения случайной величины проводится на основе различных критериев согласия: χ^2 Пирсона, Колмогорова, Смирнова и др. Наиболее универсальным считается критерий χ^2 Пирсона [29], так как он не зависит от вида распределения случайной величины и ее размерности.

Для проверки гипотезы необходимо построить вариационный ряд. Для этого интервал измерения наблюдаемых значений случайной величины

разбивают на N непересекающихся частных интервалов, где N выбирается в соответствии с условием:

$$N \geq [1 + 3,22 \cdot \lg n] + 1, \quad (10)$$

где n – объем выборки, а квадратные скобки обозначают целую часть числа.

В соответствии с гипотезой H_0 :

$$N \geq [1 + 3,22 \cdot \lg 1000] + 1 \Rightarrow N \geq [1 + 3,22 \cdot \lg 1000] + 1 \Rightarrow N \geq 11. \quad (11)$$

Таким образом, для корректных статистических выводов интервал изменения наблюдаемых значений случайной величины необходимо разбить на 11 непересекающихся частичных интервалов.

Группированный вариационный ряд для выборки из генеральной совокупности значений случайной величины (числа поступления заявок в интервалы наблюдения) представлен в таблице 3:

Таблица 3. Группированный вариационный ряд

Интервал наблюдения x_i	n_i – частота поступления заявок в i -й интервал	Относительная частота \bar{n}_i	Накопленная относительная частота w_i
0	9	0,009	0,009
1	28	0,028	0,037
2	81	0,081	0,118
3	146	0,146	0,264
4	207	0,207	0,471
5	218	0,218	0,689
6	163	0,163	0,852
7	114	0,114	0,966
8	18	0,018	0,984
9	13	0,013	0,997
10	3	0,003	1

Гистограмма частот статистических данных представлена на рисунке 14.

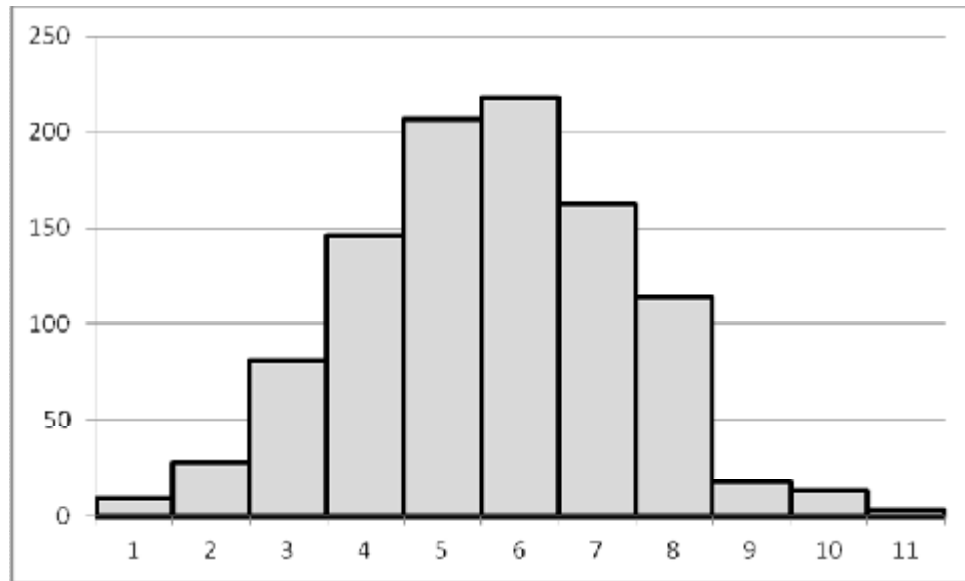


Рисунок 14 – Гистограмма частот статистических данных

Следующим этапом проверки согласованности распределений по критерию χ^2 Пирсона является определение статистического среднего:

$$\tilde{m}_x = \sum_{i=0}^{10} x_i \cdot \bar{n}_i = 4,613. \quad (12)$$

Для вычисления гипотетических вероятностей p_i попадания заявок в i -й интервал наблюдения в соответствии с законом Пуассона используется формула:

$$p_i = \frac{a^i}{i!} \cdot e^{-a}, \quad (13)$$

где $a = \tilde{m}_t = 4,613$.

Вычисленные значения вероятностей p_i попадания заявок в i -й интервал наблюдения представлены в таблице 4:

Таблица 4 – Гипотетическое распределение случайной величины

x_i	0	1	2	3	4	5
p_i	0,045770221	0,105569015	0,162329955	0,187207021	0,172717197	0,132790739
x_i	6	7	8	9	10	
p_i	0,087509097	0,050459933	0,025863519	0,011930841	0,005003361	

Гистограмма частот гипотетических данных представлена на рисунке 15.

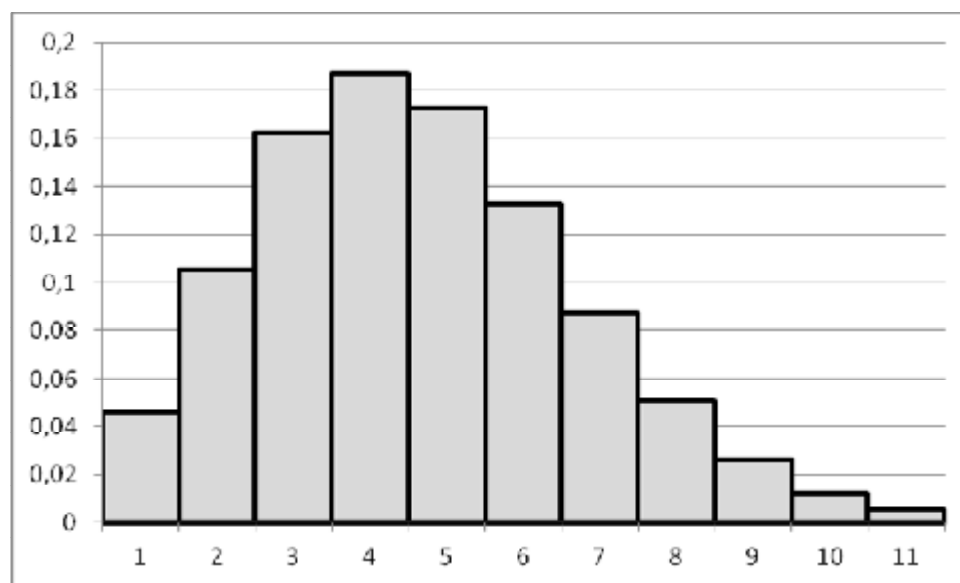


Рисунок 15 – Гистограмма относительных частот гипотетических данных

Значение выборочного критерия χ^2 Пирсона для 11 интервалов наблюдения вычисляется по формуле:

$$\chi^2_{\text{выб}} = \sum_{i=0}^{10} \frac{(n_i - n \cdot p_i)^2}{n \cdot p_i}, \quad (14)$$

где n – объем выборки.

Распределение χ^2 Пирсона зависит от числа степеней свободы r , которое определяется, как число значений случайной величины ($k = 11$)

минус единица (первое условие $\sum_{i=0}^{10} n_i = 1$) и минус еще единица – совпадение статистического и гипотетического математических ожиданий случайной величины поступления заявок в наблюдаемые временные интервалы: $r = 11 - 1 - 1 = 9$.

По таблице значений $c_{\text{выб}}^2$ в зависимости от r и p определяется значение p для $r = 9$ и $c_{\text{выб}}^2 = 3,959909$. Полученное значение вероятности p лежит в диапазоне $(0,95; 0,9)$ [19]. Следовательно, закон распределения времени между поступлением запросов в узлы РБД предприятия ГПК "ШахтИинвестКузбасс" соответствует закону Пуассона, а несущественные расхождения можно отнести к случайным выбросам.

2.2.3 Общий вид модели

Определить вид закона распределения случайной величины времени обслуживания заявок в резервных узлах не представляется возможным, ввиду наличия большого числа независимых случайных факторов, влияющих на время обслуживания заявок. К таким факторам относят: сложность запросов; количество связей РБД, задействованных в обработке запроса; типы и количество операций в запросе; размеры обрабатываемых данных. Поэтому в качестве закона распределения целесообразно использовать – произвольное распределение.

Поступление запросов в соответствии с распределением Пуассона, а их обслуживание по произвольному закону позволяет моделировать отдельные элементы рассматриваемой РБД с помощью одноканальных СМО типа 1/M/G/FCFS [52, 46]. Таким образом, на основе теории массового обслуживания модель отклика РБД на запросы при репликации можно представить в виде совокупности одноканальных СМО типа 1/M/G/FCFS, описывающих обработку запросов на главном и резервном серверах, а также передачу данных по ТКС от главного сервера до резервных и обратно, рисунок 16 [65, 87].

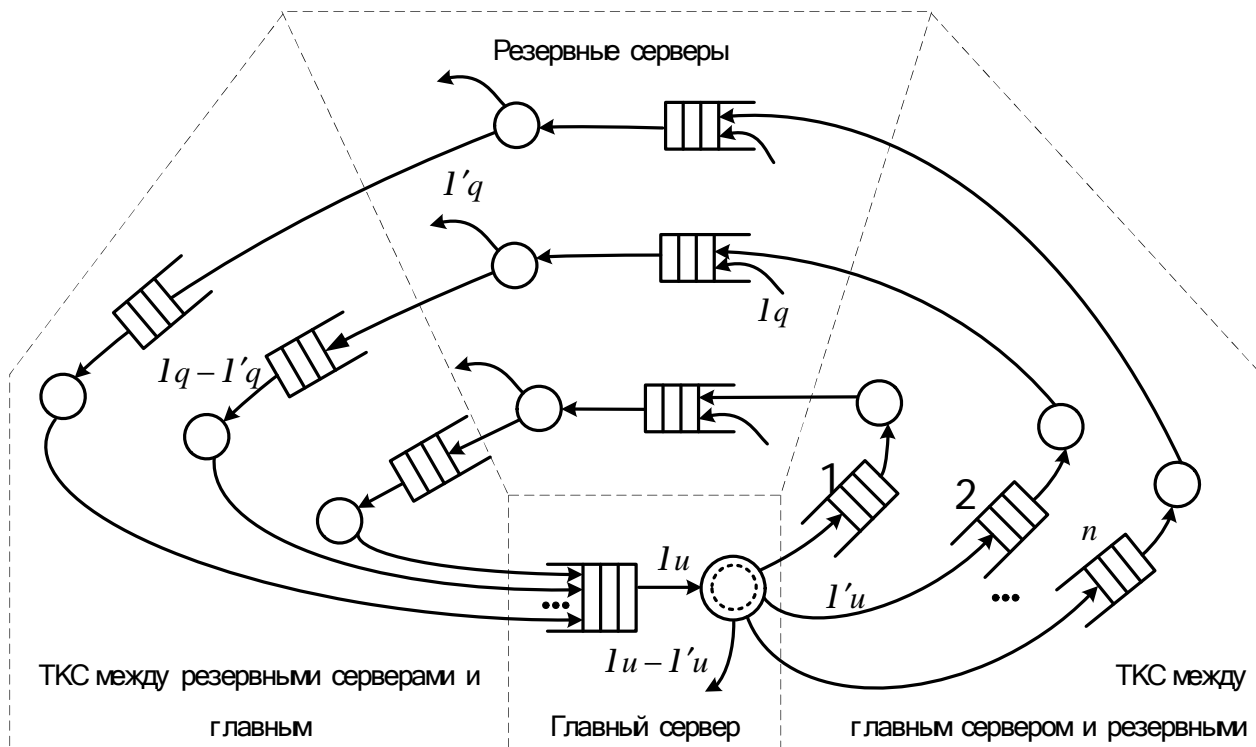


Рисунок 16 – Структура РБД при репликации

Такие модели относятся к классу непрерывно-стохастических (Q-схемы). Так как элементы рассматриваемой РБД соединены последовательно (рисунок 16), то имеет место многофазовая Q-схема с одноканальным оператором сопряжения [83]:

$$Q = \langle W, U, H, Z, R, AL \rangle, \quad (15)$$

где W – подмножество входящих потоков, U – подмножество потоков обслуживания, H – подмножество собственных параметров, Z – подмножество состояний системы, R – оператор сопряжения, AL – оператор алгоритма обслуживания заявок.

Применительно к процессу отклика РБД на запросы при репликации, в качестве подмножества входящих потоков W выступают потоки запросов на обновление и поисковые.

Подмножество потоков обслуживания U определяется задержками при обработке запросов во всех фазах обслуживания: на главном сервере; на резервном; при передаче от главного сервера до резервного и обратно.

Подмножество собственных параметров H характеризует емкость накопителей в отдельных элементах Q-схемы. В работе принято допущение о бесконечной емкости всех накопителей. Корректность такого допущения обуславливается областью применения модели в рамках функционирования РБД без перехода в режим блокировок обслуживания.

Подмножество состояний системы Z определяется состояниями накопителей, каналов и обслуживающих приборов.

Оператор сопряжения R характеризует структуру Q-схемы. Применительно к процессу отклика РБД на запросы при репликации оператор сопряжения задает одноканальное многофазовое обслуживание в разомкнутой системе.

Оператор алгоритма обслуживания заявок AL задает беспriorитетное обслуживание без прерываний и блокировок в порядке FCFS.

На основе модели, предложенной Л. И. Мейкшан [60], в соответствии с особенностями функционирования РБД в условиях воздействия случайных факторов получено, что среднее время отклика РБД на запросы можно определить как [33]:

$$\begin{aligned} \bar{T}(I'q, I'u) = & \frac{I'q}{Iq} \cdot (\bar{Tr}(I'q, I'u) + M[trq]) + \\ & + (1 - \frac{I'q}{Iq}) \cdot (\bar{Tgr}(I'q, I'u) + \bar{Trg}(I'q, I'u) + \bar{Tg}(I'q, I'u) + \\ & + M[tgq] + M[tgrqotkl] + M[trgq]), \end{aligned} \quad (16)$$

где $M[tgq]$ – математическое ожидание (МО) времени обработки поискового запроса на главном сервере;

$M[trq]$ – МО времени обработки поискового запроса на резервном сервере;

$M[trgq]$ – МО времени передачи запроса с резервного сервера на главный;

$M[tgrqotkl]$ – МО времени передачи отклика на запрос с главного сервера на резервный;

Iu – общая интенсивность запросов на обновление;

Iq – общая интенсивность поисковых запросов;

$I'u$ – интенсивность запросов на обновление, обрабатываемых на резервном узле;

$I'q$ – интенсивность поисковых запросов, обрабатываемых на резервном узле;

$\bar{Tr}(I'q, I'u)$ – среднее время ожидания обслуживания запроса на резервном сервере;

$\bar{Tg}(I'q, I'u)$ – среднее время ожидания обслуживания запроса на главном сервере;

$\bar{Trg}(I'q, I'u)$ – среднее время ожидания обслуживания запроса при передаче с резервного сервера на главный;

$\bar{Tgr}(I'q, I'u)$ – среднее время ожидания обслуживания запроса при передаче с главного сервера на резервный.

Первое слагаемое выражения (16) соответствует случаю обработки запроса по локальному циклу (сплошные прямоугольники на рисунке 17). Второе слагаемое выражения (16) определяет обработку запроса на главном сервере по удаленному циклу (пунктирные прямоугольники на рисунке 17).

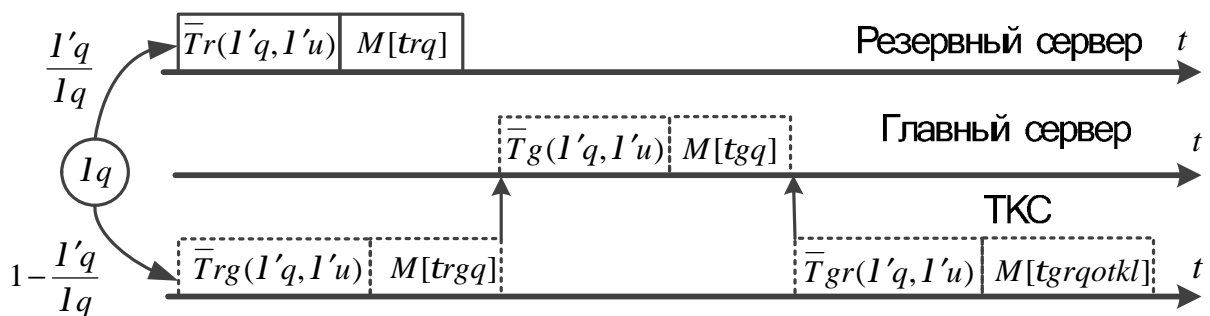


Рисунок 17 – Диаграмма обслуживания запросов по удаленному и локальному циклам

2.2.4 Модель обработки запросов на резервном сервере

На резервный сервер поступает два типа запросов:

- поисковые запросы с интенсивностью $I'q$ и МО времени обработки $M[trq]$;
- запросы на обновление БД резервного сервера с интенсивностью $I'u$ и МО времени обработки $M[tru]$.

По формуле Поллачека-Хинчина для СМО типа 1/M/G/FCFS среднее время ожидания обслуживания определяется как [86, 13, 91]:

$$\bar{Tr}(I'q, I'u) = \frac{I'q \cdot M^2[trq] + I'u \cdot M^2[tru]}{1 - (I'q \cdot M[trq] + I'u \cdot M[tru])}, \quad (17)$$

где

$$0 < I'q \cdot M[trq] + I'u \cdot M[tru] < 1. \quad (18)$$

2.2.5 Модель обработки запросов на главном сервере

Главный сервер обрабатывает три типа запросов:

- запросы на обновление БД с интенсивностью Iu и МО времени обработки $M[tgu]$;
- требования на отправку корректирующих сообщений для обновления БД на резервных узлах с интенсивностью $I'u$ и МО времени обработки $n \cdot M[trgu]$;
- поисковые запросы от n резервных серверов с интенсивностью $n \cdot (Iq - I'q)$ и МО времени обработки $M[tgq]$.

Применение формулы Поллачека-Хинчина дает следующий результат [86, 13, 91]:

$$\bar{T}g(I'q, I'u) = \frac{Iu \cdot M^2[tgu] + I'u \cdot (n \cdot M[trgu])^2 + n \cdot (Iq - I'q) \cdot M^2[tgq]}{1 - (Iu \cdot M[tgu] + I'u \cdot n \cdot M[trgu] + n \cdot (Iq - I'q) \cdot M[tgq])}, \quad (19)$$

где

$$0 < Iu \cdot M[tgu] + I'u \cdot n \cdot M[trgu] + n \cdot (Iq - I'q) \cdot M[tgq] < 1. \quad (20)$$

2.2.6 Модель обработки запросов на участке сети от главного сервера до резервного

СМО, моделирующая работу коммуникационного оборудования между главным и резервным сервером РБД, позволяет оценить среднее время ожидания передачи отклика РБД на участке сети от главного сервера до резервного. В такой СМО существует два типа заявок:

- требования на передачу обновления с главного сервера на резервный с интенсивностью $I'u$ и МО времени обработки $M[tgru]$;
- требования на передачу отклика на запрос с главного сервера на резервный с интенсивностью $Iq - I'q$ и МО времени обработки $M[tgrqotkl]$;

В соответствии с формулой Поллачека-Хинчина [86, 13, 91]:

$$\bar{T}gr(I'q, I'u) = \frac{I'u \cdot M^2[tgru] + (Iq - I'q) \cdot M^2[tgrqotkl]}{1 - (I'u \cdot M[tgru] + (Iq - I'q) \cdot M[tgrqotkl])}, \quad (21)$$

где

$$0 < I'u \cdot M[tgru] + (Iq - I'q) \cdot M[tgrqotkl] < 1. \quad (22)$$

2.2.7 Модель обработки запросов на участке сети от резервного сервера до главного

СМО, моделирующая работу коммуникационного оборудования между резервным и главным сервером РБД, позволяет оценить среднее время ожидания передачи запроса к РБД на участке сети от резервного сервера до главного.

В такой СМО существует два вида заявок:

- требования на передачу запросов на главный сервер с интенсивностью $I_q - I'_q$ и МО времени обработки $M[trgq]$;
- требования на передачу обновления БД с резервного сервера на главный с интенсивностью $\frac{I'_u}{n}$ и МО времени обработки $M[trgu]$.

В соответствии с формулой Поллачека-Хинчина [86, 13, 91]:

$$\bar{T}_{rg}(I'_q, I'_u) = \frac{(I_q - I'_q) \cdot M^2[trgq] + \frac{I'_u}{n} \cdot M^2[trgu]}{1 - ((I_q - I'_q) \cdot M[trgq] + \frac{I'_u}{n} \cdot M[trgu])}, \quad (23)$$

где

$$0 < (I_q - I'_q) \cdot M[trgq] + \frac{I'_u}{n} \cdot M[trgu] < 1. \quad (24)$$

2.3 Проверка адекватности модели отклика РБД на запросы при репликации

Под проверкой адекватности модели подразумевают оценку согласованности результатов эксперимента с теоретическими следствиями в пределах заданной точности [17, 28].

Для проверки адекватности модели отклика РБД на запросы при репликации необходимо сравнить наборы модельного и полученного в условиях производства среднего времени обработки запросов для РБД предприятия ГПК "ШахтИнвестКузбасс", таблица 5.

Таблица 5 – Модельные и полученные в условиях производства данные

№ эксперимента	$I'q, \text{мин}^{-1}$	$I'u, \text{мин}^{-1}$	Модельное время $\bar{T}(I'q, I'u)$	Полученное в условиях производства время \bar{T}^*	$\Delta\bar{T}, \%$	$\overline{\Delta\bar{T}}, \%$
1	1	1	0,381	0,391	2,6	2,96
2	50	5	0,324	0,329	1,5	
3	100	10	0,277	0,272	1,8	
4	150	15	0,249	0,258	3,5	
5	200	20	0,221	0,229	3,5	
6	250	25	0,206	0,214	3,7	
7	300	30	0,197	0,205	3,9	
8	350	35	0,219	0,213	2,7	
9	400	40	0,274	0,265	3,3	
10	450	45	0,512	0,496	3,1	

Модельные данные таблицы 5 получены на основе программы "ctrReplic", зарегистрированной в Федеральной службе по интеллектуальной собственности (свидетельство о государственной регистрации программы для ЭВМ № 2013618670 от 13 сентября 2013 года) [100].

Исходные данные для экспериментов выбраны таким образом, чтобы максимально охватить диапазон допустимых значений управляющих параметров за ограниченное число экспериментов.

Графически расхождение полученного в условиях производства и модельного среднего времени отклика РБД на запросы при увеличении значений управляющих параметров представлено на рисунке 18.

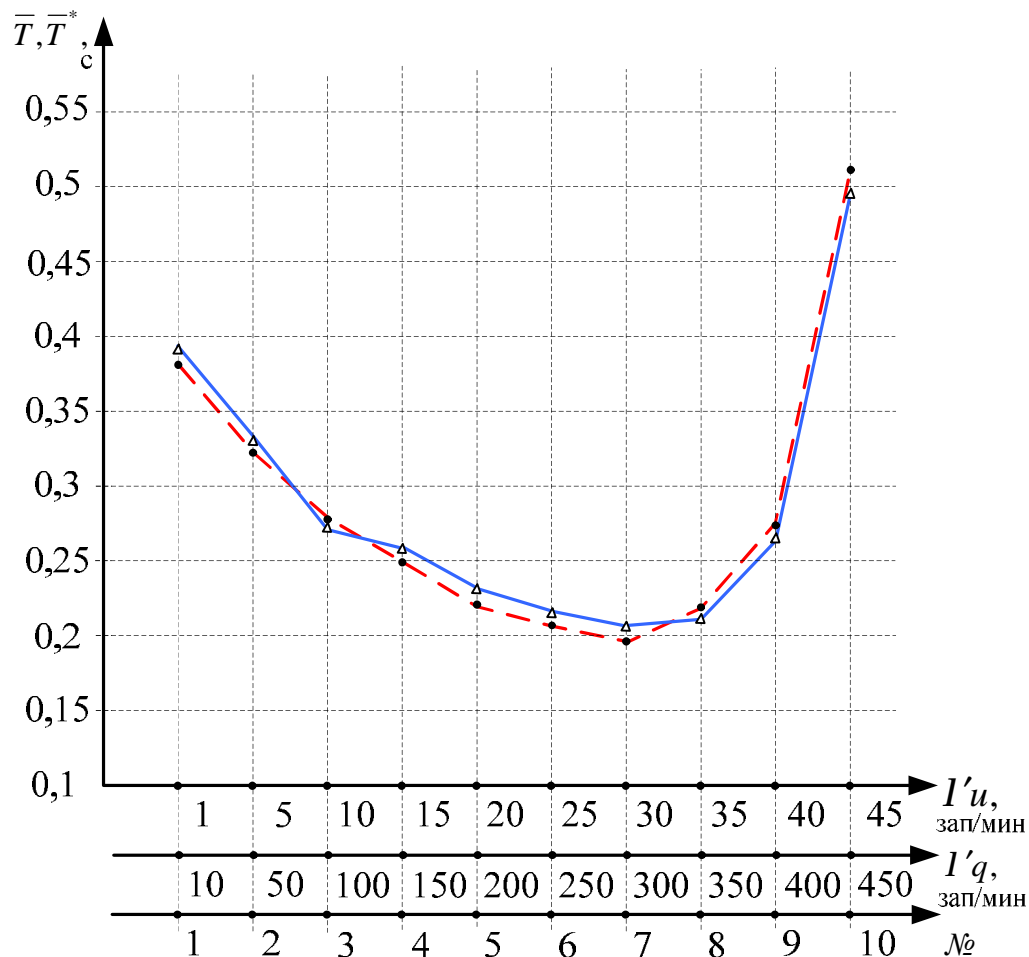


Рисунок 18 – Графики полученного в условиях производства (сплошная линия) и модельного (пунктирная линия) среднего времени отклика РБД на запросы при увлечении значений управляющих параметров

Анализ полученных результатов показывает, что среднее расхождение модельного и полученного в условиях производства среднего времени отклика РБД на запросы для 10 экспериментов составляет 2,96% и не превышает 5% по всему диапазону исходных данных экспериментов. Это позволяет сформулировать вывод о допустимой адекватности представленной модели отклика РБД на запросы при репликации [51, 31, 58].

2.4 Проверка чувствительности модели отклика РБД на запросы при репликации

Чувствительность модели – восприимчивость к изменению существенных для исследования параметров [12].

В качестве существенных для исследования параметров в диссертационной работе рассматриваются $I'u$ – интенсивность запросов на обновление, обрабатываемых на резервном сервере, и $I'q$ – интенсивность поисковых запросов, обрабатываемых на резервном сервере.

Оценку чувствительности проводят по каждому параметру в отдельности на основании приращений наблюдаемой переменной. Так как вид зависимости среднего времени отклика РБД на запросы при репликации в соответствии с данными, полученными в условиях производства на предприятии ГПК "ШахтИнвестКузбасс", (таблица 5) – параболический, то оценку чувствительности модели следует проводить по значениям параметров, соответствующим максимальному разбросу значений среднего времени отклика на запросы.

Оценка чувствительности модели проводится в три этапа [28, 78]:

1. Определяется величина относительного среднего приращения каждого параметра на основе выражений:

$$\Delta I'q = \frac{(I'q \max - I'q \min) \cdot 2}{I'q \max + I'q \min} \cdot 100\%, \quad (25)$$

где $I'q \max$ – значение интенсивности поисковых запросов, обрабатываемых на резервном узле, при котором среднее время отклика РБД на запросы максимально в соответствии с данными, полученными в условиях производства, таблица 5, а $I'q \min$ – значение интенсивности поисковых запросов, обрабатываемых на резервном узле, при котором среднее время отклика РБД на запросы минимально в соответствии с данными, полученными в условиях производства, таблица 5;

$$\Delta I'u = \frac{(I'u \max - I'u \min) \cdot 2}{I'u \max + I'u \min} \cdot 100\%, \quad (26)$$

где $I'u \max$ – значение интенсивности запросов на обновление, обрабатываемых на резервном узле, при котором среднее время отклика РБД на запросы максимально в соответствии с данными, полученными в условиях производства, таблица 5, а $I'u \min$ – значение интенсивности запросов на обновление, обрабатываемых на резервном узле, при котором среднее время отклика РБД на запросы минимально в соответствии с данными, полученными в условиях производства, таблица 5.

2. Определяются модельные значения $\bar{T}(I'q \max, I'u)$ и $\bar{T}(I'q \min, I'u)$ при фиксированном значении $I'u$, а также модельные значения $\bar{T}(I'q, I'u \max)$ и $\bar{T}(I'q, I'u \min)$ при фиксированном значении $I'q$.

3. Вычисляются относительные приращения среднего времени отклика РБД на запросы для каждого оцениваемого параметра:

$$\Delta \bar{T}q = \frac{(\bar{T}(I'q \max, I'u) - \bar{T}(I'q \min, I'u)) \cdot 2}{\bar{T}(I'q \max, I'u) + \bar{T}(I'q \min, I'u)} \cdot 100\%, \quad (27)$$

$$\Delta \bar{T}u = \frac{(\bar{T}(I'q, I'u \max) - \bar{T}(I'q, I'u \min)) \cdot 2}{\bar{T}(I'q, I'u \max) + \bar{T}(I'q, I'u \min)} \cdot 100\%. \quad (28)$$

Полученные пары $(\Delta \bar{T}q, \Delta I'q)$ и $(\Delta \bar{T}u, \Delta I'u)$ характеризуют чувствительность модели по каждому отдельному параметру.

Применительно к РБД предприятия ГПК "ШахтИнвестКузбасс":

$$\Delta I'q = \frac{(450 - 300) \cdot 2}{450 + 300} \cdot 100\% = 40\%, \quad (29)$$

$$\Delta I'u = \frac{(45 - 30) \cdot 2}{45 + 30} \cdot 100\% = 40\%. \quad (30)$$

На основе представленной модели получены модельные значения $\bar{T}(I'q_{\max}, I'u)$ и $\bar{T}(I'q_{\min}, I'u)$ при фиксированном значении $I'u = 30 \text{ мин}^{-1}$, а также модельные значения $\bar{T}(I'q, I'u_{\max})$ и $\bar{T}(I'q, I'u_{\min})$ при фиксированном значении $I'q = 300 \text{ мин}^{-1}$. Результаты вычислений представлены в таблице 6

Таблица 6 – Расчетные данные для оценки чувствительности

$I'q, \text{ мин}^{-1}$	$I'u, \text{ мин}^{-1}$	Расчитанное время $\bar{T}(I'q, I'u)$
300	30	0,197
300	45	0,282
450	30	0,477

В соответствии с формулой (27) и данными таблицы 6 относительное приращение среднего времени отклика РБД на запросы для параметра $I'q$ определяется как:

$$\Delta \bar{T}_q = \frac{(0,477 - 0,197) \cdot 2}{0,477 + 0,197} \cdot 100\% \approx 83\% \quad (31)$$

В соответствии с формулой (28) и данными таблицы 6 относительное приращение среднего времени отклика РБД на запросы для параметра $I'u$ определяется как:

$$\Delta \bar{T}_u = \frac{(0,282 - 0,197) \cdot 2}{0,282 + 0,197} \cdot 100\% \approx 35\% \quad (32)$$

Пары значений приращений оцениваемых параметров (40%, 83%) и (40%, 35%) позволяют сформулировать вывод о достаточной чувствительности модели к изменению параметров $I'u$ и $I'q$ [28].

Выводы по главе 2.

1. Уменьшение количества реплицируемых данных порождает увеличение числа удаленных заявок на обслуживание. С другой стороны, увеличение количества реплицируемых данных приводит к росту числа заявок на репликацию. При заданных ограничениях на временные задержки обработки запросов в различных элементах РБД неправильный выбор значений интенсивностей обработки запросов на резервных серверах снижает оперативность её отклика на запросы.

Существование данного противоречия обосновывает целесообразность решения задачи выбора фрагментов данных для немедленной репликации с учетом снижения среднего времени отклика РБД на запросы при репликации. Решение данной задачи возможно на основе адекватной математической модели отклика РБД на запросы при репликации.

2. Заявленная модель отклика РБД на запросы при репликации описывает схему владения данными "ведущий/ведомый" в соответствии с особенностями РБД предприятия ГПК "ШахтИнвестКузбасс".

3. Закон распределения времени между поступлением запросов на резервные узлы согласуется с распределением Пуассона, что подтверждается результатами проведенной оценки на основе критерия согласия χ^2 Пирсона. При этом закон распределения случайной величины времени обслуживания заявок в резервных узлах – произвольный.

Полученные результаты позволяют моделировать работу отдельных элементов РБД с помощью одноканальных СМО типа 1/M/G/FCFS.

4. Результаты сравнения модельного и полученного в условиях производства среднего времени обработки запросов для 10 экспериментов позволяют утверждать о допустимой адекватности представленной модели отклика РБД на запросы при репликации.

5. Оценка чувствительности полученной модели по каждому параметру в отдельности на основании приращений наблюдаемой переменной продемонстрировала достаточную чувствительность модели.

ГЛАВА 3. РАЗРАБОТКА АЛГОРИТМОВ ВЫЧИСЛЕНИЯ ЗНАЧЕНИЙ ПАРАМЕТРОВ РБД ПРИ РЕПЛИКАЦИИ

Одним из критериев эффективности функционирования РБД является минимум среднего времени отклика на запросы. Так как телекоммуникации и узлы РБД предприятия ГПК территориально рассредоточены и, как правило, размещены в агрессивной среде (повышенная влажность, запыленность, сейсмоактивность, низкие и высокие температуры, вибрации), то задачу минимизации времени отклика РБД на запросы необходимо решать применительно к заданным ограничениям на временные задержки обработки запросов в различных её элементах.

3.1 Задача минимизации среднего времени отклика РБД на запросы

В соответствии с моделью отклика РБД на запросы при репликации среднее время отклика РБД на запросы определяется как [60, 32]:

$$\begin{aligned} \bar{T}(I'q, I'u) = & \frac{I'q}{Iq} \cdot (\bar{T}r(I'q, I'u) + M[trq]) + \\ & + (1 - \frac{I'q}{Iq}) \cdot (\bar{T}gr(I'q, I'u) + \bar{T}rg(I'q, I'u) + \bar{T}g(I'q, I'u) + \\ & + M[tgq] + M[tgrqotkl] + M[trgq]), \end{aligned}$$

где:

$$\begin{aligned} \bar{T}r(I'q, I'u) = & \frac{I'q \cdot M^2[trq] + I'u \cdot M^2[tru]}{1 - (I'q \cdot M[trq] + I'u \cdot M[tru])}, \\ \bar{T}g(I'q, I'u) = & \frac{Iu \cdot M^2[tgu] + I'u \cdot (n \cdot M[trgu])^2 + n \cdot (Iq - I'q) \cdot M^2[tgq]}{1 - (Iu \cdot M[tgu] + I'u \cdot n \cdot M[trgu] + n \cdot (Iq - I'q) \cdot M[tgq])}, \\ \bar{T}gr(I'q, I'u) = & \frac{I'u \cdot M^2[tgru] + (Iq - I'q) \cdot M^2[tgrqotkl]}{1 - (I'u \cdot M[tgru] + (Iq - I'q) \cdot M[tgrqotkl])}, \end{aligned}$$

$$\bar{T}rg(I'q, I'u) = \frac{(Iq - I'q) \cdot M^2[trgq] + \frac{I'u}{n} \cdot M^2[trgu]}{1 - ((Iq - I'q) \cdot M[trgq] + \frac{I'u}{n} \cdot M[trgu])},$$

при ограничениях "А":

$$0 < I'q \cdot M[trq] + I'u \cdot M[tru] < 1,$$

$$0 < I'u \cdot M[tgru] + (Iq - I'q) \cdot M[tgrqotkl] < 1,$$

$$0 < Iu \cdot M[tgu] + I'u \cdot n \cdot M[trgu] + n \cdot (Iq - I'q) \cdot M[tgq] < 1,$$

$$0 < (Iq - I'q) \cdot M[trgq] + \frac{I'u}{n} \cdot M[trgu] < 1.$$

Здесь:

$M[tgq]$ – МО времени обработки поискового запроса на главном сервере;

$M[trq]$ – МО времени обработки поискового запроса на резервном сервере;

$M[tgu]$ – МО времени обработки запроса на обновление на главном сервере;

$M[tru]$ – МО времени обработки запроса на обновление на резервном сервере;

$M[tgru]$ – МО времени передачи обновления с главного сервера на резервный;

$M[trgu]$ – МО времени, требуемого главному серверу на отправку одного сообщения для обновления БД резервного сервера;

$M[trgq]$ – МО времени передачи запроса с резервного сервера на главный;

$M[tgrqotkl]$ – МО времени передачи отклика на запрос с главного сервера на резервный;

Iu – общая интенсивность запросов на обновление;

$I'q$ – общая интенсивность поисковых запросов;

$I'u$ – интенсивность запросов на обновление, обрабатываемых на резервном узле;

$I'q$ – интенсивность поисковых запросов, обрабатываемых на резервном узле;

n – число резервных серверов.

Таким образом, задача исследования сводится к нахождению экстремума целевой функции:

$$\bar{T}(I'q(\{X'\}, type), I'u(\{X'\}, type)) \rightarrow \min_{I'q, I'u} \left| \begin{array}{l} (\bar{T}) \\ "A", \bar{T} \leq T\delta \end{array} \right. \quad (33)$$

при наличии ограничений в виде линейных неравенств – "А".

3.2 Обоснование математического метода решения задачи минимизации среднего времени отклика РБД на запросы

Аналитически задачи данного вида решаются на основе обобщенного метода множителей Лагранжа. Суть метода состоит в том, что если точка безусловного оптимума функции $\bar{T}(I'q, I'u)$ не удовлетворяет всем ограничениям задачи, то оптимальное решение задачи с ограничениями достигается в граничной точке области допустимых решений [86].

Вычислительная процедура аналитического метода включает следующие шаги [86]:

1. Решается задача без учета ограничений – "А":

$$\bar{T}(I'q, I'u) \rightarrow \min(\bar{T}). \quad (34)$$

Для нахождения точек безусловного минимума функции $\bar{T}(I'q, I'u)$ необходимо решить систему уравнений с частными производными:

$$\begin{cases} \frac{\partial T(I'q, I'u)}{\partial I'q} = 0 \\ \frac{\partial T(I'q, I'u)}{\partial I'u} = 0. \end{cases} \quad (35)$$

Согласно правилам дифференцирования функция вида $y = \frac{u}{v}$ имеет производную [69]:

$$y' = \frac{u'v - uv'}{v^2}. \quad (36)$$

Частные производные элементов функции $\bar{T}(I'q, I'u)$ имеют вид:

$$\begin{aligned} \frac{\partial \bar{T}(I'q, I'u)}{\partial I'q} &= \frac{Iq \cdot M^2[trq]}{1 - (I'q \cdot M[trq] + I'u \cdot M[tru])} - \\ &- \frac{Iq \cdot M[trq] \cdot (I'q \cdot M^2[trq] + I'u \cdot M^2[tru])}{(1 - (I'q \cdot M[trq] + I'u \cdot M[tru]))^2} = \\ &= \frac{I'q \cdot Iq \cdot M^3[trq] + I'u \cdot Iq \cdot M^2[tru] \cdot M[trq]}{(I'q \cdot M[trq] + I'u \cdot M[tru] - 1)^2} - \frac{Iq \cdot M^2[trq]}{I'q \cdot M[trq] + I'u \cdot M[tru] - 1}, \end{aligned} \quad (37)$$

$$\begin{aligned} \frac{\partial \bar{T}(I'q, I'u)}{\partial I'u} &= \frac{Iu \cdot M^2[tru]}{1 - (I'q \cdot M[trq] + I'u \cdot M[tru])} - \\ &- \frac{Iu \cdot M[tru] \cdot (I'q \cdot M^2[trq] + I'u \cdot M^2[tru])}{(1 - (I'q \cdot M[trq] + I'u \cdot M[tru]))^2} = \\ &= \frac{I'u \cdot Iu \cdot M^3[tru] + I'q \cdot Iu \cdot M^2[trq] \cdot M[tru]}{(I'q \cdot M[trq] + I'u \cdot M[tru] - 1)^2} - \frac{Iu \cdot M^2[tru]}{I'q \cdot M[trq] + I'u \cdot M[tru] - 1}, \end{aligned} \quad (38)$$

$$\begin{aligned}
\frac{\partial \overline{Tg}(I'q, I'u)}{\partial I'q} &= \frac{-n \cdot Iq \cdot M^2[tgq]}{1 - (Iu \cdot M[tgu] + I'u \cdot n \cdot M[trgu] + n \cdot (Iq - I'q) \cdot M[tgq])} - \\
&= \frac{(Iu \cdot M^2[tgu] + I'u \cdot (n \cdot M[trgu])^2 + n \cdot (Iq - I'q) \cdot M^2[tgq]) \cdot (n \cdot Iq \cdot M[tgq])}{(1 - (Iu \cdot M[tgu] + I'u \cdot n \cdot M[trgu] + n \cdot (Iq - I'q) \cdot M[tgq]))^2} = \\
&= \frac{n \cdot Iq \cdot M^2[tgq]}{Iu \cdot M[tgu] + I'u \cdot n \cdot M[trgu] + n \cdot (Iq - I'q) \cdot M[tgq] - 1} \\
&= \frac{n \cdot Iq \cdot Iu \cdot M^2[tgu] \cdot M[tgq] + I'u \cdot Iq \cdot n^3 \cdot M[trgu]^2 \cdot M[tgq]}{(Iu \cdot M[tgu] + I'u \cdot n \cdot M[trgu] + n \cdot (Iq - I'q) \cdot M[tgq] - 1)^2} + \\
&+ \frac{n^2 \cdot (Iq^2 - Iq \cdot I'q) \cdot M^3[tgq]}{(Iu \cdot M[tgu] + I'u \cdot n \cdot M[trgu] + n \cdot (Iq - I'q) \cdot M[tgq] - 1)^2},
\end{aligned} \tag{39}$$

$$\begin{aligned}
\frac{\partial \overline{Tg}(I'q, I'u)}{\partial I'u} &= \frac{Iu \cdot (n \cdot M[trgu])^2}{1 - (Iu \cdot M[tgu] + I'u \cdot n \cdot M[trgu] + n \cdot (Iq - I'q) \cdot M[tgq])} - \\
&= \frac{(Iu \cdot M^2[tgu] + I'u \cdot (n \cdot M[trgu])^2 + n \cdot (Iq - I'q) \cdot M^2[tgq]) \cdot (-Iu \cdot n \cdot M[trgu])}{(1 - (Iu \cdot M[tgu] + I'u \cdot n \cdot M[trgu] + n \cdot (Iq - I'q) \cdot M[tgq]))^2} = \\
&= \frac{Iu^2 \cdot n \cdot M^2[tgu] \cdot M[trgu] + I'u \cdot Iu \cdot (n \cdot M[trgu])^3}{(Iu \cdot M[tgu] + I'u \cdot n \cdot M[trgu] + n \cdot (Iq - I'q) \cdot M[tgq] - 1)^2} + \\
&+ \frac{n^2 \cdot (Iq - I'q) \cdot M[trgu] \cdot M^2[tgq]}{(Iu \cdot M[tgu] + I'u \cdot n \cdot M[trgu] + n \cdot (Iq - I'q) \cdot M[tgq] - 1)^2} - \\
&= \frac{Iu \cdot (n \cdot M[trgu])^2}{Iu \cdot M[tgu] + I'u \cdot n \cdot M[trgu] + n \cdot (Iq - I'q) \cdot M[tgq] - 1},
\end{aligned} \tag{40}$$

$$\begin{aligned}
\frac{\partial \overline{Tgr}(I'q, I'u)}{\partial I'q} &= \frac{-Iq \cdot M^2[tgrqotkl]}{1 - (I'u \cdot M[tgru] + (Iq - I'q) \cdot M[tgrqotkl])} - \\
&= \frac{Iq \cdot M[tgrqotkl] \cdot (I'u \cdot M^2[tgru] + (Iq - I'q) \cdot M^2[tgrqotkl])}{(1 - (I'u \cdot M[tgru] + (Iq - I'q) \cdot M[tgrqotkl]))^2} = \\
&= \frac{Iq \cdot M^2[tgrqotkl]}{I'u \cdot M[tgru] + (Iq - I'q) \cdot M[tgrqotkl] - 1} \\
&= \frac{Iq \cdot M[tgrqotkl] \cdot I'u \cdot M^2[tgru] + (Iq^2 - Iq \cdot I'q) \cdot M^3[tgrqotkl]}{(I'u \cdot M[tgru] + (Iq - I'q) \cdot M[tgrqotkl] - 1)^2},
\end{aligned} \tag{41}$$

$$\begin{aligned}
\frac{\partial \bar{Tgr}(I'q, I'u)}{\partial I'u} &= \frac{Iu \cdot M^2[tgru]}{1 - (I'u \cdot M[tgru] + (Iq - I'q) \cdot M[tgrqotkl])} - \\
&= \frac{-Iu \cdot M[tgru] \cdot (I'u \cdot M^2[tgru] + (Iq - I'q) \cdot M^2[tgrqotkl])}{(1 - (I'u \cdot M[tgru] + (Iq - I'q) \cdot M[tgrqotkl]))^2} = \\
&= \frac{I'u \cdot Iu \cdot M^3[tgru] + (Iq - I'q) \cdot M^2[tgrqotkl] \cdot Iu \cdot M[tgru]}{(I'u \cdot M[tgru] + (Iq - I'q) \cdot M[tgrqotkl] - 1)^2} - \\
&= \frac{Iu \cdot M^2[tgru]}{I'u \cdot M[tgru] + (Iq - I'q) \cdot M[tgrqotkl] - 1},
\end{aligned} \tag{42}$$

$$\begin{aligned}
\frac{\partial \bar{Trg}(I'q, I'u)}{\partial I'q} &= \frac{-Iq \cdot M^2[trgq]}{1 - ((Iq - I'q) \cdot M[trgq] + \frac{I'u}{n} \cdot M[trgu])} - \\
&= \frac{Iq \cdot M[trgq] \cdot ((Iq - I'q) \cdot M^2[trgq] + \frac{I'u}{n} \cdot M^2[trgu])}{(1 - ((Iq - I'q) \cdot M[trgq] + \frac{I'u}{n} \cdot M[trgu]))^2} = \\
&= \frac{Iq \cdot M^2[trgq]}{(Iq - I'q) \cdot M[trgq] + \frac{I'u}{n} \cdot M[trgu] - 1} - \\
&= \frac{(Iq^2 - Iq \cdot I'q) \cdot M^3[trgq] + \frac{I'u}{n} \cdot Iq \cdot M[trgq] \cdot M^2[trgu]}{((Iq - I'q) \cdot M[trgq] + \frac{I'u}{n} \cdot M[trgu] - 1)^2},
\end{aligned} \tag{43}$$

$$\begin{aligned}
\frac{\partial \bar{Trg}(I'q, I'u)}{\partial I'u} &= \frac{\frac{Iu}{n} \cdot M^2[trgu]}{1 - ((Iq - I'q) \cdot M[trgq] + \frac{I'u}{n} \cdot M[trgu])} - \\
&= \frac{-\frac{Iu}{n} \cdot M[trgu] \cdot ((Iq - I'q) \cdot M^2[trgq] + \frac{I'u}{n} \cdot M^2[trgu])}{(1 - ((Iq - I'q) \cdot M[trgq] + \frac{I'u}{n} \cdot M[trgu]))^2} = \\
&= \frac{(Iq - I'q) \cdot \frac{Iu}{n} \cdot M[trgu] \cdot M^2[trgq] + \frac{Iu \cdot I'u}{n^2} \cdot M^3[trgu]}{((Iq - I'q) \cdot M[trgq] + \frac{I'u}{n} \cdot M[trgu] - 1)^2} - \\
&= \frac{\frac{Iu}{n} \cdot M^2[trgu]}{(Iq - I'q) \cdot M[trgq] + \frac{I'u}{n} \cdot M[trgu] - 1}.
\end{aligned} \tag{44}$$

В соответствии с правилами дифференцирования функция вида $y = u \cdot v$ имеет производную [69]:

$$y' = u' \cdot v + u \cdot v'. \quad (45)$$

Тогда, частные производные функции $\bar{T}(I'q, I'u)$ определяются как:

$$\begin{aligned} \frac{\partial \bar{T}(I'q, I'u)}{\partial I'q} = & \bar{T}r(I'q, I'u) + M[trq] + \frac{I'q}{Iq} \cdot \frac{\partial \bar{T}r(I'q, I'u)}{\partial I'q} - \\ & - (\bar{T}gr(I'q, I'u) + \bar{T}rg(I'q, I'u) + \bar{T}g(I'q, I'u) + M[tgq] + M[tgrqotkl] + \\ & + M[trgq]) + (1 - \frac{I'q}{Iq}) \cdot (\frac{\partial \bar{T}gr(I'q, I'u)}{\partial I'q} + \frac{\partial \bar{T}rg(I'q, I'u)}{\partial I'q} + \frac{\partial \bar{T}g(I'q, I'u)}{\partial I'q}), \end{aligned} \quad (46)$$

$$\begin{aligned} \frac{\partial \bar{T}(I'q, I'u)}{\partial I'u} = & \frac{I'q}{Iq} \cdot \frac{\partial \bar{T}r(I'q, I'u)}{\partial I'u} + \\ & + (1 - \frac{I'q}{Iq}) \cdot (\frac{\partial \bar{T}gr(I'q, I'u)}{\partial I'u} + \frac{\partial \bar{T}rg(I'q, I'u)}{\partial I'u} + \frac{\partial \bar{T}g(I'q, I'u)}{\partial I'u}). \end{aligned} \quad (47)$$

Дальнейшее аналитическое решение сводится к нахождению корней системы уравнений:

$$\left\{ \begin{aligned} & r(I'q, I'u) + M[trq] + \frac{I'q}{Iq} \cdot \frac{\partial \bar{T}r(I'q, I'u)}{\partial I'q} - \\ & - (\bar{T}gr(I'q, I'u) + \bar{T}rg(I'q, I'u) + \bar{T}g(I'q, I'u) + M[tgq] + M[tgrqotkl] + \\ & + M[trgq]) + (1 - \frac{I'q}{Iq}) \cdot (\frac{\partial \bar{T}gr(I'q, I'u)}{\partial I'q} + \frac{\partial \bar{T}rg(I'q, I'u)}{\partial I'q} + \frac{\partial \bar{T}g(I'q, I'u)}{\partial I'q}) = 0 \\ & \frac{I'q}{Iq} \cdot \frac{\partial \bar{T}r(I'q, I'u)}{\partial I'u} + \\ & + (1 - \frac{I'q}{Iq}) \cdot (\frac{\partial \bar{T}gr(I'q, I'u)}{\partial I'u} + \frac{\partial \bar{T}rg(I'q, I'u)}{\partial I'u} + \frac{\partial \bar{T}g(I'q, I'u)}{\partial I'u}) = 0. \end{aligned} \right. \quad (48)$$

Однако решение полученной системы уравнений аналитически сопряжено с нахождением корней уравнений шестого порядка. Степень

проявляется при приведении элементов уравнений к общему знаменателю. В связи с этим для решения необходимо использовать численные методы решения нелинейных задач с линейными ограничениями.

В таблице 7 представлены характеристики аналитических и численных методов для решения задачи минимизации среднего времени отклика РБД на запросы [86, 67]:

Таблица 7 – Характеристики методов решения задачи нахождения минимума среднего времени отклика РБД на запросы

Свойства	Методы	
	Аналитические	Численные
Вычислительная сложность	Высокая	низкая
Точность	Максимальная	удовлетворительная
Сложность реализации алгоритмов	Высокая	низкая

Видно, что численные методы, в отличие от аналитических, позволяют решать задачу нахождения минимума среднего времени отклика РБД на запросы с удовлетворительной точностью при условии снижения вычислительной сложности и сложности реализации алгоритмов.

3.3 Алгоритм вычисления оптимальной загруженности резервного узла распределенной базы данных при репликации

В общем виде решение задачи нахождения минимума среднего времени отклика РБД на запросы численными методами состоит из нескольких этапов и может быть реализовано в рамках метода линейных комбинаций, в основе которого лежит градиентный метод наискорейшего спуска [4, 86], модифицированный для применения при наличии линейных ограничений.

Схема алгоритма вычисления оптимальной загруженности резервного узла РБД при репликации, оформленная согласно [21], представлена на рисунке 19.

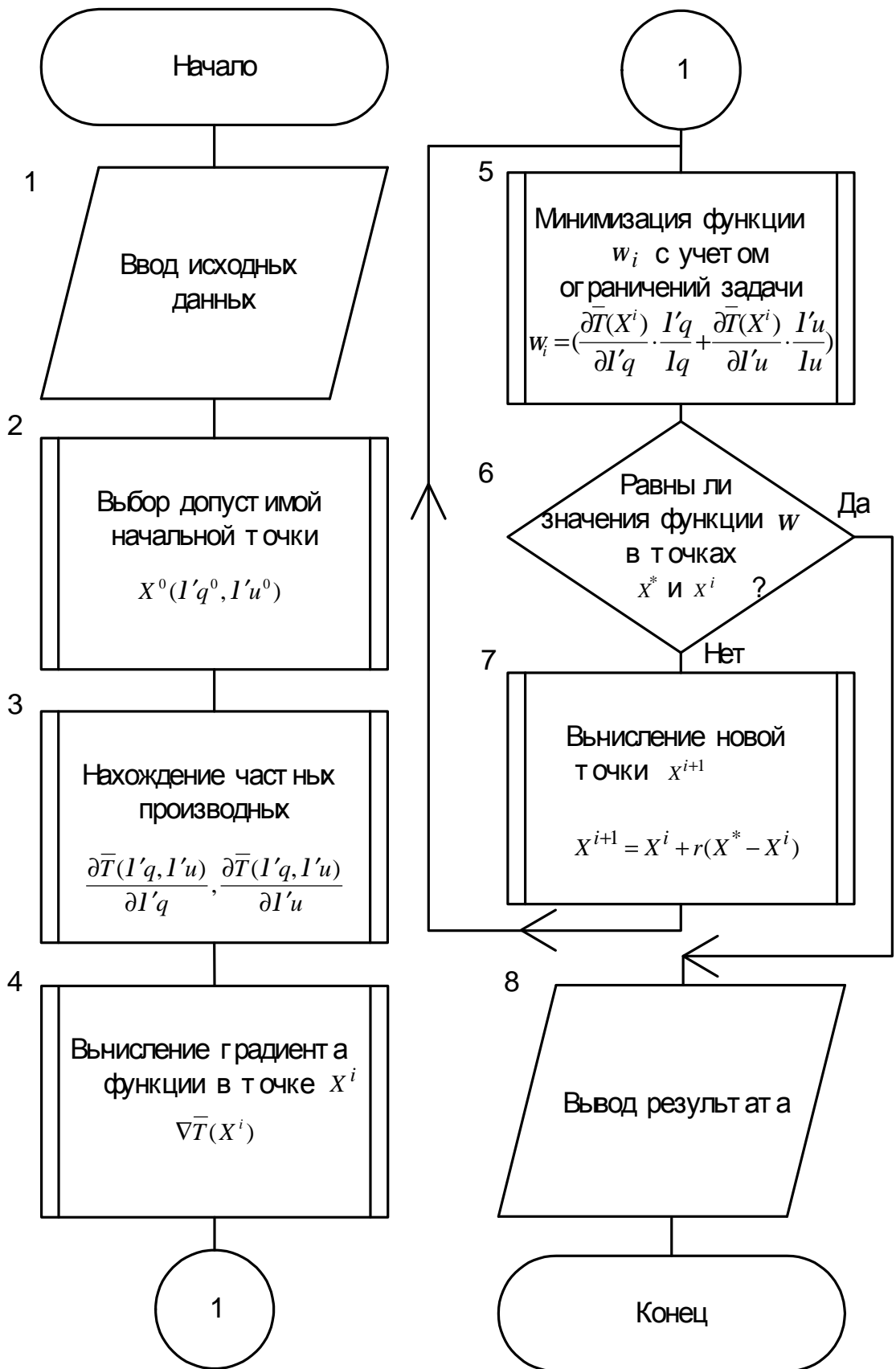


Рисунок 19 – Схема алгоритма вычисления оптимальной загруженности резервного узла РБД при репликации

Алгоритм содержит следующие этапы:

1. Ввод исходных данных.

На этом этапе производится сбор и обработка данных необходимых для корректной работы алгоритма:

- МО времени обработки поискового запроса на главном сервере – $M[tgq]$, на резервном сервере – $M[trq]$;
- МО времени обработки запроса на обновление на главном сервере – $M[tgu]$, на резервном сервере – $M[tru]$;
- МО времени, требуемого главному серверу на отправку одного сообщения для обновления БД резервного сервера – $M[trgu]$
- МО времени передачи обновления с главного сервера на резервный – $M[tgru]$;
- МО времени передачи запроса с резервного сервера на главный – $M[trgq]$ и МО времени передачи отклика на запрос с главного сервера на резервный – $M[tgrqotkl]$;
- I_u – общая интенсивность запросов на обновление;
- I_q – общая интенсивность поисковых запросов;
- I'_u – интенсивность запросов на обновление, обрабатываемых на резервном узле;
- I'_q – интенсивность поисковых запросов, обрабатываемых на резервном узле;
- n – число резервных серверов.
- e_{treb} – точность проверки результата.

2. Выбор допустимой начальной точки.

На данном этапе выбирается начальная точка $X^0(I'_q^0, I'_u^0)$, которая принадлежит области допустимых решений и находится на пересечении линейных ограничений. Для этого все ограничения модели приводятся к стандартной (канонической) форме:

$$I'q \cdot M[trq] + I'u \cdot M[tru] + s_1 = 0, \quad (49)$$

$$Iu \cdot M[tgu] + I'u \cdot n \cdot M[trgu] + n \cdot (Iq - I'q) \cdot M[tgq] + s_2 = 0, \quad (50)$$

$$I'u \cdot M[tgru] + (Iq - I'q) \cdot M[tgrqotkl] + s_3 = 0, \quad (51)$$

$$(Iq - I'q) \cdot M[trgq] + \frac{I'u}{n} \cdot M[trgu] + s_4 = 0, \quad (52)$$

$$I'q \cdot M[trq] + I'u \cdot M[tru] + s_5 = 1, \quad (53)$$

$$Iu \cdot M[tgu] + I'u \cdot n \cdot M[trgu] + n \cdot (Iq - I'q) \cdot M[tgq] + s_6 = 1, \quad (54)$$

$$I'u \cdot M[tgru] + (Iq - I'q) \cdot M[tgrqotkl] + s_7 = 1, \quad (55)$$

$$(Iq - I'q) \cdot M[trgq] + \frac{I'u}{n} \cdot M[trgu] + s_8 = 1, \quad (56)$$

$$\frac{I'q}{Iq} + s_9 = 1, \quad (57)$$

$$\frac{I'u}{Iu} + s_{10} = 1, \quad (58)$$

$$\frac{I'q}{Iq}, \frac{I'u}{Iu}, s_1, s_2, s_3, s_4, s_5, s_6, s_7, s_8, s_9, s_{10} \geq 0, \quad (59)$$

где s_i – вспомогательные переменные, включаемые в уравнения ограничений (остаточные для неравенства типа " \leq " и избыточные для неравенства типа " \geq ") для приведения неравенств к виду равенств [86].

Так как каждое из ограничений может быть изображено на плоскости оптимизируемых переменных в виде прямой, проходящей через точки пересечения с осями координат $I'q = 0$ и $I'u = 0$, то в результате получается многогранник ограничений, точки внутри и на границе которого относятся к допустимым, рисунок 20.

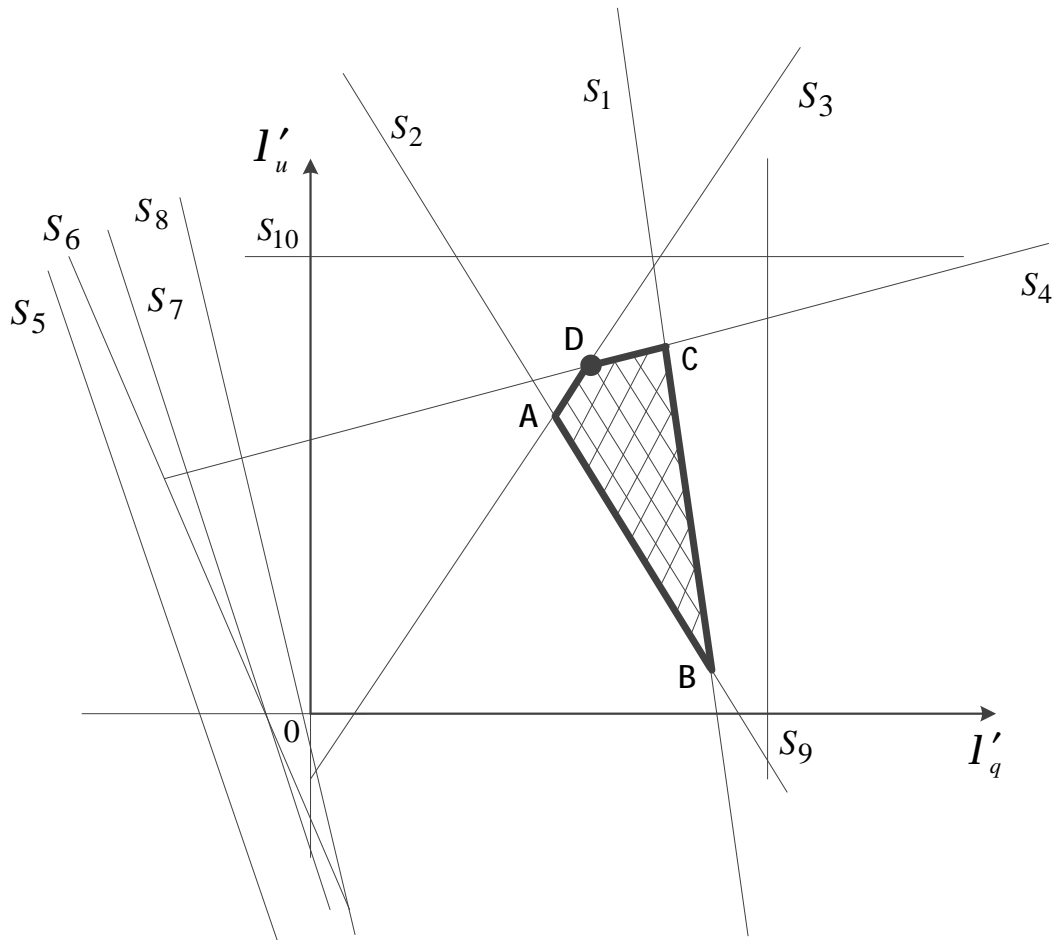


Рисунок 20 – Выбор начальной точки

Угловые точки называются *экстремальными*, так как им соответствуют нулевые значения переменных. Однозначное определение экстремальных точек возможно путем решения системы уравнений, приравнивая к нулю две переменные s_i :

$$\begin{cases}
I'q \cdot M[trq] + I'u \cdot M[tru] + s_1 = 0 \\
Iu \cdot M[tgu] + I'u \cdot n \cdot M[trgu] + n \cdot (Iq - I'q) \cdot M[trgq] + s_2 = 0 \\
I'u \cdot M[tgru] + (Iq - I'q) \cdot M[tgrqotkl] + s_3 = 0 \\
(Iq - I'q) \cdot M[trgq] + \frac{I'u}{n} \cdot M[trgu] + s_4 = 0 \\
I'q \cdot M[trq] + I'u \cdot M[tru] + s_5 = 1 \\
Iu \cdot M[tgu] + I'u \cdot n \cdot M[trgu] + n \cdot (Iq - I'q) \cdot M[trgq] + s_6 = 1 \\
I'u \cdot M[tgru] + (Iq - I'q) \cdot M[tgrqotkl] + s_7 = 1 \\
(Iq - I'q) \cdot M[trgq] + \frac{I'u}{n} \cdot M[trgu] + s_8 = 1 \\
\frac{I'q}{Iq} + s_9 = 1 \\
\frac{I'u}{Iu} + s_{10} = 1.
\end{cases} \quad (60)$$

3. Нахождение частных производных $\frac{\partial \bar{T}(I'q, I'u)}{\partial I'q}$ и $\frac{\partial \bar{T}(I'q, I'u)}{\partial I'u}$ в

соответствии с исходными данными.

4. Вычисление градиента функции в точке $\nabla \bar{T}(X^i)$.

5. Минимизация функции $w_i = \left(\frac{\partial \bar{T}(I'q^i, I'u^i)}{\partial I'q} \cdot \frac{I'q}{Iq} + \frac{\partial \bar{T}(I'q^i, I'u^i)}{\partial I'u} \cdot \frac{I'u}{Iu} \right)$ с

учетом ограничений исходной задачи. Данный этап алгоритма представляет собой задачу линейного программирования по отысканию экстремума выпуклой линейной целевой функции при наличии ограничений в виде линейных неравенств:

$$Z = \min \sum_i c_i \cdot x_i, \quad (61)$$

при ограничениях:

$$\sum_i a_{ij} \cdot x_i \leq b_j, \quad (62)$$

$$x_i \geq 0, \quad (63)$$

где $i = 1, n$; $j = 1, m$; n – число переменных; m – число линейно независимых ограничений; b_j – значения ограничений; a_{ij} – коэффициенты при переменных в ограничениях.

Такие задачи эффективно решаются симплекс-методом [86]. Схема алгоритма, адаптированного под условия исходной задачи, оформленная согласно [21], изображена на рисунке 21.

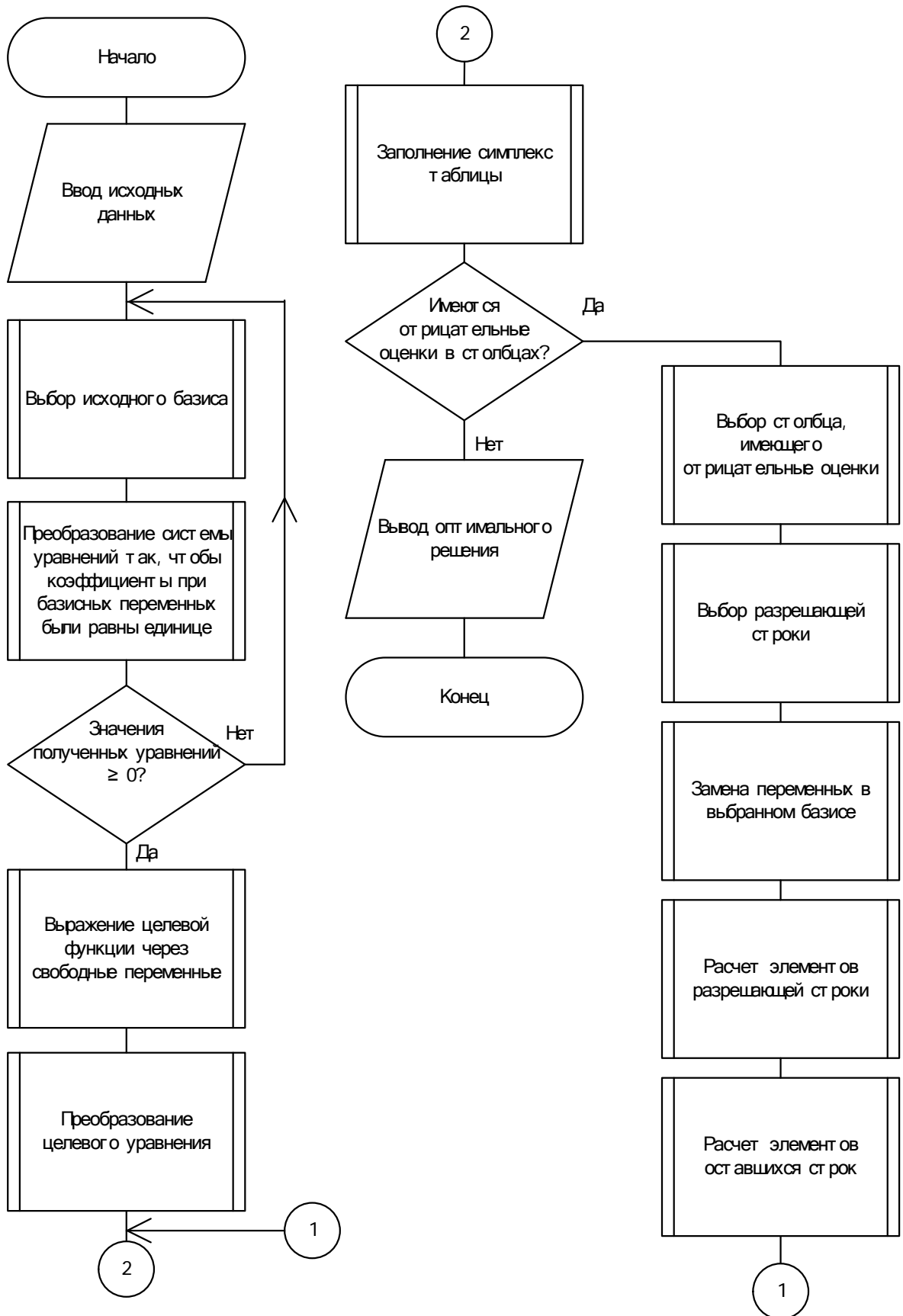


Рисунок 21 – Схема алгоритма адаптированного симплекс-метода

6. Условие проверки допустимости полученного результата.

Выбор требуемой точности проверки условия равенства результатов работы алгоритма в смежных итерациях целесообразно проводить с учетом максимальных погрешностей вычислений, возникающих в алгоритме. Основная погрешность алгоритма проявляется в блоке ввода исходных данных на этапе сбора и обработки данных при вычислении математических ожиданий исходных величин. Помимо этого снижение точности вычисления результатов алгоритма возникает в блоке вычисления новой точки для метода линейных комбинации, так как выполняется на основе приближенного метода золотого сечения. Однако погрешность в блоке вычисления новой точки на несколько порядков ниже погрешности, возникающей при сборе и обработке исходных данных, в связи с чем, ею можно пренебречь.

Погрешность, возникающая при сборе и обработке исходных данных, составляет микросекунды, следовательно, точность проверки условия равенства должна составлять $\frac{1}{2} \cdot 10^{-5} \leq e_{ireb} \leq 10^{-6}$ секунд, чтобы не противоречить условиям допустимой точности алгоритма.

7. Вычисление новой точки X^{i+1} для очередной итерации алгоритма.

На данном этапе определяется значение коэффициента r выражения $X^{i+1} = X^i + r(X^* - X^i)$, устремляющее к минимуму целевую функцию (6), что относится к задаче одномерной минимизации функции на отрезке.

На практике существует два наиболее простых специальных метода решения задач одномерной минимизации унимодальной функции на отрезке: метод поиска Фибоначчи и метод золотого сечения [24].

Метод поиска Фибоначчи применяется при фиксированном количестве обращений к процедуре расчета функции. Тогда как метод золотого сечения не требует задания окончательного интервала неопределенности процедуры

поиска. Метод золотого сечения просматривает точки, дробящие интервал неопределенности в отношениях, заданных выражением [24]:

$$\lim_{k \rightarrow \infty} \frac{F_{k-1}}{F_k} = \frac{2}{1 + \sqrt{5}} \equiv t \approx 0.6180, \quad (64)$$

где t – решение квадратного уравнения $t^2 + t - 1 = 0$.

8. Вывод результата.

Полученные значения $I'q$ и $I'u$ позволяют принимать решения о выборе на резервных серверах фрагментов данных для немедленной репликации таким образом, чтобы время отклика РБД на запросы при заданных условиях её функционирования стремилось к минимуму.

3.4 Свойства алгоритма вычисления оптимальной загруженности резервного узла распределенной базы данных при репликации

3.4.1 Оценка корректности алгоритма

Алгоритм называется *корректным*, если выполняются следующие условия [11, 7]:

1. После выполнения конечного числа элементарных операций алгоритм позволяет преобразовать любые входные данные в результат.
2. Результат устойчив по отношению к малым возмущениям входных данных.

Доказательства корректности алгоритма состоит в следующем:

1. В алгоритме выделяются критические фрагменты. Применительно к алгоритму вычисления оптимальной загруженности резервного узла РБД при репликации критическими фрагментами являются:

– расчет составляющих функции $\bar{T}(l'q, l'u)$: $\bar{Tr}(l'q, l'u)$, $\bar{Tgr}(l'q, l'u)$, $\bar{Trg}(l'q, l'u)$, $\bar{Tg}(l'q, l'u)$.

В соответствии с формулой Поллачека-Хинчина для СМО типа 1/M/G/FCFS для среднего времени ожидания в очереди [13]:

$$w = \frac{\frac{1}{2} \sum_{i=1}^N I_i b_i^{(2)}}{1 - R_N}, \quad (65)$$

где N – число типов заявок на обслуживание, b_i – среднее время обслуживания заявок i -го типа, I_i – интенсивность поступления заявок i -го типа на обслуживание, вычисления должны проводиться при ограничении:

$$0 \leq R_N < 1, \quad (66)$$

где $R_N = \sum_{i=1}^N r_i$, а r_i – коэффициент загрузки прибора заявками i -го типа.

– цикл с условием продолжения 4-7.

2. Определяются предусловия и постусловия для выделенных фрагментов.

Отсутствие аварийного останова алгоритма при расчете составляющих функции $\bar{T}(l'q, l'u)$ определяется корректным подбором значений параметров РБД в соответствии с ограничением $0 \leq R_N < 1$.

Условием выхода из цикла 4-7 является признак достижения требуемой точности вычисления e_{treb} в соответствии с применяемым численным градиентным методом линейных комбинаций.

3. Включение полученных предусловий в систему условий корректности алгоритма.

Таким образом, можно считать доказанной корректность алгоритма по Бейберу [7] при соблюдении следующих предусловий:

$$U = \begin{cases} |X^i - X^*| \leq e_{\text{treb}}; \\ 0 \leq R_N < 1 \end{cases} \quad (67)$$

То есть, корректность алгоритма гарантирована при заданной точности проверки условия равенства результатов его работы в смежных итерациях

$\frac{1}{2} \cdot 10^{-5} \leq e_{\text{treb}} \leq 10^{-6}$ секунд и выполнении требований по загруженности РБД

$0 \leq R_N < 1$.

3.4.2 Оценка сложности алгоритма

Для оценки сложности алгоритма необходимо рассчитать число элементарных шагов, которые выполняет алгоритм при заданном размере входных данных с учетом времени их выполнения, таблица 8 [4, 2].

Таблица 8 – Распределение времени выполнения этапов алгоритма

Номер этапа	Этап алгоритма	Временная стоимость выполнения этапа	Число раз	Общее время выполнения этапа
1	Ввод исходных данных	c_1	1	c_1
2	Выбор начальной точки	c_2	1	c_2
3	Вычисление частных производных	c_3	1	c_3
4	Вычисление градиента функции в точке	c_4	n	$c_4 \cdot n$
5	Минимизация функции	c_5	n	$c_5 \cdot n$
6	Условие проверки получения результата удовлетворяющего заданной точности	c_6	n	$c_6 \cdot n$
7	Вычисление новой точки	c_7	$n - 1$	$c_7 \cdot (n - 1)$
8	Вывод результатов	c_8	1	c_8

Временная сложность алгоритма определяется как сумма временных сложностей каждого из этапов алгоритма с учетом размера входных данных:

$$T(n) = c_{1-3,8} + c_{4-6} \cdot n + c_7 \cdot (n - 1), \quad (68)$$

где c – временная стоимость соответствующего шага алгоритма, а n – размер входных данных.

Из формулы видно, что функция времени работы алгоритма имеет линейную зависимость от размера входных данных, следовательно, алгоритм относится к классу полиномиальных по времени.

3.4.3 Оценка точности алгоритма

В ходе выполнения алгоритма возникают погрешности различного рода, снижающие точность получаемых оценок [49, 4].

Как правило, общая погрешность вычислений обусловлена рядом причин:

1. используемые в вычислениях модели описывают реальные процессы приближенно;
2. неточность исходных данных;
3. погрешность расчетов в ЭВМ.

В результате возникают: неустранимые погрешности d_n , вычислительные погрешности d_g и погрешности метода d_m .

При этом для удовлетворительной точности алгоритма отношение неустранимой погрешности d_n и погрешности метода d_m должно лежать в диапазоне от 2 до 10, включительно. А отношение вычислительной погрешности d_g и погрешности метода d_m должно составлять не менее одного порядка [11]:

$$U = \begin{cases} \frac{d_n}{d_m} = [2...10] \\ \frac{d_m}{d_g} \geq 10. \end{cases} \quad (69)$$

Неустраняемая погрешность определяется как:

$$d_n \approx 10^{-N+1}, \quad (70)$$

где N – длина мантииссы исходных данных.

Наибольшая погрешность значений исходных данных алгоритма проявляется в значениях МО заданных характеристик и не превышает микросекунд. Следовательно, $d_n \approx 10^{-5}$.

Что касается погрешности метода, то она определяется заданной точностью проверки условия равенства результатов работы алгоритма в смежных итерациях $\frac{1}{2} \cdot 10^{-5} \leq e_{\text{treb}} \leq 10^{-6}$ секунд. Следовательно, погрешность метода колеблется в диапазоне $d_m = [10^{-6}, \frac{1}{2} \cdot 10^{-5}]$.

Вычислительная погрешность алгоритма d_g определяется внутренней архитектурой ЭВМ. В современных условиях точность ЭВМ на много порядков выше величины 10^{-6} . Следовательно, $d_g \ll 10^{-6}$.

Полученные значения погрешностей удовлетворяют системе критериев (69) и позволяют сформулировать вывод о достаточной точности алгоритма.

3.4.4 Оценка вычислительной устойчивости алгоритма

Оценка вычислительной устойчивости алгоритма проводится на основе анализа погрешности результатов вычисления в зависимости от числа шагов алгоритма (рисунок 22).

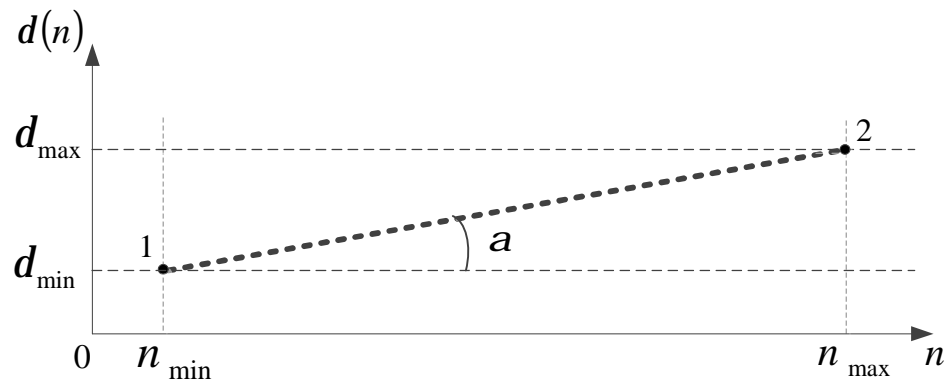


Рисунок 22 – Аппроксимация линейной функцией зависимости погрешности результатов работы алгоритма вычисления оптимальной загруженности резервного узла распределенной базы данных при репликации от числа шагов

Вычислительная устойчивость алгоритма определяется углом наклона линии 1-2, аппроксимирующей линейной функцией зависимость погрешности результатов работы алгоритма от числа шагов. При этом анализ данных показывает, что для оцениваемого алгоритма угол наклона линии 1-2 стремится к нулю $\angle a \rightarrow 0$, поскольку разность $d_{\max} - d_{\min}$ более высокого порядка малости, чем разность $n_{\max} - n_{\min}$, что говорит о достаточной вычислительной устойчивости алгоритма.

3.5 Алгоритм выбора фрагментов данных для немедленной репликации

Фрагменты данных для немедленной репликации целесообразно выбирать с учетом полученных оптимальных значений интенсивностей поисковых и запросов на обновление таким образом, чтобы средний объем пересылаемых по сети данных был минимальным.

В соответствии с этапом сбора статистической информации об интенсивности поисковых запросов и запросов на обновление к фрагментам РБД имеются характеристики:

- $V = (v_1, v_2, \dots, v_n)$, где n – количество рассматриваемых фрагментов РБД, а V – вектор-строка характеристик размера рассматриваемых фрагментов РБД;

- $\Lambda q = (Iq_1, Iq_2, \dots, Iq_n)$, где Λq – вектор-строка характеристик поисковых запросов к рассматриваемым фрагментам РБД;

- $\Lambda u = (Iu_1, Iu_2, \dots, Iu_n)$, где Λu – вектор-строка характеристик запросов на обновление к рассматриваемым фрагментам РБД.

В качестве рассматриваемых фрагментов РБД могут выступать целые таблицы, отдельные столбцы, отдельные строки, ячейки, в зависимости от детализации решаемой задачи и возможностей конкретных СУБД.

Минимум среднего объема пересылаемой по сети информации достигается в том случае, если совокупный размер фрагментов РБД, задействованных в немедленной репликации, минимальный, при условии обслуживания этими фрагментами максимального числа поисковых запросов на резервных серверах в соответствии с доступными сетевыми и вычислительными ресурсами.

Таким образом, возникает необходимость решения задачи минимизации размера пересылаемой по сети информации с учетом ограничений на полученные значения Iq^{opt} и Iu^{opt} . При этом:

$$Vq(x_1, x_2, \dots, x_n) = \bar{x}_1 \cdot lq_1 \cdot n_1 + \bar{x}_2 \cdot lq_2 \cdot n_2 + \dots + \bar{x}_n \cdot lq_n \cdot n_n, \quad (71)$$

где Vq – средний объем информации, пересылаемой по сети за единицу времени в рамках удаленного обслуживания поисковых запросов к фрагментам РБД ($x_i = 0$, если i -й фрагмент данных не задействован в немедленной репликации), а:

$$Vu(x_1, x_2, \dots, x_n) = x_1 \cdot lu_1 \cdot n_1 + x_2 \cdot lu_2 \cdot n_2 + \dots + x_n \cdot lu_n \cdot n_n, \quad (72)$$

где Vu – средний объем информации, пересылаемой по сети за единицу времени в рамках обновления фрагментов данных для немедленной репликации ($x_i = 1$, если i - й фрагмент данных задействован в немедленной репликации).

Тогда задача минимизации приобретает вид:

$$Vq(x_1, x_2, \dots, x_n) + Vu(x_1, x_2, \dots, x_n) \rightarrow \min(x_1, x_2, \dots, x_n). \quad (73)$$

После подстановки слагаемых:

$$n_1 \cdot (x_1 \cdot (lu_1 - lq_1) + lq_1) + n_2 \cdot (x_2 \cdot (lu_2 - lq_2) + lq_2) + \dots + n_n \cdot (x_n \cdot (lu_n - lq_n) + lq_n) \rightarrow \min(x_1, x_2, \dots, x_n), \quad (74)$$

при наличии ограничений:

$$\begin{cases} x_1 \cdot lu_1 + x_2 \cdot lu_2 + \dots + x_n \cdot lu_n \leq lu^{opt} + d \\ lu^{opt} - d \leq x_1 \cdot lu_1 + x_2 \cdot lu_2 + \dots + x_n \cdot lu_n \\ x_1 \cdot lq_1 + x_2 \cdot lq_2 + \dots + x_n \cdot lq_n \leq lq^{opt} + d \\ lq^{opt} - d \leq x_1 \cdot lq_1 + x_2 \cdot lq_2 + \dots + x_n \cdot lq_n \\ x_1, x_2, \dots, x_n \in \{0,1\} \end{cases} \quad (75)$$

Задачи такого типа относятся к, так называемым, задачам многомерной двоичной оптимальной упаковки или, по-другому, к задачам о "0-1 рюкзаке" [67, 4].

Решение возможно на основе метода полного перебора, частично-целочисленного линейного программирования на основе аддитивного алгоритма для задач с двоичными переменными (частный случай метода ветвей и границ), динамического программирования.

Характеристики математических методов решения задачи многомерной двоичной оптимальной упаковки представлены в таблице 9.

Таблица 9 – Характеристики методов решения задачи упаковки

Свойства	Методы		
	Полный перебор	Ветвей и границ	Динамического программирования
Чувствительность к множеству ограничений	высокая	низкая	высокая
Требовательность к памяти	низкая	низкая	высокая
Вычислительная сложность	экспоненциальная	сводима к полиномиальной	сводима к полиномиальной
Сложность реализации алгоритмов	низкая	средняя	высокая
Возможность получения промежуточных результатов	нет	есть	есть

Сравнительный анализ представленных методов показывает, что наиболее подходящий для решения задачи многомерной двоичной оптимальной упаковки – метод ветвей и границ. Однако для сведения его вычислительной сложности к полиномиальной, необходимо использовать оптимизированный метод ветвей и границ с точностью вычисления результатов сниженной в допустимых пределах.

Для представления задачи в виде пригодном для решения методом ветвей и границ необходимо, чтобы в выражении целевой функции все коэффициенты были неотрицательны (это возможно только после подстановки в целевую функцию исходных значений характеристик фрагментов РБД). Все ограничения должны быть типа " \leq ".

Схема алгоритма выбора фрагментов данных для немедленной репликации, оформленная согласно [21], представлена на рисунке 23.

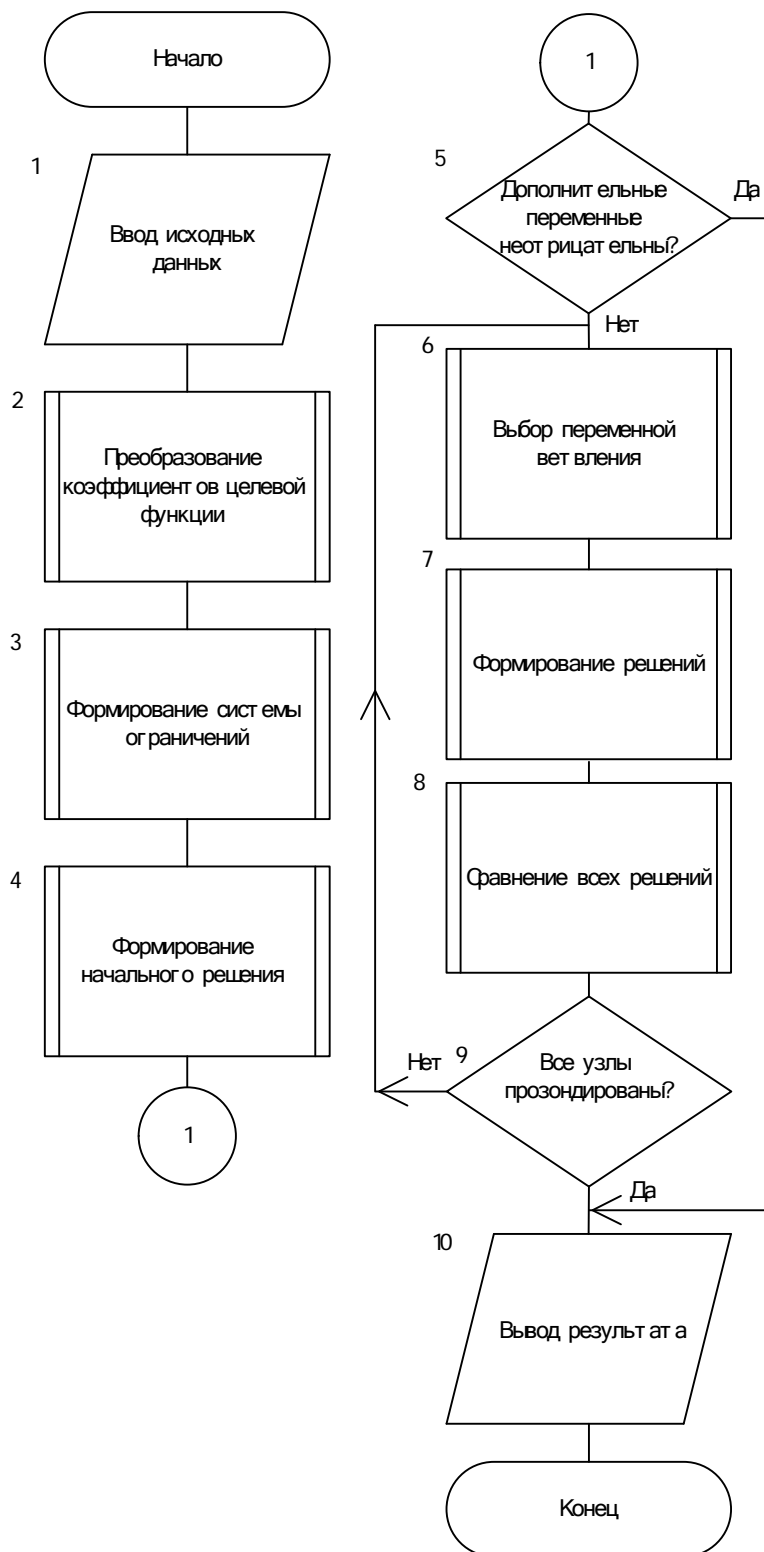


Рисунок 23 – Схема алгоритма выбора фрагментов данных для немедленной репликации

Алгоритм выбора фрагментов данных для немедленной репликации состоит из следующих этапов:

1. Ввод исходных данных. В качестве исходных данных выступают:

- $V = (v_1, v_2, \dots, v_n)$ – вектор-строка характеристик размера рассматриваемых фрагментов РБД, где n – количество рассматриваемых фрагментов РБД;

- $\Lambda q = (Iq_1, Iq_2, \dots, Iq_n)$ – вектор-строка характеристик поисковых запросов к рассматриваемым фрагментам РБД;

- $\Lambda u = (Iu_1, Iu_2, \dots, Iu_n)$ – вектор-строка характеристик запросов на обновление к рассматриваемым фрагментам РБД;

- Iu^{opt}, Iq^{opt} – оптимальные значения параметров репликации рассматриваемой РБД, полученные при помощи алгоритма вычисления оптимальных характеристик узла РБД при репликации.

2. Преобразование коэффициентов целевой функции.

После подстановки в целевую функцию исходных данных переменные x_i с отрицательными коэффициентами заменяются на $(1 - x'_i)$, а переменные x_i с положительными коэффициентами заменяются на (x'_i) :

$$z_1 \cdot x'_1 + z_2 \cdot x'_2 + \dots + z_n \cdot x'_n = Z. \quad (76)$$

3. Формирование системы ограничений.

На данном этапе переменные неравенств преобразуются в соответствии с целевой функцией, а каждое из неравенств преобразуется в равенство путем добавления дополнительных переменных:

$$\begin{cases} x'_1 \cdot a_1 + x'_2 \cdot a_2 + \dots + x'_n \cdot a_n + s_1 = A \\ -x'_1 \cdot b_1 - x'_2 \cdot b_2 - \dots - x'_n \cdot b_n + s_2 = B \\ x'_1 \cdot c_1 + x'_2 \cdot c_2 + \dots + x'_n \cdot c_n + s_3 = C \\ -x'_1 \cdot d_1 + x'_2 \cdot d_2 + \dots + x'_n \cdot d_n + s_4 = D \\ x'_1, x'_2, \dots, x'_n \in \{0,1\} \end{cases} \quad (77)$$

4. Формирование начального решения.

В начальном решении все двоичные переменные должны равняться нулю $x'_1, x'_2, \dots, x'_n = 0$. В этом случае дополнительные переменные будут базисными, а их значения определяются правыми частями ограничений [86]. Начальное решение представлено в таблице 10.

Таблица 10 – Начальное решение алгоритма выбора фрагментов реплик

Базис	x'_1	x'_2	...	x'_n	s_1	s_2	s_3	s_4	Решение
s_1	a_1	a_2	a_i	a_n	1	0	0	0	A
s_2	$-b_1$	$-b_2$	$-b_i$	$-b_n$	0	1	0	0	B
s_3	c_1	c_2	c_i	c_n	0	0	1	0	C
s_4	$-d_1$	$-d_2$	$-d_i$	$-d_n$	0	0	0	1	D
Коэффициент целевой функции	z_1	z_2	z_i	z_n					

Так как все двоичные переменные равны нулю, то дополнительные переменные принимают значения: $(s_1, s_2, s_3, s_4) = (A, B, C, D)$.

5. Проверка оптимальности начального решения.

Если значения всех дополнительных переменных (A, B, C, D) при нулевых значениях двоичных переменных – неотрицательные, то рассматриваемое решение – оптимальное. В противном случае, необходимо прозондировать все узлы графа переходов.

6. Выбор переменной ветвления.

Переменная ветвления необходима для уменьшения абсолютной величины отрицательных значений дополнительных переменных. Выбор переменной ветвления производится на основе меры недопустимости дополнительной переменной:

$$I_j = \sum_i \min\{0, s_i - a_{ij}\}, \quad (78)$$

где a_{ij} – коэффициент при переменной x'_j в i -ом ограничении.

Процесс выбора узлов ветвления можно представить в виде графа переходов (рисунок 24).

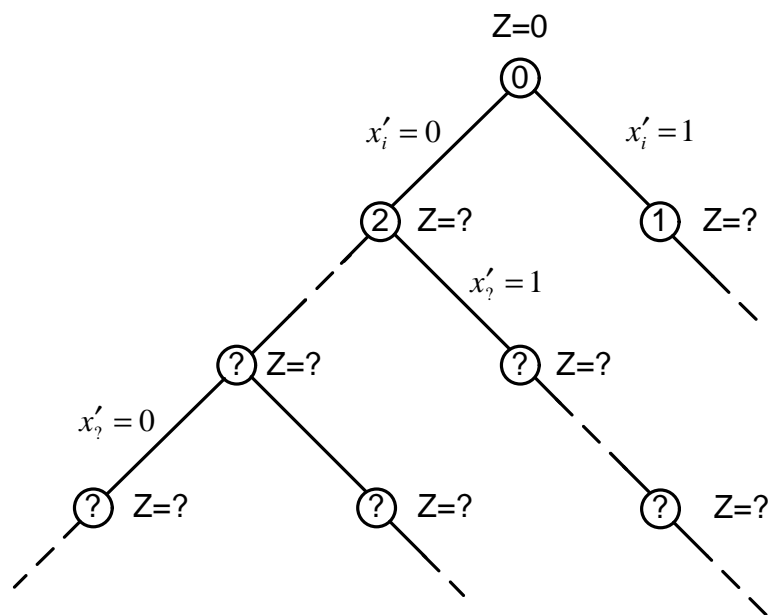


Рисунок 24 – Граф переходов переменных ветвления

7. Формирование решений.

Рассчитываются значения целевой функции для выбранного набора переменных ветвления.

8. Сравнение всех решений.

Определяются возможные варианты дальнейшего ветвления. Те узлы, в которых значения дополнительных переменных приобретают отрицательные значения, а целевая функция удаляется от оптимального решения, убираются из дальнейшего рассмотрения.

9. Проверка, прозондированы ли все узлы?

Если все узлы прозондированы, то в качестве решения принимается набор двоичных переменных, определяющих минимум целевой функции. В противном случае, выполняется переход на этап № 6 – выбор переменной ветвления.

10. Вывод результата.

Полученное на предыдущих этапах алгоритма оптимальное решение или сообщение об его отсутствии выводится в качестве результата работы алгоритма.

3.6 Свойства алгоритма выбора фрагментов данных для немедленной репликации

3.6.1 Оценка корректности алгоритма

Критической частью алгоритма выбора фрагментов данных для немедленной репликации является цикл 6-9 с выходом по условию. Условием корректного завершения цикла 6-9 является конечное число точек зондирования, определяемое степенью фрагментации данных. Таким образом, алгоритм выбора фрагментов данных для немедленной репликации корректен по Бейберу [7] при соблюдении предусловия:

$$n \neq \infty, \quad (79)$$

где n – количество обрабатываемых фрагментов данных РБД с учетом ресурсов ЭВМ, на которой будут производиться вычисления.

3.6.2 Оценка вычислительной сложности алгоритма

Задача нахождения точного набора фрагментов данных для немедленной репликации относится к классу NP – полных. В худшем случае сложность ее решения классическим методом ветвей и границ сопоставима со сложностью полного перебора – $O(n \cdot 2^n)$, что является недопустимым при больших значениях n . В работе [67] доказано, что за счет допустимого снижения точности результата можно свести решение рассматриваемой задачи к полностью полиномиальному с вычислительной сложностью – $O(\frac{n^4}{d})$, где d – погрешность решения. Чем больше n , тем больше выигрыш в соотношении точность/сложность по сравнению с классическим методом решения.

Вычислительная сложность представленного алгоритма выбора фрагментов данных для немедленной репликации соизмерима со сложностью $O(\frac{n^4}{d})$, так как дополнительные этапы преобразования и ввода/вывода данных не оказывают существенного влияния на сложность известного алгоритма [2].

3.6.3 Оценка точности алгоритма

Оценка погрешностей алгоритма показывает, что основной вклад в погрешность результата вычислений алгоритма вносит погрешность метода d_m .

Очевидно, что вычислительная погрешность алгоритма d_e , определяемая внутренней архитектурой ЭВМ, не оказывает ощутимого влияния на точность алгоритма.

Неустраняемая погрешность d_n проявляется при получении исходных данных.

В работе [67] доказано, что погрешность оптимизированного метода ветвей и границ при снижении сложности до полиномиальной не превышает 5%.

Проведенная оценка погрешностей позволяет сформулировать вывод о достаточной точности алгоритма.

3.6.4 Оценка вычислительной устойчивости

Оценка вычислительной устойчивости алгоритма проводится на основе анализа погрешности результатов вычисления в зависимости от числа шагов алгоритма.

На рисунке 25 представлен график зависимости погрешности результатов вычислений алгоритма от числа шагов.

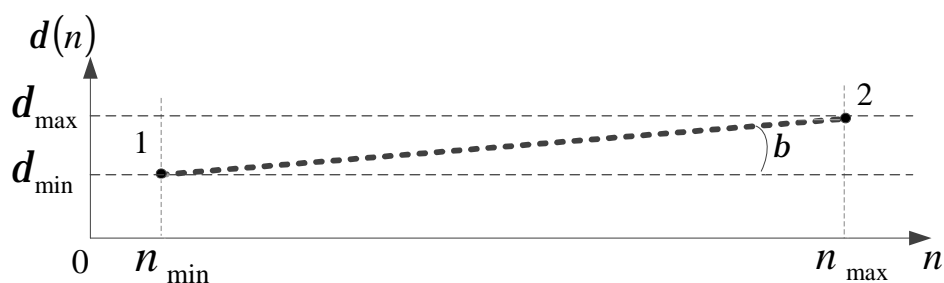


Рисунок 25 – Аппроксимация линейной функцией зависимости погрешности результатов работы алгоритма выбора фрагментов данных для немедленной репликации от числа шагов

Вычислительная устойчивость алгоритма определяется углом наклона линии 1-2, аппроксимирующей линейной функцией зависимость погрешности результатов работы алгоритма от числа шагов. Для оцениваемого алгоритма угол наклона линии 1-2 стремится к нулю $\angle b \rightarrow 0$, а увеличение погрешности распределено равномерно по всей области

определения n и фактически совпадает с линией 1-2, что говорит о достаточной вычислительной устойчивости алгоритма.

Выводы по главе 3

1. Представленные алгоритмы направлены на решение задачи минимизации среднего времени отклика РБД на запросы за счет обоснованного выбора фрагментов данных для немедленной репликации.

2. Решение задачи минимизации среднего времени отклика РБД на запросы аналитическими методами сопряжено с решением уравнений шестого порядка. В связи с этим решение целесообразно осуществлять в рамках численных методов.

3. Алгоритм вычисления оптимальной загруженности резервного узла РБД при репликации целесообразно строить на основе метода линейных комбинаций. Полученный алгоритм является корректным, обладает достаточной точностью, вычислительной устойчивостью и допустимой сложностью.

4. Задачу выбора фрагментов данных для немедленной репликации целесообразно решать на основе метода ветвей и границ, модифицированного с учетом приведения к полиномиальной сложности.

5. Полученный алгоритм выбора фрагментов данных для немедленной репликации позволяет находить решения за полиномиальное время с учетом снижения точности вычислений не более чем на 5 %. Алгоритм является корректным, обладает достаточной точностью, вычислительной устойчивостью и допустимой сложностью.

ГЛАВА 4. РАЗРАБОТКА АЛГОРИТМА АВТОМАТИЗАЦИИ ПРОЦЕССА КОНФИГУРИРОВАНИЯ РЕПЛИКАЦИИ В РБД ПРЕДПРИЯТИЯ ГПК

Данная глава посвящена реализации способа управления репликацией в РБД, защищенного патентом на изобретение ("Способ репликации информации в распределенных базах данных с конкурентным распределением потоков" (заявка о выдаче патента на изобретение № 2012116021)), в виде алгоритма автоматизации процесса конфигурирования репликации в РБД предприятия ГПК, отличающегося обеспечением условий функционирования РБД предприятия ГПК с пониженным средним временем отклика на запросы, позволяющего вычислять значения параметров репликации в РБД предприятия ГПК и формировать решения для администратора по её конфигурированию.

4.1. Алгоритм автоматизации процесса конфигурирования репликации в РБД предприятия ГПК

Для получения решений по конфигурированию репликации в РБД предприятия ГПК с учетом минимума среднего времени отклика РБД на запросы необходима разработка алгоритма, состоящего из следующих этапов:

1. Вычисление значений характеристик РБД.
2. Вычисление допустимых оптимальных значений параметров РБД при репликации на основе алгоритма вычисления оптимальной загруженности резервного узла.
3. Обоснованный выбор фрагментов данных для немедленной репликации на основе алгоритма выбора фрагментов данных для немедленной репликации.
4. Формирование решений для администратора РБД предприятия ГПК по конфигурированию репликации.

Схема алгоритма автоматизации процесса конфигурирования репликации в РБД предприятия ГПК, оформленная согласно [21], представлена рисунке 26 [30]:

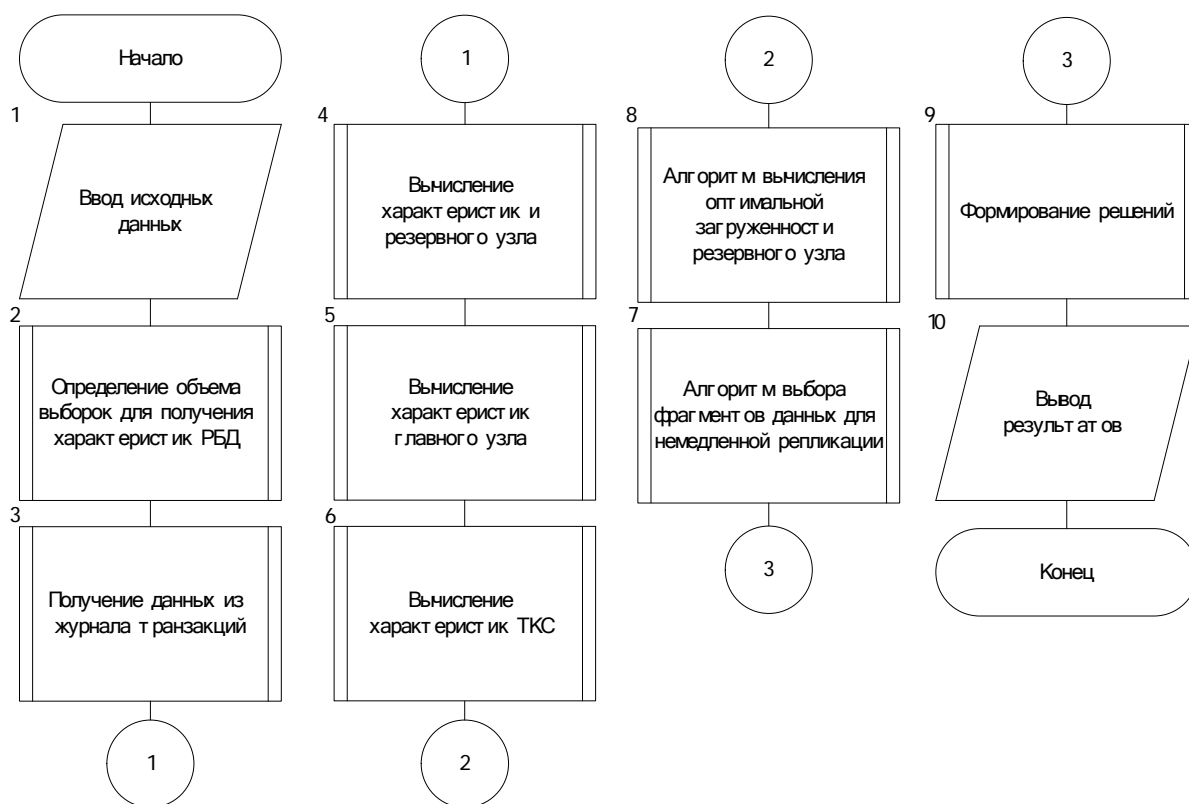


Рисунок 26 – Схема алгоритма автоматизации процесса конфигурирования репликации в РБД предприятия ГПК

4.2. Этап вычисления значений характеристик РБД

На первом этапе алгоритма конфигурирования репликации в РБД предприятия ГПК выполняется вычисление значений характеристик РБД, необходимых для решения задачи минимизации среднего времени отклика РБД на запросы.

При этом важным вопросом является определение объема выборочной совокупности наблюдений над количественным признаком для обеспечения требуемой точности вычислений [23].

Так как генеральная средняя x_g неизвестна, то в качестве её оценки принимают выборочную среднюю [20]:

$$\bar{x}_B = (x_1 + x_2 + \dots + x_n) / n, \quad (80)$$

где n – объем выборки; x_1, x_2, \dots, x_n – значения признака.

Величина \bar{x}_B является несмещенной оценкой генеральной средней x_G . Данное утверждение подтверждается тем, что значения признака x_1, x_2, \dots, x_n можно рассматривать как случайные величины X_1, X_2, \dots, X_n , а поскольку они имеют одинаковое распределение, то числовые характеристики, в частности математические ожидания, одинаковы. Другими словами, так как математическое ожидание среднего арифметического одинаково распределенных случайных величин равно математическому ожиданию каждой из величин, то числовые характеристики этих величин и генеральной совокупности одинаковы [20].

В соответствии с теоремой Чебышева, при увеличении объема выборки выборочная средняя стремится по вероятности к генеральной и является её состоятельной оценкой.

С другой стороны, центральная предельная теорема для одинаково распределенных слагаемых гласит, что если X_1, X_2, \dots, X_n – независимые случайные величины, имеющие одно и то же распределение с математическим ожиданием m и дисперсией s^2 , то при увеличении объема выборки закон распределения суммы:

$$Y_n = \sum_{k=1}^n X_k, \quad (81)$$

неограниченно приближается к нормальному [19].

Данные положения позволяют для нахождения минимального объема выборки с заданной точностью и надежностью использовать выражение [20]:

$$n = \frac{t^2 \cdot s^2}{d^2}, \quad (82)$$

где t – коэффициент доверия, s – среднее квадратическое отклонение вариационного признака, d – предельная ошибка выборки.

Предельная ошибка выборки d определяется в соответствии с заданной доверительной вероятностью g на основе выражения [47, 20, 19]:

$$P[|\Theta - \Theta^*| < d] = g, \quad (83)$$

где Θ – значение неизвестного параметра, Θ^* – оценка неизвестного параметра.

На практике задаются надежностью g равной 0,95; 0,99 и 0,999 [20]. Значение $g = 0,999$ характеризует максимальную надежность определения минимального объема выборки, применяемое на практике.

Значение коэффициента доверия t определяется на основе функции Лапласа [20]:

$$\Phi(t) = \frac{g}{2}, \quad (84)$$

Для $g = 0,999$, в соответствии с табличными значениями функции Лапласа:

$$\Phi(x) = \frac{1}{\sqrt{2p}} \cdot \int_0^x e^{-\frac{x^2}{2}} dx, \quad (85)$$

коэффициент доверия $t \approx 3,2$.

Величина предельной ошибки выборки d задается, как правило, в пределах до 10% от предполагаемого среднего уровня признака [64]. Погрешности используемых в методике алгоритмов не превышают 5%. Следовательно, величина предельной ошибки выборки d не должна превышать 5% от предполагаемого среднего уровня признака.

Среднее квадратическое отклонение вариационного признака s определяется на основе оценки генеральной дисперсии выборки по исправленной выборочной [20]. Если в качестве оценки генеральной дисперсии использовать выборочную, то будут возникать систематические ошибки, занижающие значения генеральной дисперсии. Это связано с тем, что выборочная дисперсия является смещенной оценкой генеральной. Чтобы получить несмещенную оценку, выборочную дисперсию необходимо преобразовать в соответствии с выражением [20]:

$$s^2 = \frac{n}{n-1} \cdot D_B, \quad (86)$$

где s^2 – обозначение исправленной выборочной дисперсии.

Однако при больших значениях n объема выборки выборочная и исправленная дисперсия различается мало. На практике пользуются исправленной дисперсией, если выполняется условие $n < 30$ [20].

Представленные рассуждения позволяют при $n > 30$ в качестве генерального среднего квадратического отклонения использовать выборочное, полученное на основе выражений:

$$s_B = \sqrt{D_B}, \quad (87)$$

$$D_B = \frac{\sum_{i=1}^n (x_i - \bar{x}_B)^2}{n}, \quad (88)$$

где \bar{x}_B определяется с помощью формулы (80).

Благодаря представленным выражениям можно рассчитать минимально достаточный объем выборки для получения значения количественного признака с учетом обеспечения требуемой точности и надежности вычислений.

4.2.1 Оценка применимости этапа вычисления значений характеристик РБД предприятия ГПК "ШахтИнвестКузбасс"

Для оценки возможности реализации этапа вычисления значений характеристик РБД предприятия ГПК "ШахтИнвестКузбасс" необходимо оценить возможность получения математических ожиданий времени обслуживания запросов на различных этапах их обработки: $M[tgru]$, $M[tgq]$, $M[trq]$, $M[tgu]$, $M[tru]$, $M[trgu]$, $M[trgq]$, $M[tgrgotkl]$.

Для оценки достаточного объема выборки по каждой характеристике необходимо оценить значения времени обслуживания запросов на различных этапах их обработки более чем для 30 элементов выборки [20]. В таблице 11 представлены значения времени передачи обновления с главного сервера на резервный $tgru$ для РБД предприятия ГПК "ШахтИнвестКузбасс" (31 значение).

Таблица 11 – Начальная выборка значений $tgru$

№	Значение	№	Значение	№	Значение
1	0,0392	12	0,017	23	0,02
2	0,0012	13	0,0162	24	0,0167
3	0,0195	14	0,0178	25	0,0416
4	0,0328	15	0,0232	26	0,0182
5	0,011	16	0,0149	27	0,0186
6	0,0188	17	0,0049	28	0,0181
7	0,0091	18	0,0034	29	0,0353
8	0,0243	19	0,0121	30	0,0305
9	0,0346	20	0,0141	31	0,0456
10	0,0338	21	0,0153		
11	0,0385	22	0,021		

По формуле (80) $\bar{x}_B = 0,0215$. В соответствии с формулами (88) и (87) $s_B^2 = 0,0001274$. При уровне надежности $g = 0,999$ коэффициент доверия $t \approx 3,2$. Величина предельной ошибки выборки d при максимальном отклонении в 5% определяется как:

$$d = 0,05 \cdot \bar{x}_B = 0,05 \cdot 0,0215 = 0,001075 . \quad (89)$$

Тогда, минимальный объем выборки для получения значения характеристики $M[tgru]$ при заданной надежности $g = 0,999$ и точности 5% рассчитывается на основе формулы (82) как:

$$n = \frac{3,2^2 \cdot 0,0001274}{(0,001075)^2} \approx 1129 . \quad (90)$$

Полученный объем выборки говорит о том, что при обработке 1129 измерений времени передачи обновлений с главного сервера на резервный данная характеристика приближенно равна её МО $M[tgru]$.

В среднем каждую минуту в РБД предприятия ГПК "ШахтИнвестКузбасс" передача обновлений с главного сервера на резервный выполняется 12 раз. Тогда, для получения необходимого объема выборки – 1129 элемента следует проанализировать таблицы транзакции меньше чем за 2 часа функционирования РБД, что приемлемо.

В таблице 12 представлены значения времени обработки поискового запроса на главном сервере tgq для РБД предприятия ГПК "ШахтИнвестКузбасс" (31 значение).

Таблица 12 – Начальная выборка значений tgq

№	Значение	№	Значение	№	Значение
1	0,0026	12	0,0012	23	0,0042
2	0,0016	13	0,0011	24	0,0028
3	0,0031	14	0,0019	25	0,0027
4	0,0042	15	0,0021	26	0,0024
5	0,0041	16	0,0018	27	0,0049
6	0,0046	17	0,0023	28	0,0038
7	0,0007	18	0,0025	29	0,0053
8	0,0027	19	0,0030	30	0,0022
9	0,0040	20	0,0025	31	0,0046
10	0,0018	21	0,0026		
11	0,0022	22	0,0025		

По формуле (80) $\bar{x}_B = 0,0028$. В соответствии с формулами (88) и (87) $s_B^2 = 0,000001321$. Величина предельной ошибки выборки d при максимальном отклонении в 5% $d = 0,00014$.

Тогда, минимальный объем выборки для получения значения характеристики $M[*tgq*]$ при заданной надежности $g = 0,999$ и точности 5% рассчитывается на основе формулы (82) как:

$$n = \frac{3,2^2 \cdot 0,000001321}{(0,00014)^2} \approx 690. \quad (91)$$

Полученный объем выборки говорит о том, что при обработке 690 значений времени обработки поискового запроса на главном сервере tgq данная характеристика приближенно равна её МО $M[*tgq*]$.

В среднем каждую минуту в РБД предприятия ГПК "ШахтИнвестКузбасс" обработка поисковых запросов на главном сервере исчисляется сотнями. Тогда, для получения необходимого объема выборки следует проанализировать таблицы транзакций за несколько минут функционирования РБД, что приемлемо.

В таблице 13 представлены значения времени обработки поискового запроса на резервном сервере trq для РБД предприятия ГПК "ШахтИнвестКузбасс" (31 значение).

Таблица 13 – Начальная выборка значений trq

№	Значение	№	Значение	№	Значение
1	0,0044	12	0,0038	23	0,0046
2	0,0045	13	0,0042	24	0,0036
3	0,0062	14	0,0032	25	0,0051
4	0,0048	15	0,0021	26	0,0062
5	0,0047	16	0,0039	27	0,0071
6	0,0044	17	0,0041	28	0,0066
7	0,0059	18	0,0038	29	0,0027
8	0,0058	19	0,0086	30	0,0039
9	0,0073	20	0,0045	31	0,0060
10	0,0042	21	0,0069		
11	0,0066	22	0,0045		

По формуле (80) $\bar{x}_B = 0,005$. В соответствии с формулами (88) и (87) $s_B^2 = 0,000002095$. Величина предельной ошибки выборки d при максимальном отклонении в 5% $d = 0,00025$.

Тогда, минимальный объем выборки для получения значения характеристики $M[trq]$ при заданной надежности $g = 0,999$ и точности 5% рассчитывается на основе формулы (82) как:

$$n = \frac{3,2^2 \cdot 0,000002095}{(0,00025)^2} \approx 343. \quad (92)$$

Полученный объем выборки говорит о том, что при обработке 343 значений времени обработки поискового запроса на резервном сервере trq данная характеристика приближенно равна её МО $M[trq]$.

В среднем каждую минуту в РБД предприятия ГПК "ШахтИнвестКузбасс" обработка поисковых запросов на резервном сервере исчисляется десятками. Тогда, для получения необходимого объема выборки следует проанализировать таблицы транзакций менее чем за час функционирования РБД, что приемлемо.

В таблице 14 представлены значения времени обработки запроса на обновление на главном сервере tgu для РБД предприятия ГПК "ШахтИнвестКузбасс" (31 значение).

Таблица 14 – Начальная выборка значений tgu

№	Значение	№	Значение	№	Значение
1	0,0072	12	0,0058	23	0,0088
2	0,0046	13	0,0057	24	0,0074
3	0,0117	14	0,0054	25	0,0073
4	0,0058	15	0,0067	26	0,0062
5	0,0108	16	0,0064	27	0,0095
6	0,0092	17	0,0069	28	0,0084
7	0,0053	18	0,0041	29	0,0119
8	0,0063	19	0,0076	30	0,0068
9	0,0086	20	0,0071	31	0,0092
10	0,0064	21	0,0072		
11	0,0082	22	0,0084		

По формуле (80) $\bar{x}_B = 0,0075$. В соответствии с формулами (88) и (87) $s_B^2 = 0,000003493$. Величина предельной ошибки выборки d при максимальном отклонении в 5% $d = 0,000375$.

Тогда, минимальный объем выборки для получения значения характеристики $M[tgu]$ при заданной надежности $g = 0,999$ и точности 5% рассчитывается на основе формулы (82) как:

$$n = \frac{3,2^2 \cdot 0,000003493}{(0,000375)^2} \approx 254. \quad (93)$$

Полученный объем выборки говорит о том, что при обработке 254 значений времени обработки запроса на обновление на главном сервере tgu данная характеристика приближенно равна её МО $M[tgu]$.

В среднем каждую минуту в РБД предприятия ГПК "ШахтИнвестКузбасс" обработка запросов на обновление на главном сервере исчисляется десятками. Тогда, для получения необходимого объема выборки следует проанализировать таблицы транзакций менее чем за час функционирования РБД, что приемлемо.

В таблице 15 представлены значения времени обработки запроса на обновление на резервном сервере tru для РБД предприятия ГПК "ШахтИнвестКузбасс" (31 значение).

Таблица 15 – Начальная выборка значений tru

№	Значение	№	Значение	№	Значение
1	0,0083	12	0,0091	23	0,0081
2	0,0087	13	0,0050	24	0,0061
3	0,0047	14	0,0107	25	0,0079
4	0,0076	15	0,0093	26	0,0101
5	0,0054	16	0,0092	27	0,0106
6	0,0086	17	0,0089	28	0,0111
7	0,0083	18	0,0114	29	0,0072
8	0,0088	19	0,0103	30	0,0084
9	0,0090	20	0,0108	31	0,0105
10	0,0085	21	0,0087		
11	0,0090	22	0,0111		

По формуле (80) $\bar{x}_B = 0,0088$. В соответствии с формулами (88) и (87) $s_B^2 = 0,000002933$. Величина предельной ошибки выборки d при максимальном отклонении в 5% $d = 0,00044$.

Тогда, минимальный объем выборки для получения значения характеристики $M[tru]$ при заданной надежности $g = 0,999$ и точности 5% рассчитывается на основе формулы (82) как:

$$n = \frac{3,2^2 \cdot 0,000002933}{(0,00044)^2} \approx 155. \quad (94)$$

Полученный объем выборки говорит о том, что при обработке 155 значений времени обработки запроса на обновление на резервном сервере tru данная характеристика приближенно равна её МО $M[tru]$.

В среднем каждую минуту в РБД предприятия ГПК "ШахтИнвестКузбасс" обработка запросов на обновление на резервном сервере исчисляется единицами. Тогда, для получения необходимого объема выборки следует проанализировать таблицы транзакций за несколько часов функционирования РБД, что приемлемо.

В таблице 16 представлены значения времени отправки сообщения для обновления БД резервного сервера $trgu$ РБД предприятия ГПК "ШахтИнвестКузбасс" (31 значение).

Таблица 16 – Начальная выборка значений $trgu$

№	Значение	№	Значение	№	Значение
1	0,0017	12	0,0014	23	0,0022
2	0,0006	13	0,0017	24	0,0018
3	0,0021	14	0,0011	25	0,0017
4	0,0022	15	0,0011	26	0,0014
5	0,0021	16	0,0014	27	0,0019
6	0,0022	17	0,0013	28	0,0028
7	0,0013	18	0,0015	29	0,0013
8	0,0017	19	0,0020	30	0,0012
9	0,0024	20	0,0015	31	0,0016
10	0,0010	21	0,0016		
11	0,0012	22	0,0015		

По формуле (80) $\bar{x}_B = 0,0016$. В соответствии с формулами (88) и (87) $s_B^2 = 0,000000207$. Величина предельной ошибки выборки d при максимальном отклонении в 5% $d = 0,00008$.

Тогда, минимальный объем выборки для получения значения характеристики $M[trgu]$ при заданной надежности $g = 0,999$ и точности 5% рассчитывается на основе формулы (82) как:

$$n = \frac{3,2^2 \cdot 0,000000207}{(0,00008)^2} \approx 331. \quad (95)$$

Полученный объем выборки говорит о том, что при обработке 331 значений времени отправки сообщения для обновления БД резервного сервера $trgu$ данная характеристика приближенно равна её МО $M[trgu]$.

В среднем каждую минуту в РБД предприятия ГПК "ШахтИнвестКузбасс" отправка сообщений для обновления БД резервных серверов исчисляется сотнями. Тогда, для получения необходимого объема выборки следует проанализировать статистические данные меньше чем за минуту функционирования РБД, что приемлемо.

В таблице 17 представлены значения времени передачи запроса с резервного сервера на главный $trgq$ в РБД предприятия ГПК "ШахтИнвестКузбасс" (31 значение).

Таблица 17 – Начальная выборка значений $trgq$

№	Значение	№	Значение	№	Значение
1	0,0094	12	0,0080	23	0,0115
2	0,0084	13	0,0094	24	0,0096
3	0,0099	14	0,0087	25	0,0095
4	0,0110	15	0,0089	26	0,0092
5	0,0109	16	0,0086	27	0,0117
6	0,0114	17	0,0091	28	0,0121
7	0,0075	18	0,0093	29	0,0121
8	0,0095	19	0,0098	30	0,0090
9	0,0108	20	0,0098	31	0,0114
10	0,0086	21	0,0094		
11	0,0090	22	0,0093		

По формуле (80) $\bar{x}_B = 0,0098$. В соответствии с формулами (88) и (87) $s_B^2 = 0,000001417$. Величина предельной ошибки выборки d при максимальном отклонении в 5% $d = 0,00049$.

Тогда, минимальный объем выборки для получения значения характеристики $M[trgq]$ при заданной надежности $g = 0,999$ и точности 5% рассчитывается на основе формулы (82) как:

$$n = \frac{3,2^2 \cdot 0,000001417}{(0,00049)^2} \approx 60. \quad (96)$$

Полученный объем выборки говорит о том, что при обработке 60 значений времени передачи запроса с резервного сервера на главный $trgq$ данная характеристика приближенно равна её МО $M[trgq]$.

В среднем каждую минуту в РБД предприятия ГПК "ШахтИнвестКузбасс" передача запросов с резервного сервера на главный исчисляется десятками. Тогда, для получения необходимого объема выборки следует проанализировать таблицы транзакций за несколько минут функционирования РБД, что приемлемо.

В таблице 18 представлены значения времени передачи отклика на запрос с главного сервера на резервный $tgrqotkl$ в РБД предприятия ГПК "ШахтИнвестКузбасс" (31 значение).

Таблица 18 – Начальная выборка значений $tgrqotkl$

№	Значение	№	Значение	№	Значение
1	0,0394	12	0,0340	23	0,0370
2	0,0344	13	0,0439	24	0,0356
3	0,0419	14	0,0347	25	0,0355
4	0,0370	15	0,0349	26	0,0352
5	0,0369	16	0,0296	27	0,0377
6	0,0374	17	0,0351	28	0,0363
7	0,0335	18	0,0353	29	0,0381
8	0,0397	19	0,0358	30	0,0350
9	0,0368	20	0,0353	31	0,0374
10	0,0346	21	0,0319		
11	0,0350	22	0,0353		

По формуле (80) $\bar{x}_B = 0,0361$. В соответствии с формулами (88) и (87) $s_B^2 = 0,000006987$.

Величина предельной ошибки выборки d при максимальном отклонении в 5% $d = 0,0018$.

Тогда, минимальный объем выборки для получения значения характеристики $M[tgrqotkl]$ при заданной надежности $g = 0,999$ и точности 5% рассчитывается на основе формулы (82) как:

$$n = \frac{3,2^2 \cdot 0,000006987}{(0,0018)^2} \approx 22. \quad (97)$$

Полученный объем выборки говорит о том, что при обработке 22 значений времени передачи отклика на запрос с главного сервера на резервный $tgrqotkl$ данная характеристика приближенно равна её МО $M[tgrqotkl]$.

В среднем каждую минуту в РБД предприятия ГПК "ШахтИнвестКузбасс" передача отклика на запрос с главного сервера на резервный исчисляется десятками. Тогда, для получения необходимого объема выборки следует проанализировать таблицы транзакций меньше чем за минуту функционирования РБД, что приемлемо.

4.3. Этап вычисления оптимальной загруженности резервного узла и выбора фрагментов данных для немедленной репликации

Выбор фрагментов данных для немедленной репликации выполняется на основе алгоритма, схема которого представлена в главе 3, рисунок 23. В качестве исходных данных алгоритма выступают результаты расчетов в соответствии с алгоритмом вычисления оптимальной загруженности резервного узла РБД при репликации $(I'_{q_{opt}}, I'_{u_{opt}})$, а также характеристики фрагментов реплик РБД.

При этом степень фрагментации РБД выбирается исходя из допустимой сложности решения задачи выбора фрагментов данных для немедленной репликации и возможностей конкретной СУБД [50].

В общем виде выделяют три уровня фрагментации РБД [92]:

1. Уровень таблиц.
2. Уровень атрибутов/строк.
3. Уровень отдельных записей.

При выборе конкретных фрагментов данных должны учитываться требования [48]:

1. Полнота. Когда экземпляр отношения R разбивается на фрагменты (R_1, R_2, \dots, R_n) , каждый элемент данных, присутствующий в отношении, должен содержаться, как минимум, в одном из созданных фрагментов. Выполнение этого правила гарантирует, что никакие данные не будут потеряны после фрагментации.

2. Восстанавливаемость. Должна применяться операция реляционной алгебры, способная восстановить первоначальное отношение R из его фрагментов. Выполнение этого правила гарантирует сохранение функциональных зависимостей.

3. Непересекаемость. Если элемент данных d_i присутствует во фрагменте R_i , то он не должен одновременно быть представлен в любом другом фрагменте. Выполнение данного правила гарантирует минимальную избыточность данных.

Логическая модель данных РБД предприятия ГПК "ШахтИнестКузбас", выполненная с помощью CASE-средства Erwin, представлена на рисунке 27 [59].



Рисунок 27 – Логическая модель данных РБД предприятия ГПК "ШахтИнестКузбас"

В настоящее время производительность типовой ЭВМ, полученная на основе тестов Linpack для измерения производительности компьютеров, составляет порядка 40 Gflops [5, 66]. В соответствии со стандартом IEEE 754 – 2008 величина 40 Gflops представляет собой $4 \cdot 10^{10}$ элементарных операций с плавающей точкой в секунду [104]. При этом операции с плавающей точкой являются наиболее трудоемкими и определяют производительность ЭВМ для самого пессимистичного варианта решения задачи. На рисунке 28 представлен график зависимости вычислительной сложности алгоритма выбора фрагментов данных для немедленной репликации от размера входных данных (количества оцениваемых фрагментов).

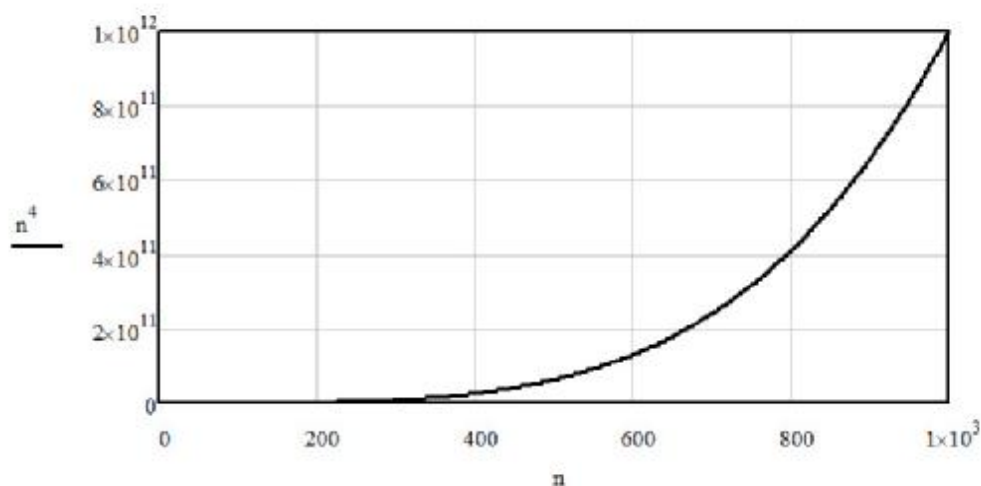


Рисунок 28 – Вычислительная сложность алгоритма выбора фрагментов данных

Видно, что с ростом размера входных данных вычислительная сложность алгоритма резко увеличивается. При $n=1000$ число операций с плавающей точкой для получения конечного результата алгоритма составляет порядка 10^{12} . Такое количество операций современная ЭВМ с производительностью 40 Gflops выполняет за время ≈ 25 секунд. Дальнейшее увеличение размерности задачи приведет к значительному росту времени ожидания выполнения алгоритма. Таким образом, в условиях функционирования РБД предприятия ГПК "ШахтИнвестКузбасс" для применения представленного алгоритма автоматизации процесса конфигурирования репликации в РБД предприятия ГПК с учетом минимума среднего времени её отклика на запросы целесообразно ограничится числом фрагментов не превышающим 1000 [35]. Такое количество фрагментов реплик достигается в рамках стратегии фрагментации на уровне снимков.

4.3.1 Пример выбора фрагментов данных для немедленной репликации

Совокупность исходных данных, полученных от РБД предприятия ГПК "ШахтИнвестКузбасс", для реализации этапа выбора фрагментов данных для немедленной репликации представлена в таблице 19.

Таблица 19 – Характеристики РБД предприятия ГПК «ШахтИнвестКузбасс»

Наименование характеристики	Значение	Размерность
$M[tgq]$	0,000049	секунд
$M[trq]$	0,002984	секунд
$M[tgu]$	0,000012	секунд
$M[tru]$	0,003192	секунд
$M[tgru]$	0,009831	секунд
$M[trgu]$	0,00112	секунд
$M[trgq]$	0,002839	секунд
$M[tgrqotkl]$	0,000765	секунд
I_u	39	запросов в минуту
I_q	412	запросов в минуту
n	43	безразмерная (серверы)

На основе программы "ctrlReplic" [100], реализующей решение задачи в соответствии с алгоритмом вычисления допустимых оптимальных значений параметров РБД при репликации, получены следующие результаты: $I'q^{opt} = 214,31$ запросов в минуту, $I'u^{opt} = 2,07$ запроса в минуту.

Совокупность характеристик фрагментов РБД предприятия ГПК "ШахтИнвестКузбасс" с учетом детализации до 1000 элементов для 15 фрагментов представлена в таблице 20. Полная таблица характеристик фрагментов приведена в приложении А (таблица А.1).

Таблица 20 – Характеристики фрагментов

Номер фрагмента	Размер, байт	Интенсивность поисковых запросов в минуту	Интенсивность запросов на обновление в минуту
1	66197	0,3875	0,0201
2	79828	0,3454	0,0557
3	73997	0,6153	0,0495
4	63414	0,3738	0,0417
5	83023	0,6472	0,0560
6	34731	0,6697	0,0440
7	90672	0,3572	0,0378
8	46684	0,4306	0,0523
9	10997	0,3491	0,0331
10	54419	0,2879	0,0424
11	59833	0,3079	0,0268
12	27277	0,2295	0,0544
13	36742	0,2971	0,0354
14	99398	0,5155	0,0207
15	23904	0,5909	0,0593

С помощью программы "ctrlReplic" [100] получен вариант набора фрагментов данных для немедленной репликации в условиях допустимого отклонения $d = 5\%$ от оптимальных значений: $I'q^{opt} = 214,31$ запросов в минуту и $I'u^{opt} = 2,07$ запроса в минуту.

При данном варианте набора фрагментов данных для немедленной репликации на узлах РБД предприятия ГПК "ШахтИнвестКузбасс" среднее время отклика на один запрос составляет $\bar{T} = 0,223$ секунды, тогда как

существующее среднее время отклика $\bar{T}_{\text{реальное}} = 0,239$ секунды. Полученный эффект от применения представленного алгоритма в РБД предприятия ГПК «ШахтИнвестКузбасс» рассчитывается по формуле [37]:

$$\Delta = \frac{\bar{T}_{\text{реальное}} - \bar{T}}{\bar{T}_{\text{реальное}}} \cdot 100\% = 6,69\%. \quad (98)$$

При реагировании на аварийные ситуации функционирования предприятия ГПК, отличающимся значительным всплеском интенсивности поисковых запросов, данный выигрыш обеспечит запас времени для принятия обоснованных решений по оперативному управлению.

На основе алгоритма автоматизации процесса конфигурирования репликации РБД предприятия ГПК получены значения интенсивностей поисковых и запросов на обновление при изменении одного из параметров во всем рабочем диапазоне значений, при которых среднее время отклика РБД на запросы минимально, рисунок 29, 30.

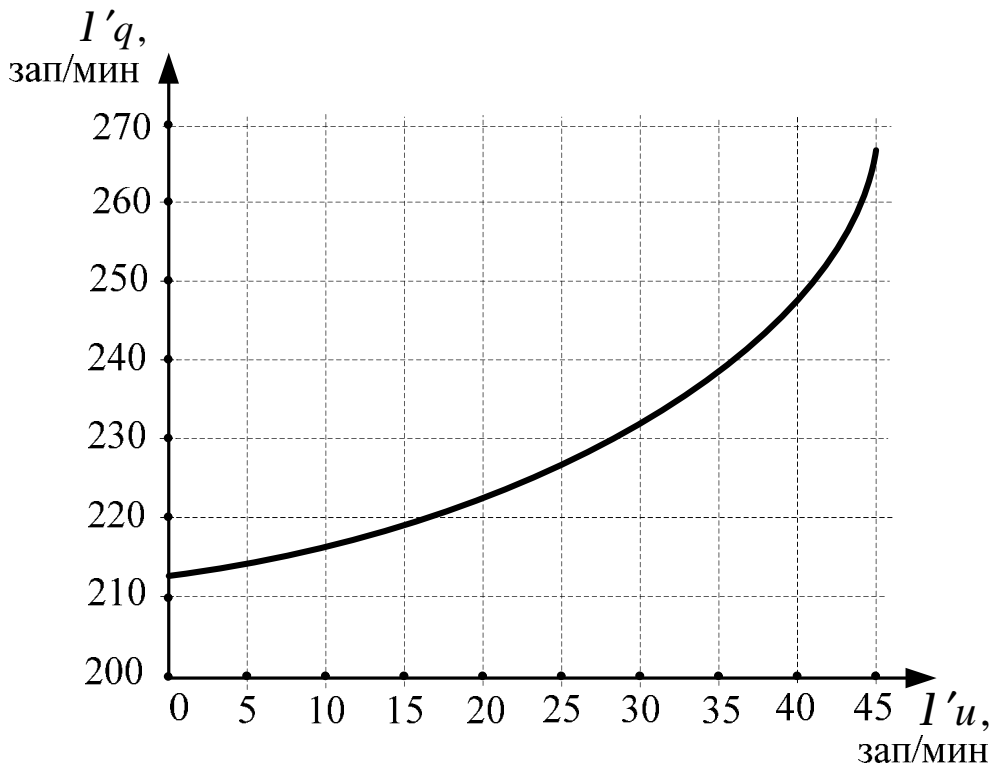


Рисунок 29 – График зависимости интенсивности поисковых запросов от изменения интенсивности запросов на обновление во всем рабочем диапазоне значений, при которых среднее время отклика РБД на запросы минимально

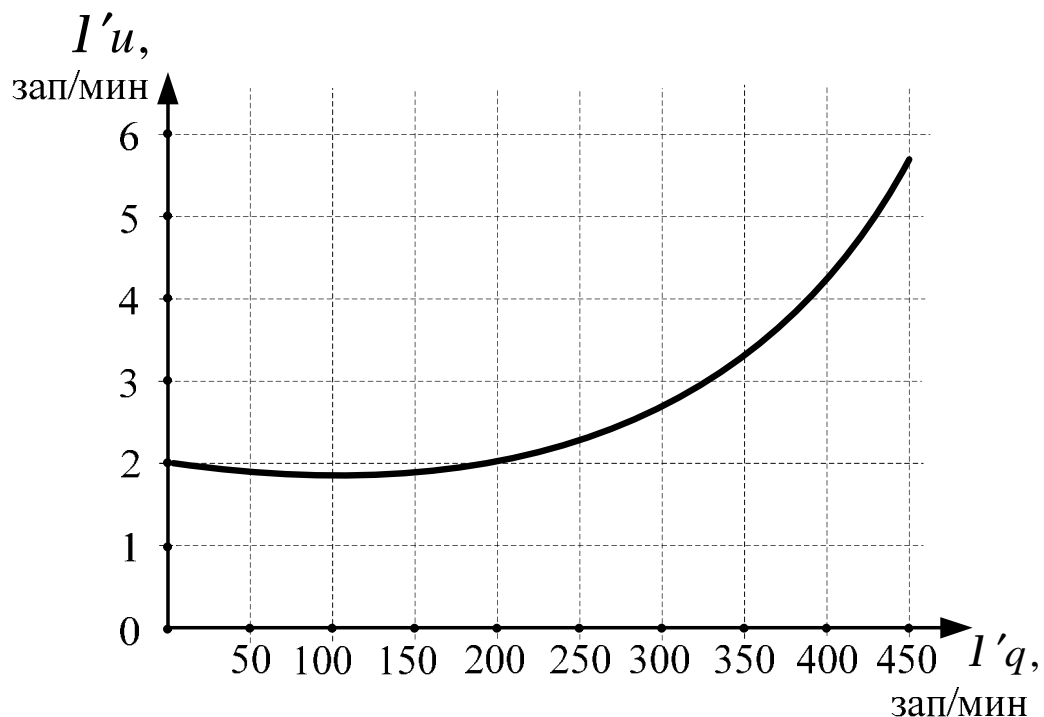


Рисунок 30 – График зависимости интенсивности запросов на обновление от изменения интенсивности поисковых запросов во всем рабочем диапазоне значений, при которых среднее время отклика РБД на запросы минимально

4.4 Этап формирования рекомендаций для администратора РБД предприятия ГПК по репликации

Решения для администратора РБД предприятия ГПК по конфигурированию репликации представляют собой наборы фрагментов данных для немедленной репликации.

Шаблоны SQL-запросов для выборки характеристик РБД из журнала транзакций и репликации выбранных фрагментов по резервным узлам имеют обобщенный вид [16, 62]:

```
select * from STAT_TABLE
where ID > (select count(*) from STAT_TABLE)-n
and DATE >= start_date
and ADDRESS in (SITE_2, SITE_3, ... , SITE_N);
```

```
insert into REPLICA from SITE_GL (ID_fragmenta)
```

```
insert into TABLE from SITE_i
```

```
select * from TABLE at SITE_GL
```

```
where ID_fragmenta in (select * from REPLICA)
```

Формирование решений для администратора РБД предприятия ГПК по репликации реализуется на основе программного модуля, схема которого в общем виде представлена на рисунке 31.

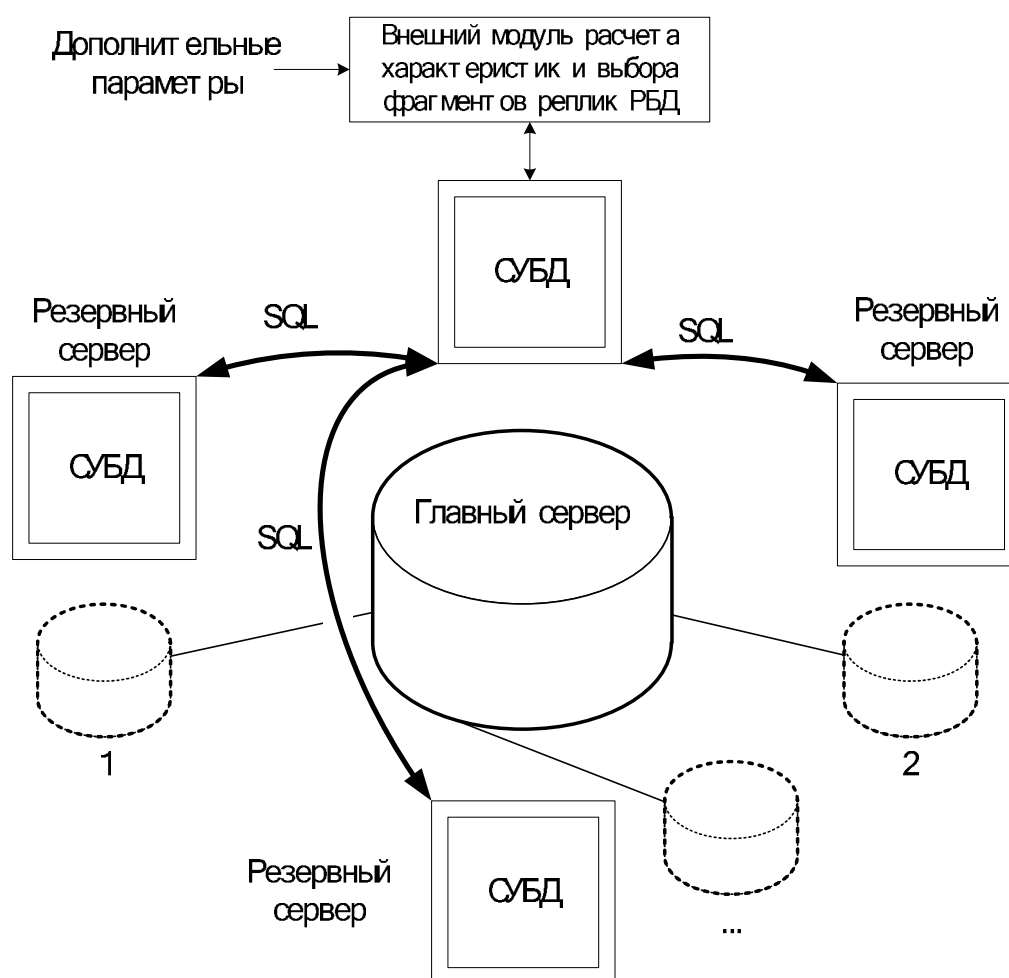


Рисунок 31 – Схема взаимодействия программного модуля с РБД

Представленный технологический алгоритм автоматизации процесса конфигурирования репликации РБД предприятия ГПК является корректным,

устойчивым, точным и обладает полиномиальной сложностью, так как состоит из этапов и алгоритмов, для которых эти свойства доказаны:

1. Вычисление значений характеристик РБД.
2. Вычисление допустимых оптимальных значений параметров РБД при репликации на основе алгоритма вычисления оптимальной загруженности резервного узла.
3. Обоснованный выбор фрагментов данных для немедленной репликации на основе алгоритма выбора фрагментов данных для немедленной репликации.
4. Формирование рекомендаций для администратора РБД предприятия ГПК по репликации.

Выводы по главе 4

1. Алгоритм автоматизации процесса конфигурирования репликации в РБД предприятия ГПК состоит из четырех основных этапов: вычисления значений характеристик РБД, вычисления допустимых оптимальных значений параметров РБД при репликации, обоснованного выбора фрагментов данных для немедленной репликации, формирования рекомендаций для администратора РБД предприятия ГПК по репликации. Алгоритм является корректным, устойчивым, точным и обладает полиномиальной сложностью

2. Аналитически доказано, что получение значений количественных признаков РБД предприятия ГПК "ШахтИнвестКузбасс" с учетом обеспечения требуемой точности и надежности, выполнимо за допустимое время. При этом в качестве значения надежности выбран наиболее трудоемкий вариант, применяемый на практике, $g = 0,999$.

3. Показано, что для реализации представленного алгоритма автоматизации процесса конфигурирования репликации в РБД предприятия ГПК с учетом минимума среднего времени её отклика на запросы необходимо ограничиться числом фрагментов не превышающим 1000, так как последующая детализация выводит время работы алгоритма за допустимые пределы.

4. Применение представленного алгоритма на примере РБД предприятия ГПК "ШахтИнвестКузбасс" позволило получить выигрыш по среднему времени отклика на запросы в 6,69 %. При реагировании на аварийные ситуации функционирования предприятия ГПК, отличающимся значительным всплеском интенсивности поисковых запросов, данный выигрыш обеспечит запас времени для принятия обоснованных решений по оперативному управлению.

5. Реализация представленного алгоритма возможна на основе программного модуля, получающего данные в рамках выполнения SQL-запросов к РБД.

ЗАКЛЮЧЕНИЕ

В диссертационной работе решена актуальная научно-техническая задача разработки модели и алгоритмов управления параметрами репликации в РБД предприятия ГПК, позволяющих снизить время отклика РБД предприятия ГПК на запросы.

В рамках проведенных исследования получены следующие основные результаты:

1. Разработана математическая модель отклика РБД на запросы при репликации, базирующаяся на модели двухуровневой информационной системы с репликацией данных, отличающаяся учетом совокупности параметров: интенсивности запросов на обновление ($I'u$) и интенсивности поисковых запросов ($I'q$), обрабатываемых на резервных серверах, на уровне физической интерпретации. Проверка модели на основе сравнения модельного и времени обработки запросов, полученного в условиях производства, выявила достаточную адекватность и точность заявленной модели.

2. На основе математической модели разработаны алгоритмы вычисления оптимальной загруженности резервного узла распределенной базы данных при репликации и выбора фрагментов данных для немедленной репликации.

Алгоритм вычисления оптимальной загруженности резервного узла при репликации в РБД, описываемой математической моделью отклика на запросы, основанный на модифицированном методе линейных комбинаций, отличающийся формированием ограничений, обеспечивающих режим функционирования РБД предприятия ГПК без блокировки, позволяет определять значения параметров репликации, при которых достигается снижение среднего времени отклика РБД на запросы. Алгоритм является корректным, обладает достаточной точностью, вычислительной устойчивостью и допустимой сложностью.

Алгоритм выбора фрагментов данных для немедленной репликации, основанный на оптимизированном методе частично-целочисленного

линейного программирования с аддитивным алгоритмом для задач с двоичными переменными, отличающийся процедурой принятия решения по критерию минимума объема пересылаемых реплик. Алгоритм позволяет находить решения за полиномиальное время с учетом снижения точности вычислений не более чем на 5 %, является корректным, вычислительно устойчивым и обладает допустимой сложностью.

3. Способ управления репликацией в РБД, защищенный патентом на изобретение, реализован в виде алгоритма автоматизации процесса конфигурирования репликации в РБД предприятия ГПК, отличающегося обеспечением условий функционирования РБД предприятия ГПК с пониженным средним временем отклика на запросы, позволяющего вычислять значения параметров репликации в РБД предприятия ГПК и формировать решения для администратора по её конфигурированию. Оценка свойств алгоритма установила его корректность, устойчивость, полиномиальную сложность и достаточную точность.

Установлено, что для применения представленного алгоритма необходимо ограничиться числом фрагментов не превышающим 1000, так как последующая детализация выводит время реализации алгоритма за допустимые пределы в соответствии с современными требованиями к производительности рабочего места администратора РБД предприятия ГПК "ШахтИнвестКузбасс".

5. Применение представленных алгоритмов на примере РБД предприятия ГПК "ШахтИнвестКузбасс" позволило получить выигрыш по среднему времени отклика на запросы в 6,69 % по сравнению со штатным функционированием системы.

Научная новизна:

1. Математическая модель отклика РБД на запросы при репликации, базирующаяся на модели двухуровневой информационной системы с репликацией данных, отличающаяся учетом совокупности параметров: интенсивности запросов на обновление ($I'u$) и интенсивности поисковых

запросов ($l'q$), обрабатываемых на резервных серверах, на уровне физической интерпретации.

2. Алгоритм вычисления оптимальной загруженности резервного узла при репликации в РБД, описываемой математической моделью отклика на запросы, основанный на модифицированном методе линейных комбинаций, отличающийся формированием ограничений, обеспечивающих режим функционирования РБД предприятия ГПК без блокировки.

3. Алгоритм выбора фрагментов данных для немедленной репликации, основанный на оптимизированном методе частично-целочисленного линейного программирования с аддитивным алгоритмом для задач с двоичными переменными, отличающийся процедурой принятия решения по критерию минимума объема пересылаемых реплик.

4. Способ управления репликацией в РБД, основанный на гибридном методе репликации, отличающийся автоматизацией подготовки принятия решения по управлению репликацией, защищенный патентом на изобретение.

Теоретические положения и созданные на их основе модель и алгоритмы разработаны с использованием теоретических и экспериментальных методов исследования. Решение задачи, поставленной в работе, стало возможным благодаря известным достижениям таких научных дисциплин, как математический анализ, математическая статистика, теория вероятностей, теория оптимизации и планирование эксперимента и не противоречит их положениям.

Теоретическая значимость полученных решений заключается в разработке нового гибридного метода репликации, позволяющего за счет управления параметрами репликации в РБД предприятия ГПК подстраиваться под имеющиеся вычислительные и сетевые ресурсы с целью повышения её реактивности.

Практическая значимость заключается в разработке совокупности алгоритмов и доведении их до программной реализации, что подтверждается свидетельствами о государственной регистрации программ для ЭВМ

№ 2013611771 от 4 февраля 2013 года и № 2013616315 от 19 июня 2013 года, патентом на полезную модель № 126161 от 20 марта 2013 года и изобретением (положительное решение от 25.10.2013 о выдаче патента на изобретение "Способ репликации информации в распределенных базах данных с конкурентным распределением потоков" по заявке № 2012116021).

Полученные результаты могут использоваться на предприятии ГПК с целью эффективной организации специализированного информационного обеспечения, создающей условия для снижения среднего времени отклика на запросы при заданных ограничениях на временные задержки обработки запросов в различных элементах РБД.

Применение представленных в работе модели и алгоритмов позволило снизить среднее время отклика РБД предприятия ГПК «ШахтИнвестКузбасс» на запросы, что подтверждено данными, полученными в условиях производства.

Направлением дальнейших исследований является: разработка теоретической базы для создания универсальной методики по настройке РБД при репликации, создание условий для её технической реализации.

Список использованных источников

1. **Аленичев, В. М.** Концепция создания справочно-информационных систем горнопромышленного комплекса / В. М. Аленичев, С. В. Корнилков, В. И. Суханов, М. А. Акоев // Горный информационно-аналитический бюллетень (научно-технический журнал). – 2009. – Том 2. – № 12. – С. 36–48.
2. **Амосов, А. А.** Вычислительные методы для инженеров : учебное пособие / А. А. Амосов, Ю. А. Дубинский, Н. В. Копченова. – М. : Высш. шк., 1994. – 544 с.
3. **Апанасевич, Д. А.** Математическое и программное обеспечение асинхронной репликации данных реляционных СУБД методом выделения объектов: дис. ... канд. тех. наук : 05.13.11 / Апанасевич Дмитрий Александрович. – Воронеж, 2008. – 120 с.
4. **Ахо, А.** Структуры данных и алгоритмы [пер. с англ.] / А. Ахо, В. Хопкрофт, Д. Ульман, Д. Джеффри. – М. : Вильямс, 2001. – 384 с.
5. **Баденко, В. Л.** Высокопроизводительные вычисления : учеб. пособие / В. Л. Баденко. – СПб. : Изд-во Политехн. ун-та, 2010. – 180 с.
6. **Баранова, И. В.** Управление предприятием на основе интегрированных средств поддержки распределённых баз данных / И. В. Баранова // Вестник южно-российского государственного технического университета Новочеркасского политехнического института. – 2013. – № 1. – С. 110–118.
7. **Бейбер, Р. Л.** Программное обеспечение без ошибок [пер. с англ.] / Р. Л. Бейбер. – М : Радио и связь, 1996. – 176 с.
8. **Белоусов, В. Е.** Алгоритмы репликации данных в распределенных системах обработки: дис. ... канд. тех. наук : 05.13.01 / Белоусов Всеволод Евгеньевич. – Пенза, 2005. – 164 с.
9. **Блохин, В. Г.** Современный эксперимент: подготовка, проведение, анализ результатов / В. Г. Блохин, О. П. Глудкин, А. И. Гуров, М. А. Ханин. – М. : Радио и связь, 1997. – 232 с.

10. **Бойко, А. А.** Система показателей качества баз данных автоматизированных систем / А. А. Бойко, С. А. Гриценко, В. Ю. Храмов // Вестник ВГУ, Серия: Системный анализ и информационные технологии. – 2010. – № 1. – С 39–45.
11. **Бочков, М. В.** Проектирование автоматизированных систем обработки информации и управления / М. В. Бочков, Е. И. Новиков, О. В. Тараканов; под ред. М. В. Бочкова. – Орел: Академия ФСО России, 2007. – 406 с.
12. **Бочков, М. И.** Имитационное моделирование информационно-телекоммуникационных систем / М. В. Бочков, Б. И. Соловьев. – Орел: Академия ФСО России, 2008. – 191 с.
13. **Бронштейн, О. И.** Модели приоритетного обслуживания / О. И. Бронштейн, И. М. Духовный. – М. : Наука, 1976. – 223 с.
14. **Брюханов, В. Н.** Теория автоматического управления, учебное издание / В. Н. Брюханов, М. Г. Косов, С. П. Протопопов, Ю. М. Соломенцев, Н. М. Султан-Заде, А. Г. Схиртладзе. – М. : Высш. шк., 2000. – 265 с.
15. **Брюханов, В. Н.** Автоматизация машиностроительного производства / В. Н. Брюханов, А. Г. Схиртладзе, В. П. Вороненко. – М. : ИЦ МГТУ «Станкин», 2003. – 288 с
16. **Бьелетич, Ш.** Microsoft SQL Server 2000. Энциклопедия пользователя [пер. с англ.] / Ш. Бьелетич, Г. Мэйбл и др. – К. : ДиаСофт, 2001. – 688 с.
17. **Васильев, К. К.** Математическое моделирование систем связи : учебное пособие / К. К. Васильев, М. Н. Служивый. – Ульяновск : УлГТУ, 2008. – 170 с.
18. **Вентцель, Е. С.** Исследование операций: задачи, принципы, методология / Е. С. Вентцель. – 4-е изд. – М. : Дрофа, 2006. – 206 с.
19. **Вентцель, Е. С.** Теория вероятностей и ее инженерные приложения / Е. С. Вентцель, Л. А. Овчаров. – М. : Высш. шк., 2000. – 480 с.
20. **Гмурман, В. Е.** Теория вероятностей и математическая статистика / В. Е. Гмурман. – 12-е изд. – М. : Высшее образование, 2008. – 479 с.

21. **ГОСТ 19.701-90.** Единая система программной документации. Схемы алгоритмов, программ, данных и систем. Условные обозначения и правила выполнения. – Введ. 1992.01.01. – М. : ФГУП "Стандартинформ", 2005. – 12 с. – (Межгосударственный стандарт).
22. **Григорьев, Ю. А.** Организация базы данных в программном комплексе анализа характеристик производительности распределённых систем обработки данных / Ю. А. Григорьев // Наука и образование: электронное научно-техническое издание. – 2012. – № 2. – 39 с.
23. **Григорьев, Ю. А.** Оценка времени выполнения SQL-запросов к базам данных / Ю. А. Григорьев // Наука и образование: электронное научно-техническое издание. – 2012. – № 1. – 30 с.
24. **Грилл, Ф.** Практическая оптимизация [пер. с англ.] / Ф. Грилл, У. Мюррей, М. Райт. – М. : Мир, 1985. – 509 с.
25. **Гришмановский, П. В.** Анализ технологий репликации данных и методы повышения эффективности разрешения конфликтов репликации / П. В. Гришмановский, Е. В. Базилевский // Вестник волжского университета им. В. Н. Татищева. – 2012. – № 2(19). – С. 98–106.
26. **Грофф, Дж.** SQL: Полное руководство : [пер. с англ.] / Дж. Грофф, П. Вайнберг. – 2-е изд. – К. : Издательская группа BHV, 2001. – 816 с.
27. **Губинский, А. И.** Информационно-управляющие человеко-машинные системы. Исследование, проектирование, испытания / А. И. Губинский, В. Г. Евграфова – М. : Машиностроение, 1993. – 528 с.
28. **Гультияев, А.** Визуальное моделирование в среде MATLAB: учебный курс / А. Гультияев. – СПб. : Питер, 2000. – 432 с.
29. **Денискина, Е. А.** Статистический анализ данных : метод. указания к расчетной работе / Е. А. Денискина, П. Э. Коломиец. – Самара : СГАУ, 2004. – 64 с.
30. **Дунаев, В. А.** Выбор фрагментов данных для размещения по узлам распределенной базы данных с учетом минимума среднего времени её отклика на запросы / В. А. Дунаев / Системы управления и информационные технологии. – 2013. – № 4(54). – С. 57–60.

31. **Дунаев, В. А.** Оценка времени реакции распределенной базы данных на запросы при гибридном механизме репликации / В. А. Дунаев // Информационные системы и технологии. – 2013. – № 6 (80). – С. 103–113.

32. **Дунаев, В. А.** Математическое моделирование информационного обмена в распределенных базах данных в режиме репликации / В. А. Дунаев, О. В. Тараканов // Информационные технологии моделирования и управления. – 2012. – № 6(78). – С. 458–465

33. **Дунаев, В. А.** Модифицированная модель обработки запросов в распределенных базах данных / В. А. Дунаев, О. В. Тараканов // Программные продукты и системы. – 2014. – № 1(105). – С. 70–76

34. **Дунаев, В. А.** Разработка модели информационного обмена в распределенных базах данных в режиме репликации / В. А. Дунаев, О. В. Тараканов // Системы управления и информационные технологии. – 2012. – № 4.1(50). – С. 192–196

35. **Дунаев, В. А.** Методика выбора фрагментов данных для размещения по узлам распределенной базы данных с учетом минимума среднего времени её отклика на запросы / В. А. Дунаев // Современные проблемы информатизации : сборник статей 19-ой Международной открытой научной конференции. – Yelm, WA, USA : Science Book Publishing House, 2014. – С. 166–169.

36. **Дунаев, В. А.** Особенности управления параметрами репликации распределенной базы данных предприятия горнопромышленного комплекса / В. А. Дунаев, О. В. Тараканов // Информационные системы и технологии. – 2014. – № 2. – С. 45–52.

37. **Дэниел, К.** Применение статистики в промышленном эксперименте [пер. с англ.] / К. Дэниел. – М. : Мир, 1979. – 294 с.

38. **Защита Г.Н. Шаповаленко:** комплексное обоснование системы оперативного контроля рабочих процессов на угольных разрезах // Уголь. – 2012. – № 6. – С. 64–66.

39. **Иванов, А. Ю.** Военно-технические основы построения и математическое моделирование перспективных средств и комплексов автоматизации / А. Ю. Иванов, С. П. Полковников, Г. Б. Ходасеевич. – СПб. : 1997. – 419 с.
40. **Иванов, А. Ю.** Модель для оценки оперативности реализации запросов к распределенным базам данных / А. Ю. Иванов // Проблемы управления рисками в техносфере. Научно-аналитический журнал. – 2008. – № 4(8). – С. 176–183.
41. **Иванов, А. Ю.** Мобильные распределенные базы данных автоматизированных информационно-управляющих систем МЧС России: дис. ... док. тех. наук : 05.25.05 / Иванов Александр Юрьевич. – Санкт-Петербург, 2009. – 277 с.
42. **Измаилов, А. Ф.** Численные методы оптимизации : учеб. пособие / А. Ф. Измаилов, М. В. Солодов. – М. : ФИЗМАТЛИТ, 2005. – 304 с.
43. **Карельская, К. А.** Иерархическая система управления распределенными информационными ресурсами: дис. ... канд. тех. наук : 05.13.01 / Карельская Катерина Александровна. – Тверь, 2006. – 107 с.
44. **Кайт, Т.** Oracle для профессионалов / Т. Кайт. – СПб. : ООО «ДиаСофтЮП», 2003. – 672 с.
45. **Кепещук, Д. Б.** Разработка модели обмена данными в гетерогенных распределенных базах данных / Д. Б. Кепещук // Вестник тюменского государственного университета. – 2006. – № 7. – С. 81–86.
46. **Клейнрок, Л.** Теория массового обслуживания [пер. с англ.] / Л. Клейнрок. – М. : Машиностроение, 1979. – 432 с.
47. **Колемаев, В. А.** Теория вероятностей и математическая статистика : учебное пособие / В. А. Колемаев, О. В. Староверов, В. Б. Турундаевский. – М. : Высш. шк., 1991. – 400 с.
48. **Коннолли, Т.** Базы данных. Проектирование, реализация и сопровождение [пер. с англ.] / Т. Коннолли, К. Бегг, А. Страчан. – 3-е изд. – М. : Издательский дом "Вильямс", 2003. – 1440 с.

49. **Кормен, Т.** Алгоритмы: построение и анализ / Т. Кормен, Ч. Лейзерсон, Р. Ривест. – М : МЦНМО, 1999. – 960 с.
50. **Костычев, Е. А.** Нацеленная генерация данных для тестирования приложений над базами данных / Е. А. Костычев, В. А. Омельченко, С. В. Зеленев // Труды института системного программирования. – М. : Институт системного программирования РАН, 2011. – С. 253–268.
51. **Крутов, В. И.** Основы научных исследований : учебник для вузов / В. И. Крутов, В. В. Попов. – М. : Высш. шк., 1989. – 400 с.
52. **Крылов, В. В.** Теория телетрафика и ее приложения / В. В. Крылов, С. С. Самохвалова. – СПб. : БХВ-Петербург, 2005. – 288 с.
53. **Кузнецов, С. Д.** Основы баз данных : учеб. пособие / С. Д. Кузнецов. – М. : БИНОМ, 2007. – 484 с.
54. **Кукушкин, А. А.** Теоретические основы автоматизированного управления. Часть 1. Основы анализа и оценки сложных систем : пособие / А. А. Кукушкин. – Орел : ВИПС, 1998. – 254 с.
55. **Кульба, В. В.** Теоретические основы проектирования оптимальных структур распределенных баз данных / В. В. Кульба. – М. : СИНТЕГ, 1990. – 660 с.
56. **Кухарев, В. Н.** Модели и алгоритмы распределения реплицированных баз данных в информационных системах: дис. ... канд. тех. наук : 05.13.18 / Кухарев Вадим Николаевич. – Новочеркасск, 2007. – 159 с.
57. **Луни, К.** Oracle 10g настольная книга администратора баз данных / К. Луни, Б. Брил. – М. : Лори, 2008. – 377 с.
58. **Макаров, Е. Г.** Инженерные расчеты в Mathcad. Учебный курс / Е. Г. Макаров. – СПб. : Питер, 2005. – 448 с.
59. **Маклаков, С. В.** ВРwin и ERwin. CASE-средства разработки информационных систем / С. В. Маклаков. – М. : ДИАЛОГ-МИФИ, 1999. – 256 с.
60. **Мейкшан, Л. И.** Анализ двухуровневой информационной системы с репликацией данных / Л. И. Мейкшан // Инфокоммуникационные технологии. – 2009. № 2. – С. 56-60.

61. **Меннинген, Й.** Система определения и обнаружения местоположения людей. Практическое испытание в шахте «Реклингхаузен» / Й. Меннинген, А. Финк // Уголь. – 2012. – № 5. – С 28–32.
62. **Молиаро, Э.** SQL. Сборник рецептов [пер. с англ.] / Э. Молиаро. – СПб. : Символ-Плюс, 2009. – 672 с.
63. **Морозов, Д. В.** Ключевые моменты построения современных СУБД в рамках информационных систем предприятия / Д. В. Морозов, Л. А. Симонова // Социально-экономические и технические системы: исследование, проектирование, оптимизация. . – 2006. – № 10.
64. **Неганова, Л. М.** Статистка: Конспект лекций / Л. М. Неганова. – М. : Издательство Юрайт, 2010. – 220 с.
65. **Олзоева, С. И.** Моделирование и расчет распределенных информационных систем : учебное пособие / С. И. Олзоева. – Улан – Удэ : ВСГТУ, 2004. – 67 с.
66. **Олифер, В. Г.** Компьютерные сети. Принципы, технологии, протоколы / В. Г. Олифер, Н. А. Олифер. – СПб. : Питер, 2001. – 672 с.
67. **Пападимитриу, Х.,** Комбинаторная оптимизация. Алгоритмы и сложность / Х. Пападимитриу, К. Стайглиц. – СПб. : Мир, 1984. – 512 с.
68. **Петухов, Г. Б.** Основы теории эффективности целенаправленных процессов. Часть 1. Методология, методы, модели / Г. Б. Петухов. – М. : МО СССР, 1989. – 660 с.
69. **Пискунов, Н. С.** Дифференциальное и интегральное исчисления. Том 1 : учебное пособие / Н. С. Пискунов. – СПб. : Мифрил, 1996. – 416 с.
70. **Питерсон, Дж.** Теория сетей Петри и моделирование систем [пер. с англ.] / Дж. Питерсон. – М. : Мир, 1984. – 264 с.
71. **Приказ Ростехнадзора** от 20.12.2010 N 1158 "О внесении изменений в Правила безопасности в угольных шахтах, утвержденные Постановлением Госгортехнадзора России от 5 июня 2003 г. № 50.
72. **Приказ** Министерства связи и массовых коммуникаций Российской Федерации (Минкомсвязь России) от 2 сентября 2011 г. № 221 г. Москва "Об утверждении Требований к информационным системам

электронного документооборота федеральных органов исполнительной власти, учитывающих в том числе необходимость обработки посредством данных систем служебной информации ограниченного распространения".

73. **Расчет** функциональной живучести информационных систем с распределенной базой данных при репликации : свидетельство о государственной регистрации программ для ЭВМ № 2013611771 / В. А. Дунаев, О. Ю. Миронов, Н. В. Покусин, Д. О. Кривошея. заявка № 2012661052 от 13.12.2012.

74. **Ригс, С.** Администрирование PostgreSQL 9. Книга рецептов [пер. с англ.] / С. Ригс, Х. Кросинг. – М. : ДМК Пресс, 2012. – 368 с.

75. **Рогачков, А. В.** Применение современных технических средств мониторинга для оценки соответствия проектных параметров анкерной крепи изменяющимся условиям проведения подземных выработок / А. В. Рогачков, А. С. Позолотин, В. Ф. Исамбетов, П. И. Муравский, П. В. Гречишкин // Уголь. – 2012. – № 12. – С 38–42.

76. **Ролланд, Ф. Д.** Основные концепции баз данных [пер. с англ.] / Ф. Д. Ролланд. – М. : Вильямс, 2002. – 256 с.

77. **Рябчинский, А. С.** Искробезопасные барьеры МАСХ Ех для приложений с высокими требованиями к функциональной безопасности / А. С. Рябчинский // Уголь. – 2011. – № 5. – С 66–67.

78. **Самарский, А. А.** Математическое моделирование: Идеи. Методы. Примеры / А. А. Самарский, А. П. Михайлов. – М. : Наука. Физматлит, 1997. – 320 с.

79. **Сергеев, И. В.** Программное и математическое обеспечение системы репликации данных СУБД независимых платформ: дис. ... канд. тех. наук : 05.13.11 / Сергеев Иван Викторович. – Москва, 2003. – 112 с.

80. **Система** децентрализованного управления структурой распределенной базы данных : пат. на полезную модель № 126161 Рос. Федерация : МПК⁸ G 06 F 12/00 / [В. А. Дунаев, Е. В. Лебедеко и др.] ; патентообладатель Гос. казенное образовательное учреждение высш. проф.

образования Академия ФСО России. – №2012116023 ; заявл. 19.04.2012, опубл. 20.03.2013.

81. **Ситников, А. А.** Моделирование процесса репликации данных в СУБД по сокращенному журналу / А. А. Ситников // Организация баз данных. – 2005. – № 2 (10).

82. **Собилёв, В. Д.** Базы данных : учебное пособие / В. Д. Собилёв. – Томск. : Томский государственный университет систем управления и радиоэлектроники, 2007. – 256 с.

83. **Советов, Б. Я.** Моделирование систем : учебник для бакалавров / Б. Я. Советов, С. А. Яковлев. – 7-е изд. – М. : Издательство Юрайт, 2012. – 343 с.

84. **Сухарев, А. Г.** Курс методов оптимизации : учеб. пособие / А. Г. Сухарев, А. В. Тимохов, В. В. Федоров. – 2-е изд. – М. : ФИЗМАТЛИТ, 2005. – 368 с.

85. **Таненбаум, Э.** Распределенные системы. Принципы и Парадигм / Э. Таненбаум, М. ванн Стеен. – СПб. : Питер, 2003. – 877 с.

86. **Таха, Х. А.** Введение в исследование операций / Х. А. Таха. – 6-е изд. – М. : Издательский дом "Вильямс", 2001. – 912 с.

87. **Уолрэнд, Дж.** Введение в теорию сетей массового обслуживания [пер. с англ.] / Дж. Уолрэнд. – М. : Мир, 1993. – 336 с.

88. **Уорсли, Дж.** PostgreSQL. Для профессионалов [пер. с англ.] / Дж. Уорсли, Дж. Дрейк. – СПб. : Питер, 2003. – 496 с.

89. **Федотов, А. В.** Автоматизация управления в производственных системах : учеб. пособие / А. В. Федотов. – Омск. : ОмГТУ, 2001. – 668 с.

90. **Федотов, Н. Н.** Средства информационного обеспечения автоматизированных систем управления / Н. Н. Федотов, Л. Б. Венчковский. – М. : Издательство стандартов, 1989. – 192 с.

91. **Хинчин, А. Я.** Работы по математической теории массового обслуживания / А. Я. Хинчин. – М. : Физматгиз, 1963. – 236 с.

92. **Черноусова, А. М.** Создание и использование баз данных : учебное пособие / А. М. Черноусова. – Оренбург : ГОУ ОГУ, 2009. – 244 с.

93. **Щекотихин, В. М.** Прикладная математика: Учебное пособие / В. М. Щекотихин, В. М. Терентьев – Орел : Академия ФАПСи, 2002. – 209 с.
94. **Шекхар, Ш.** Основы построения баз данных [пер. с англ.] / Ш. Шекхар, С. Чаула. – М. : КУДИЦ-ОБРАЗ, 2004. – 336 с.
95. **Юрасов, В. Г.** Нарушения целостности данных при асинхронной репликации и методы их решения / В. Г. Юрасов, Д. А. Апанасевич // Информационная безопасность. – 2008. – № 3 (10). – С. 423–426.
96. **Яковлев, Д. В.** Концепция построения систем контроля состояния горного массива как элемента многофункциональной системы безопасности угольных шахт / Д. В. Яковлев, Т. И. Лазаревич, А. Н. Поляков С. Н. Мулев // Сборник научных трудов ВНИМИ. Посвящен 100-летию юбилею выдающегося горного инженера Б. Ф. Братченко. – СПб. : ВНИМИ, 2012. – С 7–18.
97. **Ambler, S. W.** Test-Driven Development of Relational Databases / S. W. Ambler // IEEE Software. – 2007. – Vol. 24. – N3, – P. 37–43.
98. **Amja, A. M.** A Distributed Mobile Database Architecture / A. M. Amja, A. Obaid, N. Seguin // IEEE Asia-Pacific Services Computing Conference. – 2011. P. 62–69.
99. **Björkqvist, M.** Dynamic Replication in Service-Oriented Systems / M. Björkqvist, L. Y. Chen, W. Binder // 12th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing. – 2012. – P. 531–538.
100. **ctrlReplіc** : свидетельство о государственной регистрации программ для ЭВМ № 2013618670 / В. А. Дунаев, А. А. Цельшков, Ю. И. Федоров, А. Ю. Кузнецов, И. О. Ковыршин, О. А. Сенотрусов. заявка № 2013616315 от 19.07.2013.
101. **Garcia-Molina, H.** Database Systems. The Complete Book / H. Garcia-Molina, J. D. Ullman, J. Widom. – New Jersey : Department of Computer Science Stanford University, 2012. – 395 p.
102. **Gillenson, M. L.** Fundamentals of database management systems / M. L. Gillenson. – Memphis : Fogelman College of Business and Economics University, 2012. – 395 p.

103. **Grbac, T. G.** A Second Replicated Quantitative Analysis of Fault Distributions in Complex Software Systems / T. G. Grbac, P. Runeson, D. Huljenic // IEEE Transactions on Software Engineering. – 2013. – Vol. 39. – N4, – P. 462–476.
104. **IEEE 754–2008** Standard for Floating-Point Arithmetic, 2008.
105. **Johansson, J. M.** Modeling Network Latency and Parallel Processing in Distributed Database Design / J. M. Johansson, S. T. March, J.D. Naumann. // Decision Sciences Journal. – 2003. – Vol. 34, – № 4. – P. 677–706.
106. **Lin, X.** Query optimization strategies and implementation based on distributed database / X. Lin // 2nd IEEE International Conference on Computer Science and Information Technology. – 2009. – P. 480–484.
107. **Lin, Y.** Consistent Data Replication / Y. Lin, B. Kemme, M. Patino-Martinez, R. Jimenez Peris // Conf. Lisabon. – 2005. – P. 633–643.
108. **Sciascia, D.** Geo-replicated storage with scalable deferred update replication / D. Sciascia, F. Pedone // 43rd Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN). – 2013. – P. 1–12.
109. **Thakar, A. R.** The Catalog Archive Server Database Management System / A. R. Thakar, A. Szalay, G. Fekete, J. Gray // Computing in Science and Engineering. . – 2008. – Vol. 10. – N1, – P. 33–37.
110. **Weng, L.** Optimizing multiple queries on scientific datasets with partial replicas / L. Weng, U. Catalyurek, T. Kurc, G. Agrawal, J. Saltz // 8th IEEE/ACM International Conference on Grid Computing. – 2007. – P 259–266.
111. **Wong, L.** Oracle Streams: A High Performance Implementation for Near Real Time Asynchronous Replication / L. Wong, N. S. Arora, L. Gao, T. Hoang, J. Wu // IEEE International Conference on Data Engineering. – 2009. – P 1363–1374.

Приложение А. Характеристики фрагментов РБД предприятия ГПК

"ШахтИнвестКузбасс"

Таблица А.1 – Характеристики фрагментов

Раз мер, байт	Интенсивность поисковых запросов в минуту	Интенсивность запросов на обновление в минуту	Раз мер, байт	Интенсивность поисковых запросов в минуту	Интенсивность запросов на обновление в минуту	Раз мер, байт	Интенсивность поисковых запросов в минуту	Интенсивность запросов на обновление в минуту	Раз мер, байт	Интенсивность поисковых запросов в минуту	Интенсивность запросов на обновление в минуту
6619	0,3875	0,0201	4656	0,6430	0,0243	4444	0,2997	0,0300	4567	0,6068	0,0523
7982	0,3454	0,0557	3745	0,4950	0,0291	6790	0,2815	0,0512	8585	0,4848	0,0303
7399	0,6153	0,0495	4372	0,5287	0,0583	8057	0,5597	0,0596	6462	0,5683	0,0401
6341	0,3738	0,0417	2990	0,2177	0,0465	1334	0,4711	0,0440	8925	0,6110	0,0228
8302	0,6472	0,0560	9611	0,3523	0,0554	4814	0,4985	0,0455	3089	0,5153	0,0474
3473	0,6697	0,0440	1916	0,3769	0,0454	8310	0,3824	0,0471	5935	0,2196	0,0474
9067	0,3572	0,0378	6209	0,6660	0,0330	5895	0,3299	0,0324	3898	0,5622	0,0295
4668	0,4306	0,0523	1057	0,5159	0,0347	5768	0,4463	0,0597	8447	0,2979	0,0264
1099	0,3491	0,0331	8765	0,6486	0,0440	9208	0,2004	0,0263	4153	0,6617	0,0377
5441	0,2879	0,0424	3097	0,4521	0,0286	6394	0,4378	0,0476	6461	0,5240	0,0439
5983	0,3079	0,0268	7597	0,4398	0,0497	9784	0,5485	0,0414	1302	0,3255	0,0278
2727	0,2295	0,0544	6519	0,5214	0,0445	2865	0,5323	0,0444	6857	0,2571	0,0391
3674	0,2971	0,0354	7254	0,6758	0,0365	9083	0,5312	0,0299	2341	0,2018	0,0248
9939	0,5155	0,0207	2752	0,6658	0,0207	7906	0,2386	0,0267	9795	0,6555	0,0227
2390	0,5909	0,0593	3038	0,6330	0,0588	5935	0,5875	0,0246	9154	0,6611	0,0223
6111	0,4413	0,0418	7677	0,2800	0,0568	9066	0,6263	0,0218	3093	0,6037	0,0497
2786	0,6057	0,0564	3245	0,3627	0,0533	7381	0,5076	0,0497	4197	0,4695	0,0494
5569	0,3804	0,0571	4656	0,2844	0,0503	4724	0,3600	0,0310	8719	0,3254	0,0221
1316	0,2582	0,0281	8861	0,5371	0,0468	3090	0,5963	0,0219	7365	0,5457	0,0359
4721	0,2518	0,0554	3082	0,3944	0,0347	4511	0,4079	0,0342	7153	0,3262	0,0366
9277	0,4406	0,0335	6982	0,2208	0,0582	4077	0,4587	0,0544	9035	0,4945	0,0566
4678	0,4635	0,0567	9907	0,5648	0,0472	9237	0,4592	0,0469	2871	0,2086	0,0262
6631	0,5824	0,0431	2812	0,2578	0,0584	3660	0,4188	0,0580	8779	0,4131	0,0410
3722	0,5489	0,0287	3875	0,6681	0,0418	2755	0,6480	0,0363	3712	0,4980	0,0487
6830	0,2727	0,0382	3371	0,2252	0,0471	8665	0,4135	0,0327	6702	0,6811	0,0589
9171	0,2002	0,0486	4261	0,6737	0,0438	8995	0,6761	0,0500	6570	0,4694	0,0578
5224	0,4627	0,0273	3665	0,2691	0,0365	1383	0,2045	0,0384	5341	0,3887	0,0447
6133	0,2657	0,0230	3109	0,5083	0,0270	4897	0,4394	0,0340	6746	0,6379	0,0575
9825	0,6594	0,0563	4673	0,4192	0,0591	3380	0,2194	0,0209	8352	0,5386	0,0261
9109	0,5984	0,0292	8217	0,4366	0,0457	6182	0,3400	0,0297	2647	0,6327	0,0575
3832	0,5866	0,0303	7608	0,4783	0,0574	3551	0,6017	0,0415	4212	0,4891	0,0582
2473	0,4594	0,0323	5679	0,2905	0,0255	6053	0,6796	0,0314	6306	0,6732	0,0315
1021	0,3206	0,0550	2299	0,6330	0,0582	1425	0,4079	0,0275	4720	0,6046	0,0247
8753	0,5728	0,0255	6726	0,2438	0,0510	6050	0,4347	0,0386	5707	0,3775	0,0473
1545	0,2361	0,0439	6715	0,3211	0,0353	8311	0,6742	0,0558	5614	0,5529	0,0491
9360	0,4540	0,0427	1056	0,6378	0,0279	1550	0,3819	0,0456	9113	0,2089	0,0259
1396	0,5239	0,0538	9265	0,4575	0,0515	6993	0,5734	0,0330	2980	0,4501	0,0361
7193	0,3520	0,0288	2970	0,6920	0,0413	2989	0,5250	0,0582	6720	0,5821	0,0229
8533	0,3809	0,0464	4265	0,6495	0,0352	7097	0,4334	0,0264	3682	0,4196	0,0313
5466	0,6241	0,0431	8445	0,3186	0,0284	7778	0,6438	0,0319	1125	0,5549	0,0567
6887	0,5451	0,0540	2325	0,3636	0,0299	3355	0,2795	0,0205	9669	0,4101	0,0582
9237	0,6595	0,0297	4077	0,2889	0,0254	8377	0,5996	0,0533	7348	0,6974	0,0333
1006	0,4477	0,0220	9199	0,4295	0,0282	5704	0,6237	0,0580	6111	0,6204	0,0300
4750	0,6487	0,0322	4546	0,2163	0,0229	9075	0,3514	0,0596	6725	0,2160	0,0318
8338	0,2197	0,0246	9317	0,6073	0,0574	9808	0,2604	0,0525	2741	0,4265	0,0370
2705	0,6218	0,0463	3209	0,3001	0,0353	6864	0,5309	0,0448	5912	0,2783	0,0508
3964	0,4073	0,0421	5488	0,5144	0,0265	8049	0,6545	0,0404	9871	0,5971	0,0368

Таблица А.1 – Продолжение

73566	0,3126	0,0408	75656	0,5719	0,0296	75846	0,4462	0,0200	96965	0,2963	0,0302
94172	0,3506	0,0397	67584	0,3117	0,0554	97569	0,4306	0,0320	32737	0,6475	0,0346
47866	0,6082	0,0559	89167	0,5649	0,0382	22602	0,6065	0,0537	84146	0,2501	0,0522
27337	0,5697	0,0331	47044	0,5085	0,0339	11428	0,3183	0,0232	97761	0,2159	0,0520
50939	0,6842	0,0516	16361	0,4558	0,0220	64565	0,3346	0,0483	96350	0,5220	0,0398
68037	0,4878	0,0448	82051	0,3671	0,0489	67051	0,4708	0,0419	18443	0,6712	0,0201
56235	0,4530	0,0435	61582	0,3606	0,0233	36709	0,3355	0,0257	75074	0,2383	0,0315
98635	0,2277	0,0391	65664	0,4873	0,0285	23022	0,6849	0,0451	41935	0,5898	0,0503
69062	0,2470	0,0372	23008	0,6690	0,0458	43301	0,2663	0,0599	59608	0,3351	0,0463
94070	0,3462	0,0262	84306	0,5376	0,0598	14543	0,5041	0,0209	31339	0,5765	0,0487
17636	0,2661	0,0350	79125	0,6425	0,0275	50420	0,4818	0,0296	28312	0,5400	0,0469
63736	0,5396	0,0444	27474	0,5900	0,0537	91466	0,3135	0,0250	26623	0,4769	0,0453
92312	0,3135	0,0384	12164	0,5285	0,0518	24824	0,2611	0,0229	31787	0,2627	0,0230
69136	0,3760	0,0231	19339	0,4181	0,0542	32729	0,6286	0,0310	67639	0,6296	0,0563
16534	0,2823	0,0502	97685	0,2380	0,0377	67683	0,5282	0,0395	56155	0,5655	0,0223
29147	0,2918	0,0413	17221	0,2494	0,0330	33921	0,6728	0,0600	87126	0,2690	0,0506
26584	0,2995	0,0232	97374	0,2401	0,0348	35604	0,4113	0,0481	95476	0,5122	0,0342
94347	0,2532	0,0343	18210	0,6289	0,0495	33286	0,3228	0,0336	79136	0,3090	0,0354
68150	0,4549	0,0432	64101	0,3234	0,0543	65227	0,5245	0,0321	19542	0,3268	0,0446
28521	0,4328	0,0524	29369	0,5753	0,0265	19904	0,4095	0,0580	96619	0,3372	0,0323
27947	0,6866	0,0478	25274	0,3105	0,0513	32979	0,6383	0,0367	58138	0,4982	0,0420
55400	0,3034	0,0202	75840	0,3915	0,0284	51609	0,4873	0,0293	67444	0,5378	0,0586
38711	0,2447	0,0451	84003	0,2598	0,0552	79647	0,5224	0,0312	66625	0,3032	0,0383
48176	0,5167	0,0298	60841	0,6569	0,0233	96317	0,3563	0,0501	96723	0,2488	0,0495
77387	0,2519	0,0220	78683	0,2029	0,0251	81251	0,2317	0,0488	91669	0,2534	0,0294
54567	0,6266	0,0531	98138	0,3741	0,0448	44767	0,6416	0,0392	21583	0,2645	0,0357
70751	0,3089	0,0560	21704	0,6190	0,0221	14779	0,5502	0,0297	21193	0,2603	0,0392
67523	0,4499	0,0209	17584	0,2996	0,0281	11494	0,4193	0,0288	28485	0,6226	0,0401
50722	0,3916	0,0380	18050	0,6888	0,0341	80787	0,2484	0,0522	20709	0,4253	0,0290
77167	0,4376	0,0533	73742	0,4486	0,0491	40482	0,5379	0,0227	39299	0,2841	0,0311
64917	0,5427	0,0225	76203	0,4752	0,0500	16012	0,3313	0,0210	20014	0,3511	0,0269
49769	0,6795	0,0443	22838	0,4620	0,0564	38604	0,2367	0,0597	36577	0,6373	0,0461
27796	0,6181	0,0467	58017	0,5998	0,0289	99146	0,6003	0,0538	41296	0,3139	0,0303
20305	0,3876	0,0476	80724	0,2492	0,0583	99970	0,2116	0,0486	27362	0,4375	0,0275
55010	0,6756	0,0470	52351	0,2912	0,0222	53035	0,6980	0,0305	73190	0,5780	0,0591
43174	0,2150	0,0244	49835	0,4089	0,0359	33616	0,2481	0,0313	44526	0,4588	0,0573
89925	0,5579	0,0400	90497	0,4304	0,0338	58330	0,4085	0,0343	26395	0,6527	0,0362
86742	0,3838	0,0255	63530	0,6733	0,0375	52425	0,2298	0,0558	41180	0,5829	0,0421
27647	0,3779	0,0438	41279	0,4782	0,0307	97594	0,4058	0,0586	46517	0,4429	0,0280
58196	0,6230	0,0287	96806	0,3457	0,0298	38258	0,5653	0,0303	43806	0,6019	0,0267
69734	0,5286	0,0427	70921	0,4988	0,0220	97234	0,4526	0,0478	19915	0,5847	0,0392
61626	0,6700	0,0503	80482	0,2160	0,0386	83998	0,6574	0,0394	29056	0,4775	0,0414
33275	0,5780	0,0316	49118	0,4559	0,0241	77854	0,2992	0,0277	80023	0,3643	0,0599
60811	0,3871	0,0363	97067	0,5496	0,0582	23203	0,2865	0,0281	64576	0,4925	0,0363
32416	0,5229	0,0331	29207	0,6138	0,0364	93273	0,2925	0,0275	65722	0,5395	0,0302
85000	0,6872	0,0206	57177	0,2772	0,0438	69029	0,5038	0,0389	41372	0,4609	0,0419
39939	0,4451	0,0335	28068	0,2920	0,0237	95899	0,5344	0,0416	61615	0,6883	0,0398
17358	0,2723	0,0418	62052	0,2639	0,0538	11846	0,6421	0,0244	10895	0,2456	0,0487
62467	0,6782	0,0555	11967	0,4562	0,0270	22385	0,3910	0,0584	83237	0,4615	0,0491
87805	0,4775	0,0586	16059	0,3341	0,0579	76420	0,6664	0,0450	30045	0,5816	0,0411
16501	0,2404	0,0565	98910	0,2490	0,0301	89931	0,3093	0,0517	60020	0,3734	0,0420
61195	0,5843	0,0423	10027	0,3703	0,0288	69921	0,6154	0,0558	38233	0,4243	0,0599
44993	0,5597	0,0447	37574	0,6233	0,0559	91356	0,3162	0,0588	96190	0,5950	0,0306
54241	0,6471	0,0423	87527	0,6174	0,0582	28870	0,6380	0,0399	28521	0,5558	0,0468
56806	0,4890	0,0410	78996	0,3163	0,0401	57734	0,3068	0,0395	65977	0,6758	0,0280
86462	0,4658	0,0332	88371	0,6054	0,0282	51348	0,4268	0,0531	36049	0,2967	0,0509
85610	0,3352	0,0236	10591	0,5423	0,0223	71545	0,3503	0,0313	86061	0,2768	0,0547
78310	0,6553	0,0564	12799	0,6842	0,0259	47879	0,4908	0,0494	17682	0,6999	0,0498
41326	0,6764	0,0528	84805	0,5553	0,0562	23000	0,3020	0,0462	45841	0,4137	0,0330
99322	0,6194	0,0386	82163	0,3363	0,0267	25560	0,5233	0,0361	63332	0,3546	0,0336
98033	0,6689	0,0359	37211	0,2054	0,0363	98173	0,5844	0,0253	46822	0,4421	0,0341
37285	0,4439	0,0391	53980	0,5262	0,0268	92246	0,4059	0,0212	87920	0,2278	0,0470

Таблица А.1 – Продолжение

87950	0,4232	0,0469	63093	0,3371	0,0293	51010	0,6387	0,0429	96575	0,6391	0,0418
91922	0,5972	0,0239	66988	0,6624	0,0247	16666	0,4632	0,0300	29795	0,6110	0,0274
13389	0,4141	0,0355	98151	0,4286	0,0363	24082	0,3264	0,0508	52474	0,2938	0,0493
10027	0,6729	0,0414	51510	0,6048	0,0475	62942	0,2617	0,0468	22714	0,5526	0,0405
28364	0,3391	0,0358	19723	0,2989	0,0566	35294	0,5260	0,0496	86247	0,2949	0,0560
49445	0,5495	0,0352	48668	0,3566	0,0205	69718	0,3369	0,0314	20676	0,4417	0,0227
40362	0,4313	0,0513	60673	0,5884	0,0576	42169	0,4007	0,0469	14007	0,3421	0,0433
92837	0,5788	0,0536	74401	0,4540	0,0320	62137	0,5473	0,0223	84047	0,4837	0,0398
61687	0,5121	0,0341	60253	0,5990	0,0555	11810	0,3601	0,0572	99074	0,6854	0,0335
25530	0,5708	0,0525	77799	0,4143	0,0338	84440	0,6150	0,0490	97830	0,4858	0,0475
90153	0,6102	0,0530	66241	0,5389	0,0262	25450	0,6674	0,0243	17303	0,2630	0,0459
20745	0,2198	0,0572	33520	0,2408	0,0267	89195	0,6169	0,0258	54924	0,4604	0,0320
11813	0,3337	0,0481	31284	0,2972	0,0342	92208	0,5236	0,0415	79090	0,3383	0,0415
86382	0,5368	0,0240	31874	0,2233	0,0579	71701	0,5631	0,0256	75011	0,3724	0,0340
45410	0,4246	0,0343	25376	0,6634	0,0474	28903	0,3397	0,0439	18559	0,2833	0,0406
37060	0,6253	0,0569	83276	0,4243	0,0386	31451	0,5510	0,0261	29356	0,2124	0,0282
60162	0,6147	0,0447	36731	0,4399	0,0214	52713	0,3653	0,0216	81465	0,5867	0,0240
58877	0,5691	0,0454	12840	0,6227	0,0365	93059	0,6151	0,0279	80120	0,2870	0,0401
77123	0,5970	0,0588	56287	0,2618	0,0216	32064	0,4269	0,0330	58800	0,5415	0,0435
88871	0,6136	0,0301	81960	0,4047	0,0344	42600	0,6392	0,0395	95070	0,4870	0,0297
20251	0,6936	0,0489	87953	0,6767	0,0557	10261	0,3133	0,0540	65593	0,4486	0,0456
53024	0,5227	0,0325	12843	0,3636	0,0203	88689	0,3446	0,0299	62242	0,6455	0,0434
78573	0,2018	0,0272	58819	0,2290	0,0372	16545	0,3819	0,0596	44028	0,6306	0,0272
51464	0,2942	0,0286	26131	0,3012	0,0447	96487	0,2510	0,0315	59314	0,2347	0,0238
25780	0,4917	0,0213	71921	0,4882	0,0554	62360	0,3261	0,0328	98951	0,6329	0,0527
73997	0,6327	0,0529	66032	0,2401	0,0419	23978	0,4264	0,0365	46261	0,3889	0,0233
30402	0,2856	0,0287	80337	0,2262	0,0538	77422	0,2804	0,0268	72588	0,2743	0,0589
39395	0,5210	0,0513	89296	0,2510	0,0420	69103	0,6095	0,0392	53024	0,3940	0,0512
58707	0,2176	0,0280	96416	0,3514	0,0516	52304	0,6330	0,0282	85844	0,2488	0,0525
54158	0,2140	0,0218	79208	0,4041	0,0224	20160	0,2269	0,0336	12925	0,6672	0,0273
95853	0,3317	0,0364	25008	0,2978	0,0442	42647	0,6109	0,0481	86286	0,3393	0,0468
86984	0,3102	0,0409	20209	0,3694	0,0368	36656	0,2346	0,0502	13689	0,6641	0,0554
18907	0,6741	0,0329	56336	0,4227	0,0440	36283	0,4673	0,0341	70067	0,5482	0,0328
60302	0,5530	0,0501	19124	0,6305	0,0488	89239	0,4161	0,0301	37082	0,4255	0,0587
10242	0,6475	0,0496	93155	0,6392	0,0296	99500	0,5031	0,0248	96863	0,6390	0,0227
56149	0,4907	0,0239	41136	0,2617	0,0593	72462	0,5705	0,0349	52214	0,4089	0,0371
11994	0,3707	0,0243	10206	0,5643	0,0417	93944	0,2471	0,0557	45556	0,2776	0,0420
26485	0,3819	0,0298	99283	0,5481	0,0564	38546	0,2392	0,0399	11994	0,6596	0,0401
73747	0,5030	0,0425	52159	0,4653	0,0569	12142	0,4730	0,0469	34338	0,2295	0,0348
44325	0,6012	0,0281	31361	0,2270	0,0316	40485	0,5630	0,0286	64098	0,5805	0,0221
36818	0,3953	0,0470	30891	0,2949	0,0301	58970	0,4041	0,0493	54081	0,4693	0,0316
27732	0,5774	0,0226	23961	0,6023	0,0530	65260	0,3418	0,0276	42482	0,3558	0,0512
71539	0,5988	0,0222	75524	0,6865	0,0301	49379	0,5119	0,0597	18836	0,6520	0,0208
80265	0,2995	0,0348	36253	0,5458	0,0383	18383	0,2038	0,0590	50261	0,3047	0,0471
99269	0,2469	0,0504	39623	0,4904	0,0249	61247	0,3157	0,0446	44078	0,2488	0,0424
84924	0,4730	0,0396	82416	0,2646	0,0420	86149	0,2861	0,0267	24722	0,3811	0,0558
50368	0,6903	0,0357	81644	0,3418	0,0473	36612	0,5997	0,0546	12401	0,3248	0,0521
10591	0,3008	0,0416	57226	0,4905	0,0408	47813	0,3045	0,0552	47780	0,5170	0,0317
23879	0,5848	0,0455	87401	0,4137	0,0350	27180	0,5428	0,0415	89618	0,4868	0,0553
66186	0,4266	0,0568	90922	0,2545	0,0591	32891	0,6505	0,0547	35538	0,3397	0,0336
85599	0,4481	0,0285	38832	0,4952	0,0345	11134	0,4680	0,0288	12428	0,5377	0,0412
74890	0,6100	0,0520	92878	0,4505	0,0210	24846	0,5195	0,0562	55911	0,2927	0,0205
79853	0,5944	0,0406	62810	0,4028	0,0545	19693	0,3260	0,0516	83652	0,3743	0,0480
25123	0,5334	0,0530	55586	0,2906	0,0406	58297	0,3531	0,0322	48959	0,6599	0,0555
87838	0,3819	0,0354	58190	0,5463	0,0470	64777	0,6926	0,0532	13214	0,2522	0,0237
12115	0,5624	0,0289	96641	0,3123	0,0425	48283	0,3802	0,0522	27650	0,2584	0,0238
73014	0,4254	0,0466	10327	0,2925	0,0265	14529	0,5439	0,0371	44534	0,2484	0,0593
11154	0,6249	0,0366	83668	0,6345	0,0577	93683	0,5895	0,0523	62868	0,4286	0,0558
23140	0,4056	0,0444	40875	0,6594	0,0296	62958	0,4435	0,0488	87220	0,6604	0,0247
76189	0,4415	0,0265	11239	0,4847	0,0530	58709	0,3583	0,0582	42196	0,2278	0,0218
90313	0,4038	0,0255	38181	0,4015	0,0584	94358	0,5866	0,0375	59915	0,2478	0,0417
52290	0,4450	0,0482	86261	0,4454	0,0349	42773	0,5244	0,0459	53601	0,2392	0,0393

Таблица А.1 – Продолжение

65919	0,3719	0,0560	41315	0,2636	0,0484	64269	0,4701	0,0563	82611	0,2054	0,0230
93562	0,6859	0,0586	85300	0,2108	0,0369	76066	0,2224	0,0329	98585	0,4524	0,0384
98382	0,4719	0,0305	99813	0,6593	0,0314	17290	0,3470	0,0347	17062	0,6249	0,0236
65675	0,6427	0,0234	40120	0,2090	0,0473	52430	0,6307	0,0300	31213	0,5664	0,0394
68562	0,6558	0,0378	32946	0,3278	0,0228	59550	0,6496	0,0358	98632	0,5180	0,0221
30548	0,4535	0,0390	12357	0,5557	0,0216	48574	0,2053	0,0428	95842	0,6256	0,0521
52760	0,5583	0,0507	68034	0,2042	0,0442	41356	0,2840	0,0495	33668	0,2931	0,0429
36192	0,4742	0,0238	91488	0,2578	0,0336	75475	0,3126	0,0244	63332	0,4361	0,0429
26884	0,6128	0,0568	58034	0,3628	0,0515	34638	0,6777	0,0216	65348	0,4697	0,0208
93287	0,5577	0,0294	67677	0,2130	0,0267	80834	0,2450	0,0435	69949	0,5338	0,0489
59325	0,3327	0,0514	25909	0,5285	0,0332	71481	0,2591	0,0589	28194	0,4388	0,0526
38455	0,3517	0,0421	99673	0,6472	0,0215	55979	0,2069	0,0240	90310	0,4044	0,0376
25428	0,4922	0,0504	78139	0,3542	0,0418	55746	0,2279	0,0307	39387	0,2990	0,0548
48069	0,5411	0,0525	83078	0,5179	0,0591	53870	0,3177	0,0388	82932	0,4928	0,0242
84454	0,5466	0,0407	23437	0,4998	0,0494	90560	0,6019	0,0270	63870	0,2733	0,0599
82474	0,6334	0,0466	63255	0,4049	0,0207	42315	0,3012	0,0318	29578	0,6118	0,0402
81677	0,6319	0,0368	56789	0,5794	0,0551	31608	0,6601	0,0282	82166	0,5103	0,0586
14936	0,4024	0,0252	65189	0,3098	0,0434	34992	0,6269	0,0302	14400	0,4608	0,0402
15576	0,4145	0,0539	34107	0,3362	0,0546	21171	0,6048	0,0394	85308	0,2556	0,0204
78329	0,2921	0,0322	36684	0,4163	0,0409	28241	0,2602	0,0204	81965	0,4283	0,0593
83265	0,5144	0,0264	40925	0,5107	0,0381	21407	0,6036	0,0292	72206	0,5017	0,0418
82279	0,2430	0,0291	97358	0,5339	0,0313	40224	0,6640	0,0355	90834	0,4697	0,0238
18581	0,2383	0,0477	80367	0,3645	0,0364	89510	0,4028	0,0333	51909	0,3630	0,0278
80175	0,3547	0,0248	80095	0,4649	0,0301	87404	0,6002	0,0261	27760	0,4550	0,0502
57256	0,5100	0,0356	52667	0,2594	0,0331	22580	0,4958	0,0397	28112	0,6225	0,0482
95298	0,3472	0,0342	90595	0,2230	0,0578	97160	0,4076	0,0313	62931	0,2634	0,0526
13692	0,2282	0,0340	92084	0,6373	0,0253	18064	0,4392	0,0346	77043	0,2368	0,0411
24870	0,2531	0,0431	20187	0,6901	0,0302	61494	0,6037	0,0440	35440	0,3975	0,0588
65521	0,3881	0,0243	97984	0,2625	0,0560	17540	0,4017	0,0371	33734	0,6154	0,0311
26117	0,4629	0,0518	83811	0,2488	0,0303	71042	0,6756	0,0568	26351	0,2854	0,0551
16361	0,2220	0,0457	97075	0,2402	0,0530	98646	0,4721	0,0319	47099	0,2834	0,0508
25535	0,4763	0,0468	14922	0,4809	0,0253	87450	0,6132	0,0493	60747	0,5775	0,0487
55271	0,4219	0,0378	13821	0,6658	0,0560	23393	0,6948	0,0516	72119	0,6446	0,0351
28392	0,2987	0,0336	38686	0,4860	0,0558	21874	0,3957	0,0540	76140	0,4053	0,0297
74750	0,2585	0,0257	37711	0,3648	0,0348	42309	0,3233	0,0409	39455	0,5563	0,0438
75365	0,6654	0,0353	79892	0,5692	0,0201	53927	0,5954	0,0509	89027	0,5971	0,0503
38865	0,6672	0,0549	90252	0,6063	0,0264	23698	0,4256	0,0303	29762	0,3649	0,0555
44166	0,5018	0,0519	99497	0,2762	0,0335	15378	0,3118	0,0414	80142	0,5251	0,0316
20347	0,2051	0,0293	82745	0,4759	0,0570	97726	0,3246	0,0480	23448	0,4995	0,0303
42023	0,5412	0,0441	19152	0,3384	0,0235	82438	0,6176	0,0238	98591	0,4225	0,0457
31902	0,2405	0,0236	29185	0,3477	0,0376	66342	0,4578	0,0241	87808	0,2334	0,0475
45852	0,6330	0,0391	83177	0,3361	0,0390	65601	0,3891	0,0369	91776	0,4570	0,0314
13296	0,3954	0,0326	98322	0,6528	0,0497	64574	0,6621	0,0264	73382	0,3787	0,0526
67919	0,3060	0,0457	78373	0,6146	0,0345	33413	0,6243	0,0271	17908	0,4022	0,0276
33674	0,6785	0,0242	56262	0,4705	0,0259	46597	0,6979	0,0202	51645	0,5173	0,0548
61118	0,4046	0,0208	69954	0,4970	0,0460	32026	0,4364	0,0468	77378	0,2105	0,0522
29001	0,4776	0,0575	47352	0,4305	0,0471	63527	0,5742	0,0346	27656	0,4795	0,0373
96635	0,4088	0,0595	73036	0,4540	0,0258	42526	0,6574	0,0327	84503	0,4853	0,0473
19034	0,2024	0,0439	37406	0,5721	0,0259	89412	0,4988	0,0397	85575	0,2018	0,0507
73492	0,6967	0,0397	49876	0,4194	0,0311	58929	0,2086	0,0564	18539	0,5830	0,0528
77549	0,4286	0,0241	71201	0,2207	0,0566	39362	0,6046	0,0296	93076	0,4004	0,0522
80913	0,3161	0,0389	99585	0,6933	0,0224	87569	0,4091	0,0520	78266	0,6553	0,0484
45792	0,4693	0,0206	56702	0,6780	0,0408	60000	0,5639	0,0235	54018	0,2504	0,0216
26499	0,6175	0,0270	45498	0,6708	0,0498	29872	0,5073	0,0590	54666	0,6920	0,0413
22522	0,4366	0,0428	42817	0,6395	0,0574	58803	0,3646	0,0444	68380	0,5178	0,0299
42416	0,6783	0,0406	42273	0,2782	0,0310	51222	0,2479	0,0460	87093	0,3464	0,0226
86607	0,6364	0,0404	17048	0,3181	0,0398	16806	0,4300	0,0457	25903	0,4198	0,0452
24170	0,6223	0,0566	16106	0,3088	0,0479	65068	0,5433	0,0592	16114	0,3400	0,0470
70817	0,6037	0,0367	45361	0,3208	0,0298	89598	0,5068	0,0437	36099	0,4532	0,0376
98649	0,5311	0,0427	83567	0,6364	0,0352	80021	0,6099	0,0462	17611	0,3552	0,0438
49544	0,3167	0,0277	10244	0,4717	0,0354	45869	0,3579	0,0576	43735	0,5853	0,0531
55666	0,4593	0,0385	63750	0,3599	0,0363	84501	0,5783	0,0472	83212	0,3541	0,0353

Таблица А.1 – Продолжение

79900	0,6747	0,0452	35371	0,5432	0,0578	62467	0,3643	0,0259	16740	0,6188	0,0364
77277	0,4254	0,0380	60866	0,6094	0,0557	94798	0,3936	0,0509	58819	0,5172	0,0361
46264	0,2549	0,0507	98028	0,3331	0,0504	30704	0,2086	0,0221	93941	0,2766	0,0313
56910	0,4545	0,0484	66370	0,3338	0,0337	14562	0,6093	0,0393	31454	0,2885	0,0337
14224	0,2415	0,0209	54540	0,2479	0,0438	72423	0,4575	0,0453	84443	0,6702	0,0495
93262	0,6230	0,0488	57427	0,2099	0,0483	52966	0,2518	0,0451	31421	0,6594	0,0332
21986	0,3004	0,0222	80611	0,6024	0,0468	46723	0,6316	0,0408	90057	0,3798	0,0273
87006	0,4043	0,0291	68430	0,6654	0,0451	11041	0,6138	0,0230	58226	0,6331	0,0330
57245	0,6302	0,0307	59193	0,2776	0,0392	81043	0,3229	0,0446	58072	0,4450	0,0424
89005	0,2511	0,0488	70210	0,2206	0,0386	87937	0,4181	0,0282	10742	0,5985	0,0243
66232	0,5071	0,0386	54323	0,2527	0,0276	44256	0,6470	0,0543	85305	0,2066	0,0375
82224	0,5898	0,0202	14244	0,2231	0,0415	26279	0,6196	0,0598	21319	0,4165	0,0521
10648	0,4199	0,0319	11461	0,4116	0,0508	56284	0,3657	0,0483	86915	0,2143	0,0344
72899	0,5557	0,0390	62560	0,3980	0,0284	55946	0,4614	0,0373	65952	0,6981	0,0234
68430	0,4368	0,0563	84888	0,4792	0,0229	49283	0,4449	0,0373	76110	0,3854	0,0242
45429	0,6782	0,0514	64403	0,2761	0,0470	73280	0,6155	0,0253	84597	0,3465	0,0512
87368	0,3002	0,0554	21303	0,5283	0,0438	91356	0,6902	0,0499	74942	0,3263	0,0242