

На правах рукописи



Чусов Андрей Александрович

**ГИБКАЯ АРХИТЕКТУРА ДЛЯ ПАРАЛЛЕЛЬНОГО АНАЛИЗА И
ВИЗУАЛИЗАЦИИ ФИЗИЧЕСКИХ ПОЛЕЙ**

05.13.11 – математическое и программное обеспечение вычислительных машин,
комплексов и компьютерных сетей

Автореферат
диссертации на соискание учёной степени
кандидата технических наук

Владивосток – 2012

Работа выполнена на кафедре электроники и средств связи Инженерной школы Дальневосточного федерального университета

Научный руководитель: доктор физико-математических наук, профессор,
Стаценко Любовь Григорьевна

Официальные оппоненты: доктор технических наук, профессор,
Артемьева Ирина Леонидовна, профессор
кафедры программного обеспечения ДВФУ

доктор технических наук,
Шахгельдян Карина Иосифовна, профессор
кафедры информационных систем и
прикладной информатики ВГУЭС

Ведущая организация: Институт систем информатики им. А.П. Ершова
Сибирского отделения РАН, г. Новосибирск

Защита состоится 11 октября 2012 г. в 12 часов на заседании диссертационного совета Д 005.007.01 в Институте автоматизации и процессов управления ДВО РАН по адресу: 690041, г. Владивосток, ул. Радио, 5.

С диссертацией можно ознакомиться в библиотеке Института автоматизации и процессов управления ДВО РАН.

Автореферат разослан « ____ » _____ 2012 г.

Ученый секретарь
диссертационного совета Д 005.007.01, к.т.н.



А.В. Лебедев

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы. В настоящее время, в связи с развитием машинных вычислений, создано большое количество программных систем, задачей которых является моделирование и анализ физических процессов. Такие программные комплексы относят к классу систем компьютерного анализа (САА – computer aided analysis). Однако компьютерный анализ и моделирование обладает рядом фундаментальных проблем, связанных, в конечном счете, с высокой временной и пространственной сложностью процесса, и, соответственно, с высокими ограничениями по точности результатов, сложности входных параметров. Поэтому существующие на данный момент системы анализа, такие как Odeon Room Acoustics Software, Agilent Technologies EMDS и др., направлены лишь на решение ограниченного круга задач, связанных с моделированием полей, либо используют для физического анализа неточные методы, налагающие значительные ограничения на исходную геометрическую модель (форму и размеры), либо выполняющие этот анализ за длительные промежутки времени (на практике – до нескольких суток). Поэтому *актуальными проблемами*, возникающими при создании систем компьютерного моделирования физических процессов вообще и физических полей в частности, являются вопросы повышения точности анализа для произвольно сложных исходных моделей при понижении временных издержек и стоимости как процесса анализа, так и процесса создания используемой для этого системы САА.

Решение проблем скорости и точности выполнения анализа, состоит, очевидно, в создании САА, которые уже на этапе проектирования архитектуры нацелены на работу в высокопроизводительных вычислительных средах перспективных типов таких, как параллельные вычислительные среды, а также среды, представляющие собой распределенные вычислительные системы. Работы, посвященные распределенному имитационному моделированию и анализу, ведутся по двум основным направлениям.

Первое – параллельное дискретно-событийное моделирование (PDES – Parallel Discrete Event Simulation) – описывает управление и взаимодействие параллельных процессов, выполняющих компоненты имитационного моделирования, которые в совокупности составляют общую задачу в предметной области. Сам процесс моделирования описывается набором переменных состояния модели, изменяющихся во время проведения компонентов имитационного моделирования. Параллельные процессы выполняют эти компоненты с приходом событий, снабженных метками времени, в общий разделяемый список.

Второе направление развития систем распределенного моделирования и анализа вводит понятие федерации, определяющее объединение разнородных систем моделирования в одну надсистему по заданным протоколам, задавая способы обмена данными между этими системами.

Вместе с тем данные технологии не направлены на описание общих подсистем САА и описывают лишь процессы имитационного моделирования в

параллельных и распределенных средах.

Выделение и детализированное описание подсистем, независимых от конкретных задач анализа, позволяет снизить общую сложность САА за счет устранения необходимости создания и развертывания этих подсистем отдельно для каждой задачи. Применительно к системам моделирования и анализа физических полей к таким подсистемам можно отнести подсистемы визуализации, подсистемы, предоставляющие клиентам необходимые в процессе анализа данные, подсистемы, обеспечивающие многопользовательское управление процессом анализа и системой САА в целом, подсистемы, обеспечивающие безопасность САА. Выделение общих принципов создания архитектуры САА без привязки к конкретной предметной области позволяет повысить эффективность создания систем САА.

Использование параллельных и распределенных вычислительных ресурсов, в свою очередь, рождает необходимость создания практически применимых, легких для использования аналитических методов оценки вычислительной эффективности и, соответственно, целесообразности применения тех или иных методов распараллеливания алгоритмов. Данный вопрос является актуальным, поскольку существующие на данный момент методы таких оценок основаны на относительно трудоемком анализе структур распараллеливаемых алгоритмов как в исходной, последовательной, форме, так и в параллельной. При этом делается ряд существенных допущений и ограничений. Не учитываются издержки на создание, обслуживание и завершение потоков (которые, по отношению к последовательной сложности исходного алгоритма могут быть весьма велики), не рассматриваются издержки, связанные с используемой моделью синхронизации потоков выполнения. На практике эффективность параллельной формы алгоритма по отношению к последовательной вычисляется эмпирически для множества конкретных целевых платформ, на основании чего и делается вывод о целесообразности применения той или иной параллельной формы алгоритма.

В связи с этим **цель работы** заключается в разработке и реализации комплексного подхода к созданию архитектур систем компьютерного анализа, моделирования и визуализации физических полей различного типа с ориентацией на работу в параллельных и распределенных средах с тем, чтобы повысить допустимую сложность анализа и, таким образом, обеспечить высокую точность и скорость работы.

Основные задачи диссертационной работы.

1. Разработка принципов создания гибких архитектур систем САА, предназначенных для решения задач моделирования произвольных физических полей в пространстве неограниченной сложности и расширяемых на произвольное множество вычислительных узлов.
2. Разработка метода реализации архитектур гибких и расширяемых САА для моделирования и анализа физических полей в пространстве неограниченной сложности.
3. Разработка эффективных в плане производительности и функциональности механизмов синхронизации параллельного доступа

компонентов системы к разделяемым ресурсам.

4. Разработка метода оценки параллельной сложности, быстродействия алгоритмов, использующих блокирующую синхронизацию, с учетом издержек, связанных с состязательностью потоков.

5. Разработка законченной системы САА, использующей предлагаемые решения на примере моделирования звукового поля.

Методы исследования. В процессе выполнения диссертации использовались положения теорий чисел, сложности, информации, алгоритмов, теории вероятности, теории принятия решений, теории защиты информации, теории систем и системотехники, методы системного, объектно-ориентированного и модульного программирования, аппарат баз данных.

Научная новизна работы

Впервые предложены принципы создания систем САА, предназначенных для моделирования произвольных физических полей. Системы, построенные в соответствии с предложенными принципами, позволяют с высокой скоростью и точностью проводить анализ физических полей в пространстве неограниченной сложности.

Разработан механизм синхронизации потоков, накладывающий меньше по сравнению с аналогами временные издержки на синхронизацию и обладающий большей функциональностью.

Впервые разработан аналитический вероятностный метод оценки издержек, связанных с использованием механизмов синхронизации, с учетом состязательности потоков и параллельной сложности алгоритмов, использующих эти механизмы.

Практическая ценность и реализация результатов работы. Все исследования носят прикладной характер и направлены на повышение эффективности разработки высокопроизводительных систем компьютерного моделирования, анализа и визуализации произвольных физических полей на различных вычислительных платформах, а также расчета параметров этих полей.

В соответствии с предложенными архитектурными принципами построена законченная система САА, ориентированная на платформы intel x86 и x64, позволяющая проводить моделирование, анализ и визуализацию акустического поля в произвольном пространстве.

Разработанный механизм многопоточной синхронизации, является более эффективным (по скорости выполнения и функциональности) по сравнению с аналогами.

Предложенный метод оценки эффективности многопоточной синхронизации и параллельной сложности распараллеливаемых алгоритмов в отличие от аналогов позволяет учитывать издержки, связанные с состязательностью потоков, а также собственные сложности алгоритмов, реализующих синхронизацию.

Исследования, представленные в работе, выполнялись в рамках приоритетного национального проекта «Образование» в центре высокопроизводительных вычислений Дальневосточного Государственного

Технического Университета, позволяющем выполнять параллельные и распределенные вычисления в удаленном доступе.

Работа выполнялась в рамках Федеральной целевой программы «Научные и научно-педагогические кадры инновационной России на 2009-2013 гг., направление 1 (Стимулирование и закрепление молодежи в сфере науки, образования и высоких технологий)», а также в рамках государственного задания Министерства образования и науки Российской Федерации высшим учебным заведениям на 2012-2014 гг.

Результаты работы используются при проведении практических занятий по дисциплинам «Электроакустика и звуковое вещание», «Распространение радиоволн и антенно-фидерные устройства», «Техника и технология телерадиовещания» в лабораториях кафедры Электроники и средств связи, а также в процессе курсового проектирования, для моделирования и анализа акустических и электромагнитных полей в помещениях различного функционального назначения с использованием удаленного доступа.

Апробация работы. Научные и практические результаты работы докладывались и обсуждались на следующих конференциях и семинарах: 16-я Всероссийская межвузовская научно-техническая конференция студентов и аспирантов «Микроэлектроника и информатика — 2009» (г. Москва, апрель 2009г.); IX международный форум студентов, аспирантов и молодых ученых стран АТР (г. Владивосток, 2009г.); Региональные научно-практические конференции «Молодёжь и научно-технический прогресс» (г. Владивосток, 2008-2011гг.); Третья всероссийская научно-техническая конференция «Технические проблемы освоения мирового океана» (г. Владивосток, 2009г.); Научные конференции «Вологдинские чтения» (г. Владивосток, 2008-2011гг.); Научная конференция «Сессия Научного совета РАН по акустике и XXIV сессия Российского акустического общества» (г. Москва, 2011г.). Работа докладывалась на семинарах кафедр электроники и средств связи, акустики и медицинской техники ДВФУ (2012) и расширенном семинаре «Информационные суперкомпьютерные технологии» в Дальневосточном Научно-образовательном центре суперкомпьютерных технологий (НОЦ "СКТ-Дальний Восток") на базе ДВФУ (2012);

Публикации. По материалам диссертации опубликовано 22 работы, в том числе 4 статьи в журналах из списка ВАК, и 3 статьи в международных изданиях.

Структура и объем диссертации. Диссертационная работа состоит из введения, пяти глав и заключения, изложенных на 142 страницах, включая 70 рисунков и 4 таблицы, списка литературы, включающего 83 работы, и 3 приложений.

СОДЕРЖАНИЕ РАБОТЫ

Во введении рассматривается состояние проблемы на сегодняшний день и актуальность темы исследования, формулируются цель и задачи диссертационной работы, приводится ее краткое содержание.

Первая глава посвящена рассмотрению общих вопросов анализа и моделирования физических полей, описываются основные проблемы, связанные с организацией вычислений в параллельных и распределенных вычислительных средах, формулируются цели и задачи исследования.

Вторая глава описывает принципы, положенные в основу построения архитектуры систем компьютерного анализа. Здесь же определяются основные требования к входным данным, определяющим параметры моделирования.

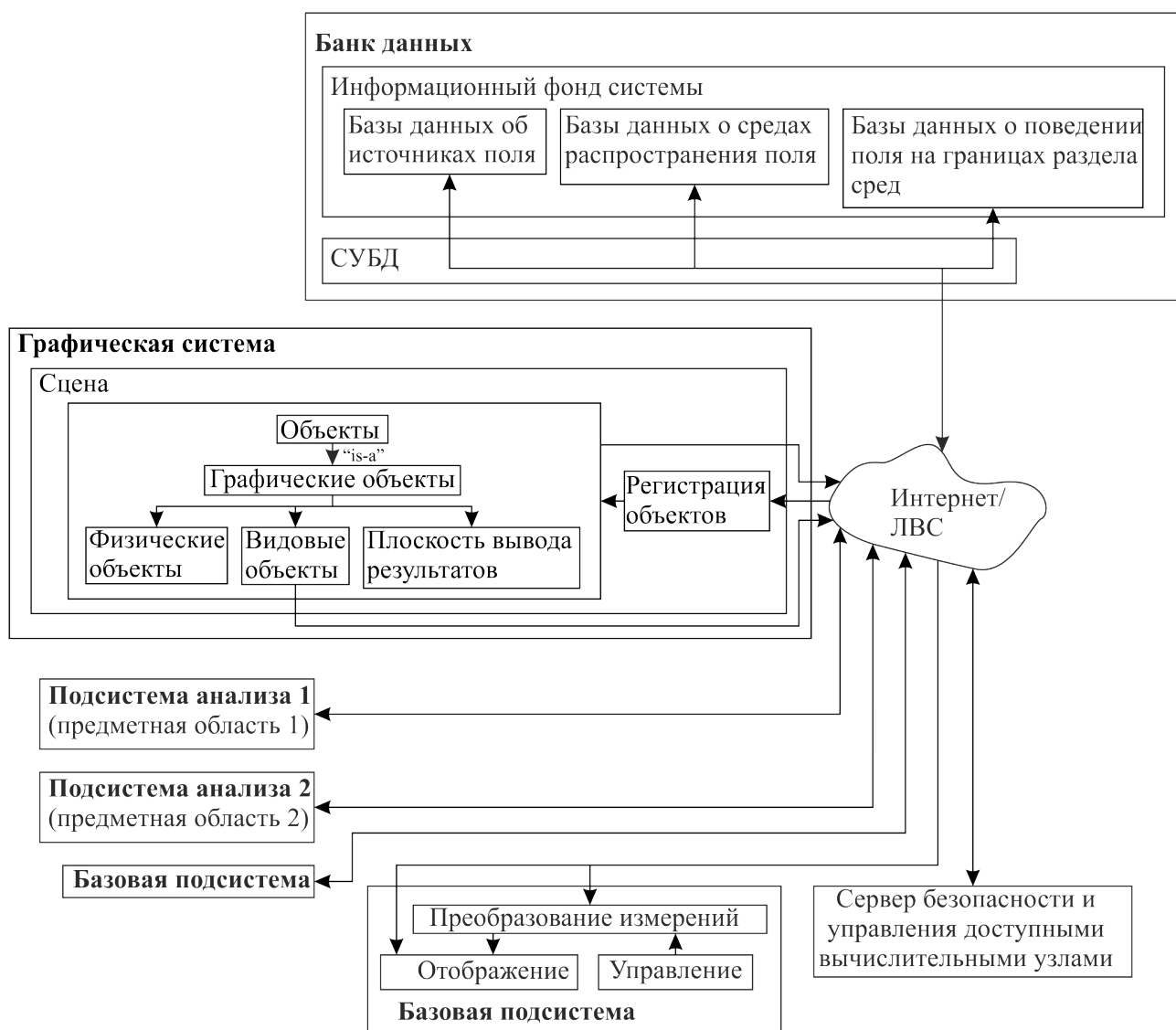


Рис. 1. Базовая структура распределенной системы САА

Принцип 1. Структура систем анализа (рис. 1) должна задаваться на основе следующих требований: 1 – гибкость системы для решения задач различных предметных областей; 2 – расширяемость на произвольное множество анализирующих и обслуживающих вычислительных узлов; 3 – возможность проведения анализа распространения любых физических полей одновременно множеством пользователей в одной среде и с использованием одной САА. Выделяются следующие обязательные компоненты систем, архитектура

которых является независимой от конкретной задачи анализа: графическая система, банк данных системы, подсистема управления доступными вычислительными узлами, совокупность базовых подсистем, имеющих общую реализацию. Кроме этого определяются подсистемы анализа, реализуемые разработчиком законченной САА для решения задач моделирования конкретных физических полей на основе требований к формату результатов анализа и с использованием средств поддержки разработчика, предоставляемых общими компонентами.

Принцип 2. Любое взаимодействие пользователя с системой должно обеспечиваться средствами базовой системы. Через нее должны определяться основные параметры модели: объекты, их аппроксимация, размеры пространства, параметры среды распространения моделируемого физического поля, а также положение секущих плоскостей, на которых рассчитываются распределения интенсивностей поля. Через нее же должны возвращаться результаты работы системы (рис. 2). Каждый пользователь должен взаимодействовать с собственным экземпляром базовой подсистемы.

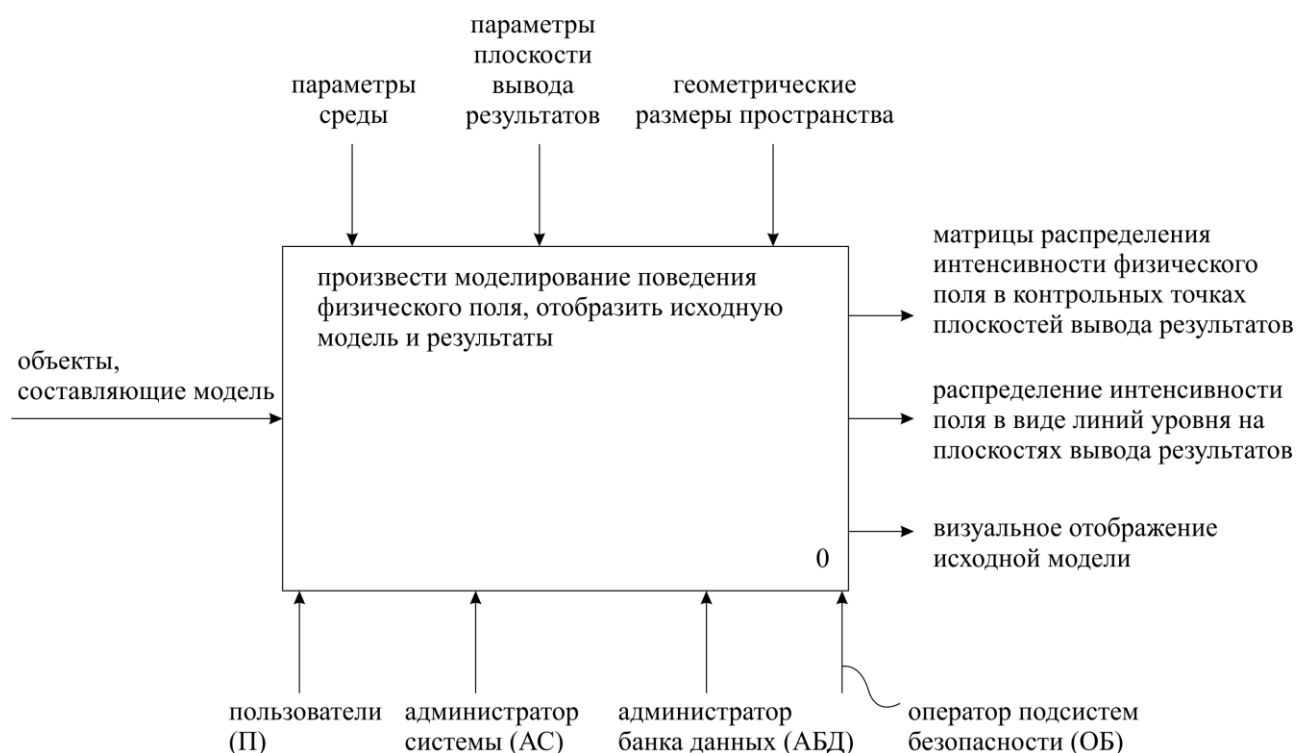


Рис. 2. Функциональная архитектура системы САА в целом (А-0)

Принцип 3. Все узлы распределенной системы при ее развертывании должны быть зарегистрированы на сервере безопасности и управления доступными вычислительными узлами. Под регистрацией узла на сервере понимается запись его идентификатора, адреса и ключа аутентификации согласно протоколам, определенным разработчиком САА. Сервер может быть выделен в отдельный узел либо являться подсистемой базовой системы. В последнем случае первый из создаваемых экземпляров базовой системы

помечается как основной – через него осуществляется управление узлами распределенной системы.

Принцип 4. Подсистема преобразования измерений необходима для реализации алгоритма взаимного отображения координат физического пространства на внутренние единицы измерения с тем, чтобы все значения могли быть заданы в единой целочисленной разрядной сетке. Этой подсистемой должны определяться цены деления и пределы измерений в системе. Вызовы, осуществляемые пользователями через подсистему управления, должны обрабатываться данной подсистемой.

Принцип 5. Графическая подсистема выносится в отдельный сервер и должна быть определена *на основе требований к обеспечению графического отображения пространства независимо от конкретных задач предметной области*, а также к обеспечению централизованного хранения элементов, составляющих исходную геометрическую модель. Подсистема готовит изображения моделируемого пространства до проведения процедуры анализа и передает их на базовые системы. На подсистему налагается дополнительное функциональное требование – хранение адресов элементов, составляющих моделируемый объект. Для этих целей графическая система должна включать в себя компонент-сцену, предоставляющую клиентам разделяемый доступ к элементам геометрической модели. Основные классы этих элементов представлены на рис. 1. Задача видового компонента — обеспечение графического отображения абстрактных объектов, определенных в другом (произвольном) архитектурном модуле. При этом в любой произвольный момент времени необходимо готовое изображение отображаемого пространства. Первая задача решается обеспечением каждого элемента (объекта) пространства унифицированным интерфейсом. Часть таких объектов относятся к классу графических, т. е. тех объектов, которые предоставляют интерфейс для их отображения. Процедура отображения должна выполняться в пространстве другого модуля — того, в котором объект определен. Решение о вынесении графической системы в отдельный вычислительный узел принимается разработчиком архитектуры на основе начальных требований к ней. При этом принимается решение о включении тех или иных методов компрессии растровых данных, либо о вынесении подсистем, реализующих видовые объекты, фактически осуществляющие отображение, на клиентские машины.

Принцип 6. Банк данных должен обеспечивать разделяемый доступ к постоянным справочным данным, необходимым при расчетах. Информационное наполнение банка составляют данные, заносимые администраторами системы и банка данных (рис. 2), и представляющие собой совокупность сведений о существующих источниках полей, таких как их энергетические характеристики, геометрические размеры; удельные параметры распространения физических полей в различных средах при различных внешних условиях; параметры, количественно описывающие поведение полей на границах раздела сред. Структура хранимых данных должна быть задана без привязки к конкретным типам данных и построена с учетом параллельного доступа к своим элементам в предположении, что этот доступ, наиболее

вероятно, сохранит имеющиеся данные неизменными.

Принцип 7. Для анализирующих подсистем определен формат вывода результатов моделирования в виде матриц распределения энергетических характеристик исследуемого поля на секущих плоскостях (рис. 2). На основе этих матриц подсистемы отображения базовых систем должны производить визуализацию распределения поля на плоскостях в виде набора изолиний.

Принцип 8. Каждому пользователю системы должна быть назначена одна из следующих ролей (рис. 2). Администратор банка данных, как указано выше, заносит необходимую информацию в банк данных. Оператор безопасности осуществляет регистрацию узлов системы на сервере безопасности. Кроме этого оператор безопасности осуществляет распределение ролей пользователей. Привилегии администратора системы объединяют привилегии оператора безопасности и администратора банка данных. Пользователи обладают минимальными правами, необходимыми для проведения анализа с использованием ограниченного оператором безопасности набора вычислительных узлов и информации банка данных, доступной на чтение.

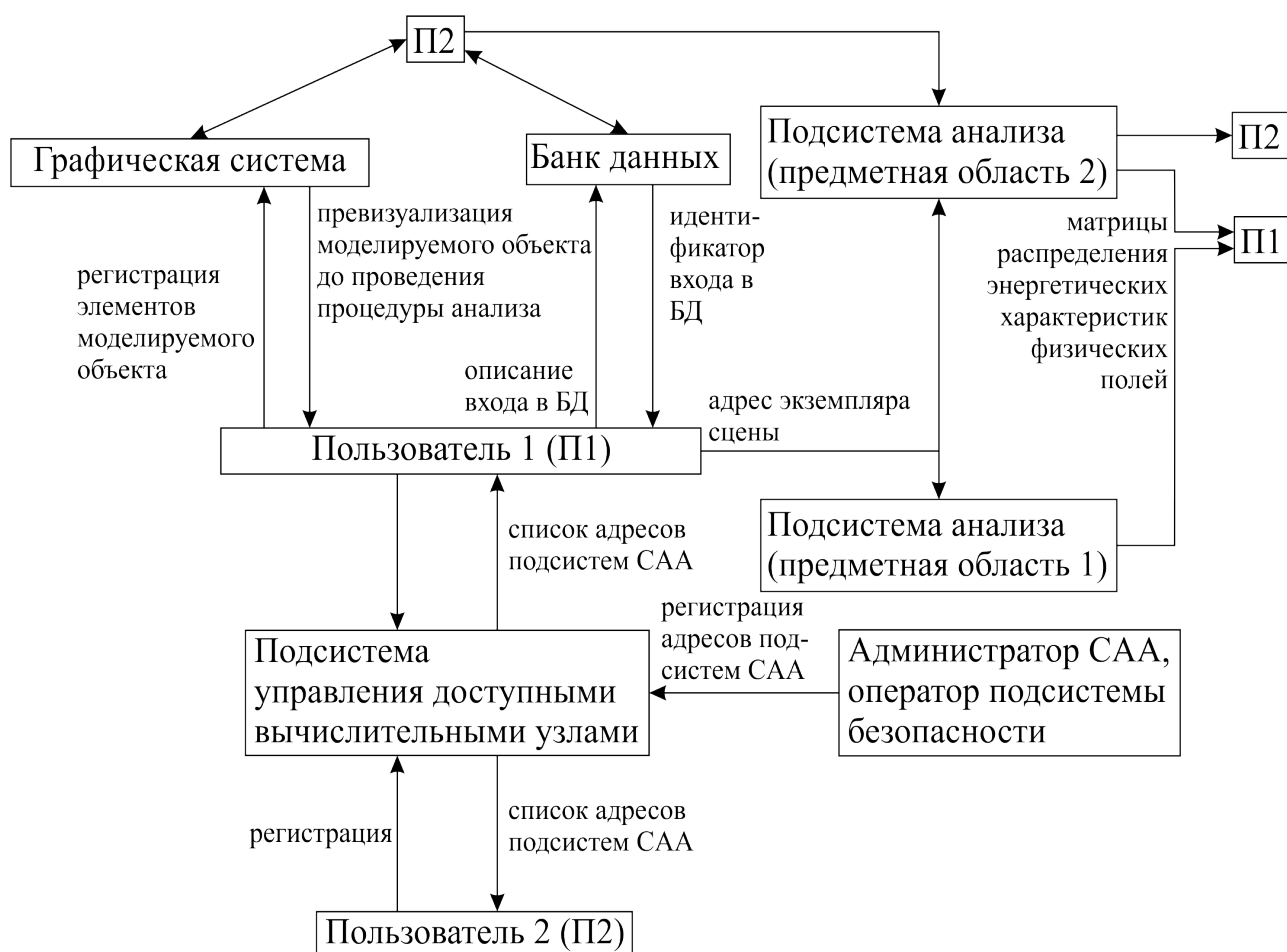


Рис. 3. Диаграмма потоков данных в системе анализа

В третьей главе описывается реализация архитектуры САА в соответствии с описанными выше принципами. Использование блочно-иерархического

системного подхода позволило максимально детализировать подсистемы, не относящиеся к предметной области, без привязки к конкретным решаемым задачам.

Диаграмма потоков данных системы анализа представлена на рис. 3. Детализированная функциональная диаграмма САА, согласно спецификации IDEF0 соответствующая рис. 2, представлена на рис. 4.

При запуске пользователем экземпляра базовой подсистемы последний по известному адресу осуществляет запрос к подсистеме управления вычислительными узлами с целью получения адресов остальных узлов, на которых развернута САА. Далее пользователь задает моделируемый объект путем создания геометрических элементов и регистрации их в графической подсистеме. Физические параметры назначаются моделируемому объекту путем присваивания его элементам идентификаторов данных банка САА. Описанная процедура выполняется функциями А1 и А3 (рис. 4), выходом которых является инициализированное моделируемое пространство, определяемое и задаваемое во время анализа компонентом сценной.

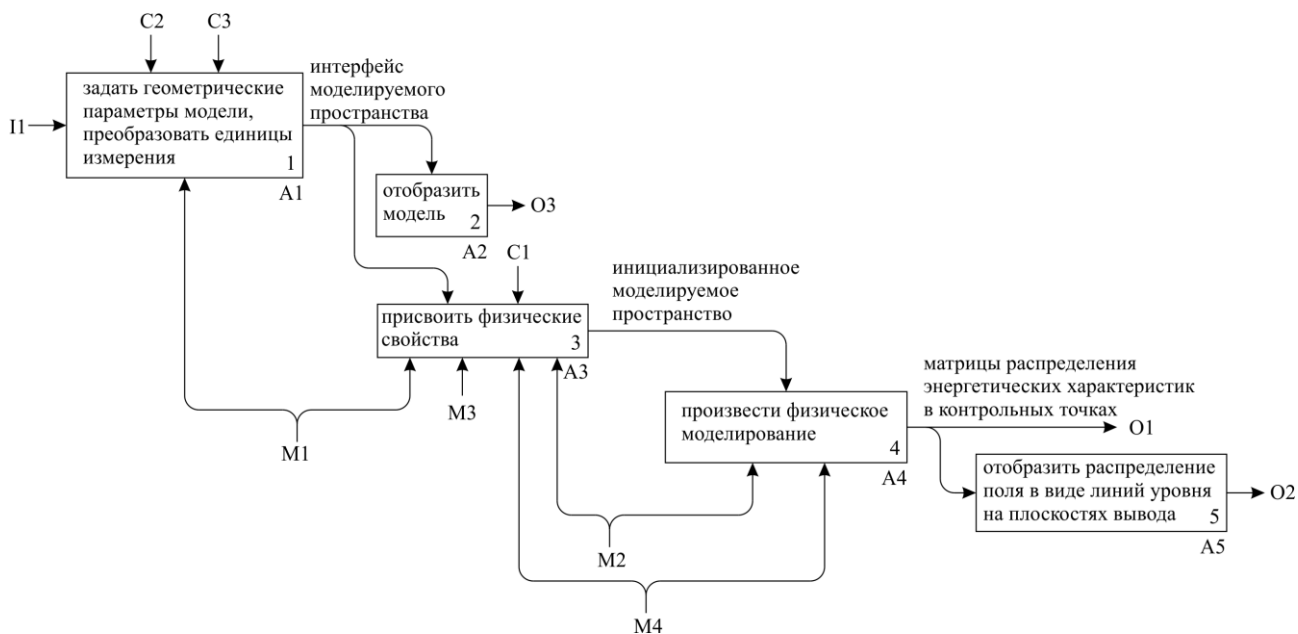


Рис. 4. Детализированная функциональная диаграмма архитектуры САА (A0)

Реализация графической системы выбрана такой, чтобы удовлетворялись требования принципа 5. При выполнении процедуры отображения используется подсистема отображения графической подсистемы, которая предоставляет функционал рисования простейших двумерных объектов в моделируемом трехмерном пространстве. Задача предоставления клиентам наиболее свежего кадра по запросу решается введением подсистемы управления буферами отображения (рис. 5). Функцией данной подсистемы является обеспечение двойной буферизации отображений с тем, чтобы в любой момент времени графическая система могла предоставить наиболее свежий готовый кадр по запросу извне. В то же время возможно проведение процедуры рисования пространства в другом (фоновом) буфере. Это обеспечивается проведением

процедуры рисования в рабочем потоке выполнения, который создается при изменении визуального состояния пространства. По завершении процедуры поток производит синхронизированную подмену буферов, и графическая подсистема через интерфейс обратного вызова уведомляет базовую систему о том, что появился свежий кадр. Если за время существования рабочего потока состояние пространства не изменилось, поток завершает выполнение.



Рис. 5. Компонент графической системы, реализующий видовой объект

Структура хранимых в банке данных, определяется подсистемой управления данными (рис. 6), которая задана на основе требований принципа 6. Фактически эта структура определена как хеш-таблица, заданная деревом,

построенным таким образом, что компоненты этого дерева обладают минимальной взаимозависимостью. Узлы этого дерева подразделяются на три иерархические категории. Корнем является супергруппа, хранящая вектор ссылок на группы. Вектор имеет фиксированный размер, равный максимальному количеству групп в файле. Элементами супергруппы являются группы, структуры которых описываются ссылкой на элементы группы и ее именем. Элементами группы являются односторонние связанные списки векторов ссылок на листья хранящегося в файле дерева элементов. Последний элемент каждого из таких векторов хранит идентификатор группы и файловый указатель на следующий вектор списка. Каждый вектор-элемент списка группы имеет фиксированный размер. Сервер банка данных отвечает за клиентский доступ к информационному фонду системы САА. Основные задачи, которые он должен при этом решать — контроль над целостностью данных, обеспечение синхронного и безопасного параллельного доступа к ним, обеспечение стойкости против несанкционированного доступа.

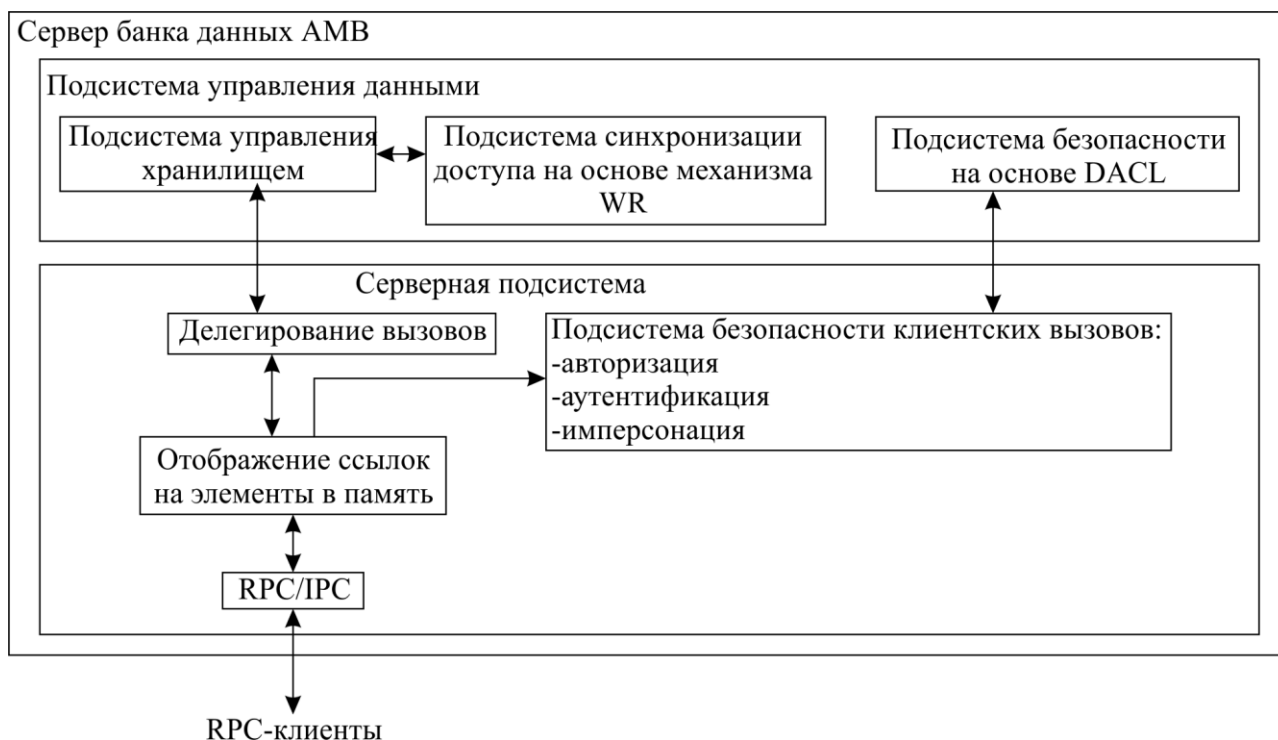


Рис. 6. Базовая структура системы управления банком данных

В четвертой главе описан механизм WR (write-read) синхронизации доступа компонентов САА к разделяемым ресурсам системы в пределах одной ЭВМ. Необходимость разработки данного механизма обусловлена, прежде всего, недостаточным набором функций, реализующих протоколы синхронизации, у существующих аналогов. Здесь же приведены количественные характеристики эффективности работы этого механизма, полученные теоретически и экспериментально.

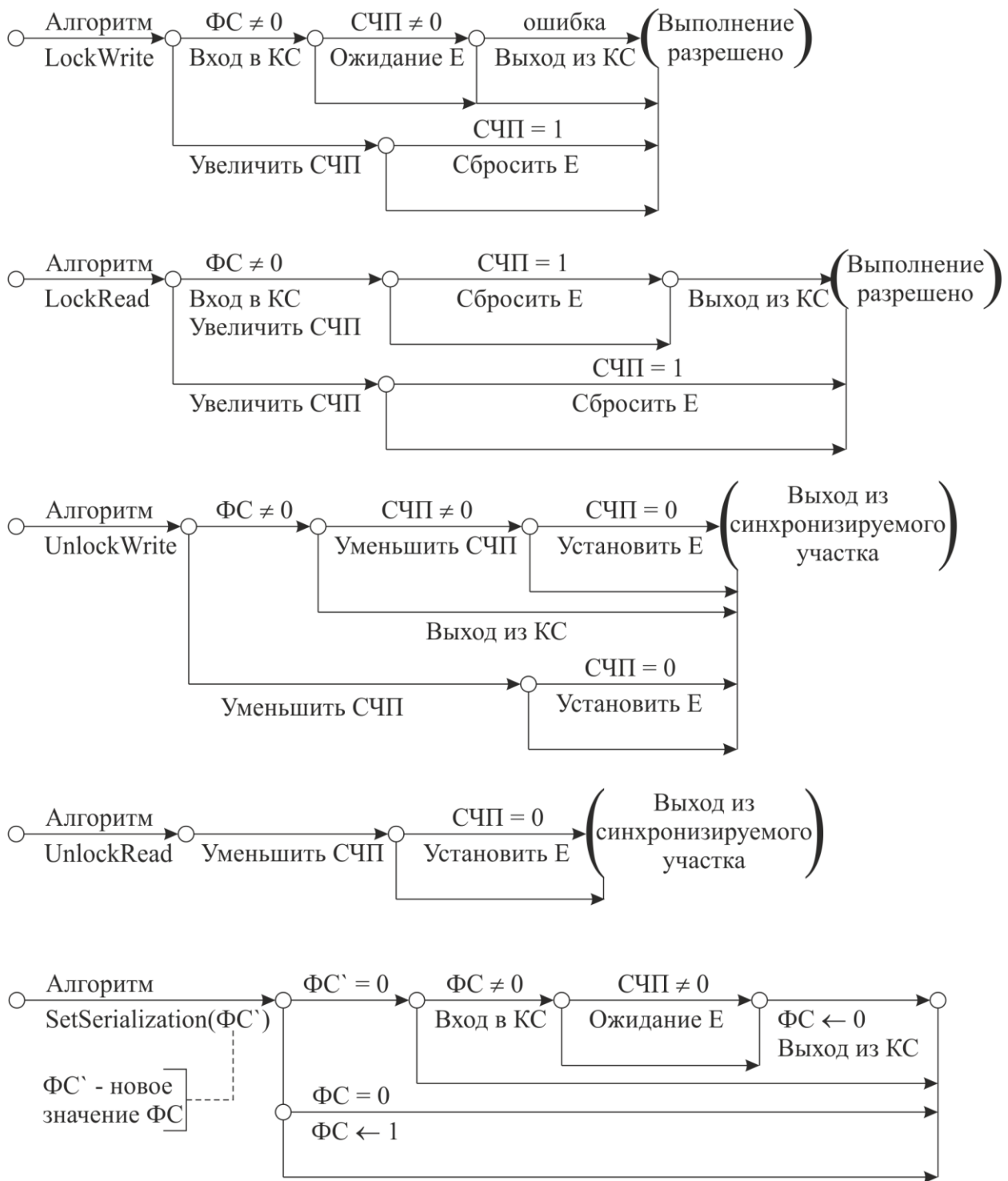


Рис. 7. Р-схемы алгоритмов синхронизации потоков WR

Все потоки выполнения, для которых применяется синхронизация WR, подразделяются на «читающие потоки» и «пишущие потоки». Задача данного механизма состоит в том, чтобы исключить параллельное выполнение пишущих потоков друг с другом, а также с читающими потоками. При этом читающим потокам должно быть разрешено выполняться параллельно. Р-схемы алгоритмов, решающих эту задачу, изображены в соответствии с единой системой программной документации (ЕСПД – ГОСТ19.005-85) на рис. 7. Используются следующие обозначения: ФС – флаг синхронизации –

фактически включает механизм синхронизации; СЧП – счетчик читающих потоков, модификации которого являются атомарными; примитивы синхронизации: КС – взаимоисключающий объект критической секции кода либо мьютекс, Е – автоматически сбрасываемое событие, при инициализации находящееся в сброшенном состоянии. Алгоритмы LockWrite и UnlockWrite организуют соответственно вход и выход в участок программного кода, называемый критическим на запись; алгоритмы LockRead и UnlockRead ограничивают участок кода, критический на чтение. Алгоритм SetSerialization реализует безопасное динамическое включение и выключение механизма синхронизации.

Второй задачей данного механизма является возможность синхронизации потоков выполнения, разнесенных по разным процессам приложения. Если имеется несколько ЭВМ, на которых производится распределенная операция, требующая синхронизации, то из них выделяется какая-либо базовая ЭВМ, т. е. сервер, который хранит и осуществляет управление разделяемым ресурсом. Далее каждому клиенту на сервере выделяется отдельный поток либо процесс, который и осуществляет непосредственный доступ к ресурсу. Данный доступ синхронизируется механизмом WR. Для удобства текст далее будет описывать ситуацию, когда несколько клиентских процессов, в рамках одной ЭВМ будут использовать один механизм WR, являющийся глобальным в этой ЭВМ. Кроме того, для удобства будет полагаться, что речь идет о платформе Windows NT4 или более новой.

Механизм может работать в двух режимах — базовом и расширенном. Базовый режим наиболее быстр, т. к. в этом случае к ядру операционной системы обращение происходит только при работе с событием. В этом случае в качестве мьютекса используется более быстрый объект критической секции кода. В расширенном режиме механизм может быть использован несколькими клиентами из разных процессов (в случае именованных экземпляров механизма). Для именованных экземпляров механизма WR действуют пространства имен объектов ядра. Также в расширенном режиме возможно задать конечный интервал времени ожидания освобождения критической секции кода на чтение и запись, а также проводить это ожидание в дежурном режиме (поддержка асинхронного вызова процедур). Состояния клиентских потоков выполнения фиксируются в структурах данных, принадлежащих клиентским процессам. Такие структуры в совокупности являются распределенной между клиентскими процессами картой, ключевыми значениями которой являются идентификаторы потоков, использующих механизм WR. Карта содержит стеки, содержащие типы блокировок, захваченных потоками. Это дает возможность, во-первых, безопасного выхода из синхронизируемых участков кода, во-вторых безопасного многократного вхождения в секции WR. Кроме того с дескриптором синхронизации связаны два значения — флаг синхронизации и счетчик читающих потоков (рис. 7). При использовании неименованных экземпляров механизма WR эти значения содержатся в структуре самого дескриптора. В случае расширенного режима с именованными экземплярами для их хранения используется разделяемая

память в виде отображения временного файла с именем, являющимся детерминированной функцией от имени объекта синхронизации.

На рис. 8 изображена структура многопоточного механизма WR, работающего в расширенном режиме, для P клиентских процессов с T_{pi} потоками, вошедшими в критическую секцию кода на чтение, в терминах Windows Multithreading.

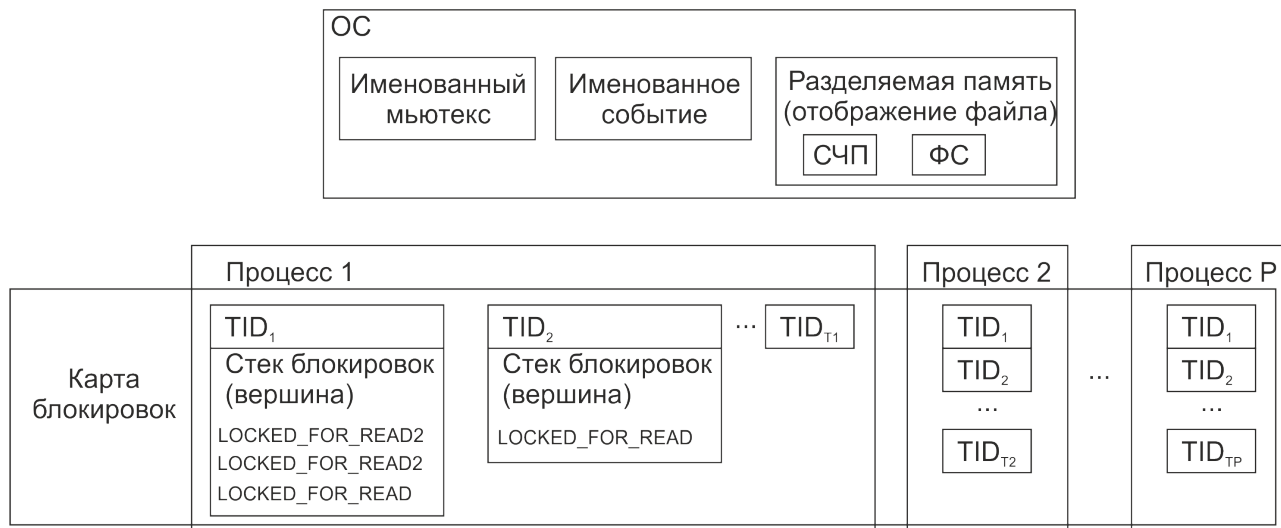


Рис. 8. Экземпляр механизма блокировки WR с $T = T_1 + T_2 + \dots + T_P$ потоками, вошедшими в синхронизируемый участок кода на чтение

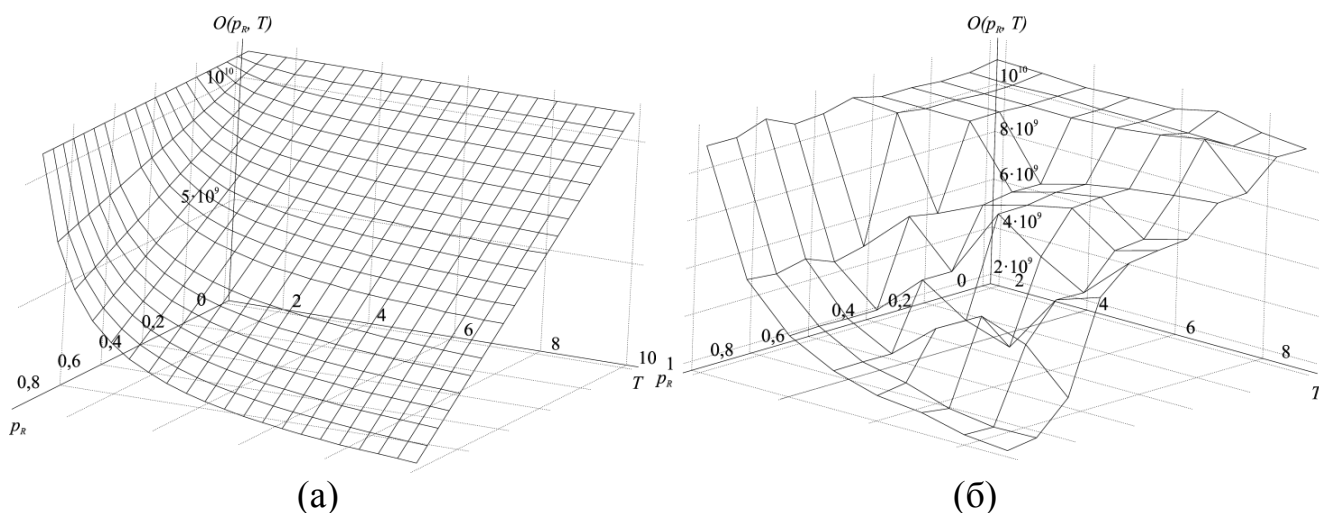


Рис. 9. Вид зависимости (1) и ее экспериментальное подтверждение для процессора Intel Core i7 860 2,6 ГГц с включенным режимом Hyper-Threading

Метод оценки эффективности механизма синхронизации WR в диссертации заключается в построении вероятностной модели вхождения потоков выполнения в критические участки кода. Пусть имеется алгоритм S , сложность последовательного выполнения которого $O_s(n)$. Пусть p_R , p_W и p_{WR} — соответственно вероятности вхождения в критические участки кода на чтение,

запись, а также вероятность того, что участок алгоритма не будет требовать синхронизации $p_R + p_W + p_{\overline{WR}} = 1$, T – число потоков выполнения. Средняя временная сложность выполнения такого алгоритма имеет вид (1).

$$O_T(p_R, p_W, T, n) = \max \left(p_W (O_S(n) + \varepsilon_W(T)) + \frac{p_R (O_S(n) + \varepsilon_R(T))}{T}, \frac{1 - p_R - p_W}{T} O_S(n) \right), \quad (1)$$

где $\varepsilon_R(T) = k_R T + C_R$ и $\varepsilon_W(T) = k_W T + C_W$ – собственные сложности алгоритмов вхождения в критические участки кода на чтение (C_R) и запись (C_W), а также время затраченное в результате состязательности единиц выполнения. В общем случае это время пропорционально числу потоков, пытающихся завладеть критической секцией кода, с коэффициентами пропорциональности k_R и k_W соответственно. Значения k_R , C_R , а также k_W и C_W определяются экспериментально отдельно для конкретной платформы. Вид зависимости (1) для случая $p_{\overline{WR}} = 0$ представлен на рис. 9.а, ее экспериментальное подтверждение приведено на рис. 9.б.

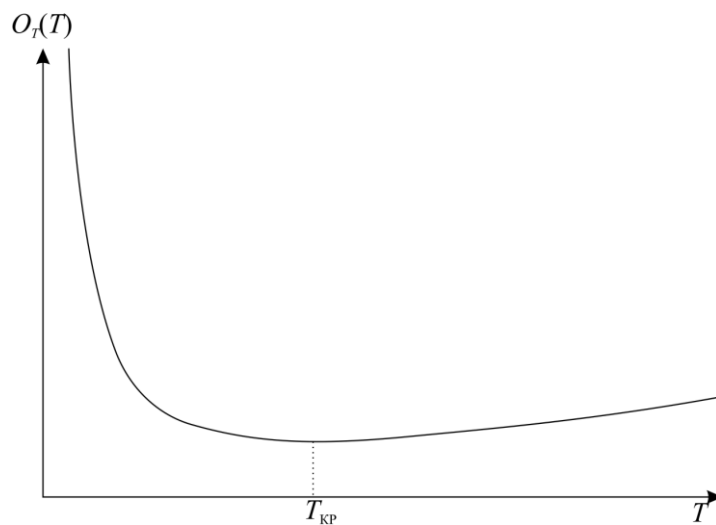


Рис. 10. Зависимость сложности параллельного алгоритма, использующего механизм WR, от количества единиц выполнения

Зависимость параллельной сложности алгоритма от количества T потоков представлена на рис. 10. Начиная с некоторого критического количества потоков T_{kp} , время выполнения алгоритма начинает повышаться за счет возрастающей роли ε_R и ε_W . Если $p_{\overline{WR}} = 0$, значение описывается формулой (2), в которой за C_R принята собственная сложность алгоритмов входа в критические участки кода на чтение.

$$T_{kp} = \sqrt{\frac{C_R + O_S p_R}{p_W k_W}} \quad (2)$$

Функционал, аналогичный базовому режиму работы механизма WR, предоставляется также средой программирования Nokia Qt - QReadWriteLock. Механизм WR при работе в базовом режиме выполняется быстрее, что показано соответствующими тестами. Их типичные результаты представлены на рис. 11 (WR – сплошная линия, Qt – пунктирная).

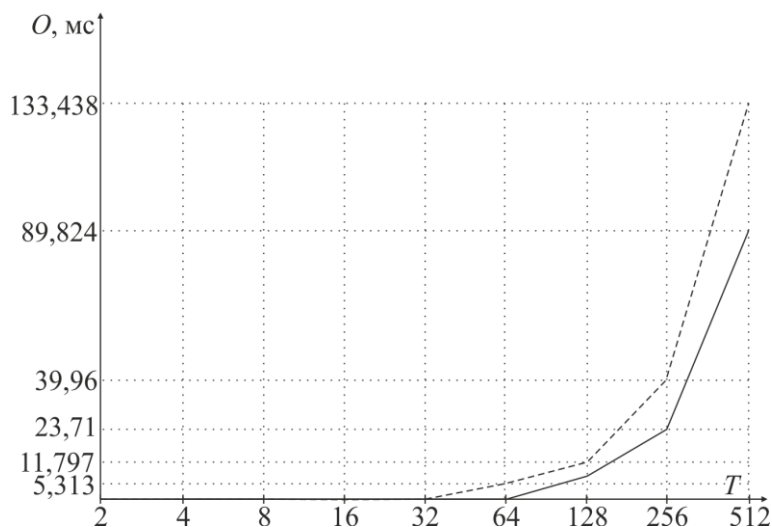


Рис. 11. Время выполнения КС алгоритма с синхронизацией WR и Nokia Qt

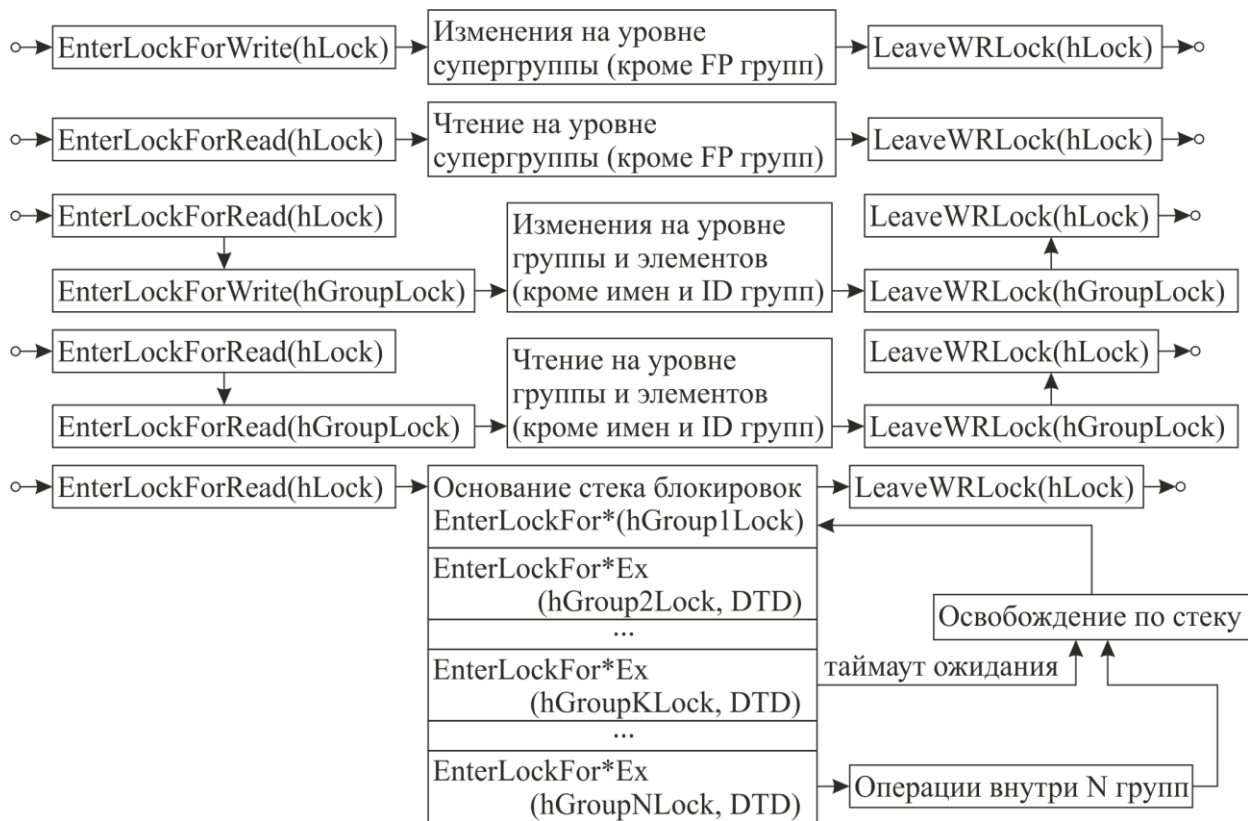


Рис. 12. Синхронизация доступа к элементам банка данных

Синхронизация параллельного доступа к описанному выше дереву банка данных осуществляется с помощью модели «читающий-пишущий» в соответствии со схемой на рис. 12. На рисунке протоколы EnterLockForRead и EnterLockForWrite реализуют вхождения в синхронизируемые участки на чтение и запись соответственно. В качестве параметров они принимают на вход дескриптор экземпляра механизма, реализующего эти протоколы. При этом hLock – дескриптор экземпляра, заданного для всей подсистемы управления данными, и обеспечивающего синхронизацию доступа к супергруппе; hGroupLock – экземпляр, обеспечивающий синхронизацию разделяемого доступа к группам, определяемый именем этой группы. Одноименные протоколы с суффиксом Ex блокируют входящий в синхронизируемый участок поток на конечное максимальное время (DTD – Deadlock Timeout Detection – определяемое эмпирически время, по истечению которого делается предположение о возникновении ситуации взаимоблокировки). Протокол LeaveWRLock осуществляет выход из синхронизируемого участка алгоритма.

В пятой главе рассматривается законченная САА Acoustic Modeler +, построенная согласно предложенным принципам, анализирующая распределение звукового поля в произвольном пространстве. Также здесь рассматривается реализация предметной области (рис. 1) с использованием многолучевого представления излучения, при котором возможно распараллеливание процесса анализа.

Подсистема анализа (рис. 1) представляет собой совокупность проектирующих подсистем, разнесенных по различным единицам выполнения, которая, используя абстрактное представление пространства, а также описанные методы распараллеливания, одновременно просчитывают различные компоненты анализируемого поля. Особенность промежуточных алгоритмов состоит в том, что выполнение задач по расчету поля по каждому лучу независимо позволяет разбить процесс моделирования на множество единиц выполнения. Во время выполнения процедуры моделирования одна из анализирующих подсистем назначается основной (управляющей). Эта подсистема разбивает задачу моделирования в многопоточную очередь из элементарных подзадач. Задачей такого элемента является расчет распространения одного луча в моделируемом пространстве и, в случае его пересечения с отражающей поверхностью, расчет параметров вторичного источника излучения. Каждая такая задача может стать источником новых подзадач. Таким образом, общий алгоритм анализа распределения физического поля в среде имеет древовидную структуру. Узлы такого дерева группируются в многопоточную очередь. При наличии суммарного количества P свободных вычислительных узлов, которые управляются подсистемой управления свободными вычислительными узлами базовой системы, P задач извлекаются и выполняются создаваемыми потоками выполнения.

Таблица 1. Время анализа звукового поля системами Acoustic Modeler + в зависимости от сложности модели пространства и количества используемых процессоров, а также системой Odeon Room Acoustics Software, в секундах

Количество граней	314	404	1509
$T_{AM+}, P = 1$	$4.7 \cdot 10^3$	$6.0 \cdot 10^3$	$2.25 \cdot 10^4$
$T_{AM+}, P = 2$	$2.3 \cdot 10^3$	$3.0 \cdot 10^3$	$1.12 \cdot 10^4$
$T_{AM+}, P = 4$	$1.2 \cdot 10^3$	$1.5 \cdot 10^3$	$5.6 \cdot 10^3$
$T_{AM+}, P = 8$	583	753	$2.8 \cdot 10^3$
$T_{AM+}, P = 24$	214	280	971
$T_{Odeon}, P = 8$	$1.44 \cdot 10^3$	$1.86 \cdot 10^3$	$6.93 \cdot 10^3$

Был проведен эксперимент по установке времени анализа, производимого представленной в диссертации САА (далее – АМ+) и наиболее близкой по сложности и методам анализа системой Odeon Room Acoustics Software (далее – Odeon). Эксперимент проводился в лаборатории телерадиовещательного комплекса ДВФУ. В качестве основной ЭВМ использовалась система с ОС Windows 7 Service Pack 1 x64 на базе четырехъядерного процессора Intel i7 860 2.8 ГГц с включенным режимом Hyper-Threading (НТТ). Система Odeon тестировалась на данной платформе без ограничений по использованию логических процессоров. Система АМ+ тестировалась с последовательным увеличением количества доступных ей процессоров в последовательности 1, 2, 4 и 8. Кроме этого система АМ+ тестировалась при использовании удаленного узла на платформе Windows Server 2008 x64 Service Pack 1 на базе четырехъядерных процессоров Intel Xeon 3.2 ГГц. Результаты эксперимента показаны в табл. 1. Здесь P – суммарное количество используемых для анализа процессоров, T_{AM+} – время затраченное предлагаемой САА, T_{Odeon} – время, затраченное системой Odeon.

РЕЗУЛЬТАТЫ РАБОТЫ

1. Разработаны общие принципы построения архитектур систем моделирования физических полей, позволяющие повысить сложность исходной модели и понизить время вычислений за счет работы в параллельной и распределенной среде. Такие системы расширяются на произвольное число вычислительных узлов, производящих процесс физического моделирования. Разработанные принципы определены с максимальной степенью детализации, что, благодаря системной блочно-иерархической структуре, дает возможность с относительной легкостью модифицировать предлагаемые архитектурные решения для анализа физических полей конкретных типов при конкретных внешних условиях.
2. Разработан метод реализации архитектур систем компьютерного анализа физических полей, основанный на представленных в диссертации принципах.

Архитектура не привязывается к конкретным предметным областям, поддерживает описанную расширяемость на произвольное множество узлов и гибкость в отношении решаемых задач.

3. Реализован механизм синхронизации единиц выполнения, позволяющий уменьшить общее время вычислений. Используемые при этом примитивы синхронизации являются широко распространенными на различных вычислительных платформах, что делает механизм переносимым. Данный механизм является более эффективным по сравнению с известными аналогами по производительности и функциональности, что подтверждается экспериментально.

4. Создан практически применимый аналитический метод вероятностной оценки эффективности применения синхронизации вычислений посредством мьютексов и посредством механизмов, реализующих модель «читающий-пишущий». Корректность данного метода подтверждена теоретическими исследованиями и экспериментально для различных платформ. В отличие от других методов оценки параллельной сложности алгоритмов, предложенный учитывает специфические для многопоточной среды издержки: состязательность единиц выполнения при их вхождении в синхронизируемые участки алгоритмов, а также собственные сложности алгоритмов, реализующих протоколы вхождения и выхода из синхронизируемых участков. В настоящее время данный метод не имеет аналогов.

5. Построена система, моделирующая и анализирующая поведение акустических полей в произвольном пространстве. Система полностью соответствует представленным принципам построения систем моделирования физических процессов и является более эффективной в плане быстродействия и точности вычислений, по сравнению с немногочисленными аналогами.

ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИИ

В изданиях, рекомендованных ВАК:

1. Параллельный поиск сигналов с заданными взаимно и автокорреляционными свойствами на многопроцессорных платформах / А.А. Чусов, А.А. Ковылин, Л.Г. Стаценко, Ю.В. Миргородская // Известия высших учебных заведений. Радиоэлектроника, г. Киев, Украина, 2011г., Т. 54, № 8, сс.29-35, ISSN 0021-3470
2. Многопоточный поиск сигналов с заданными корреляционными свойствами на SMP-системах / А.А. Чусов, А.А. Ковылин, Л.Г. Стаценко, Ю.В. Миргородская // Наука и образование: электронное научно-техническое издание, № 12, г. Москва, 2010г., Режим доступа: <http://technomag.edu.ru/doc/163224.html>, 12с. ISSN 1994-0408
3. Система эффективного кодирования информации при передаче по каналам связи в случае чрезвычайных ситуаций на море / А.А. Чусов, Л.Г. Стаценко, Ю.В. Миргородская // Известия Южного Федерального Университета, № 7 (96), 2009г., сс.200-206, ISSN 1999-9429
4. Кодирование донного изображения в канале связи автономного

необитаемого подводного аппарата (АНПА) / Чусов А.А. // Известия Южного Федерального Университета, № 9 (122), 2011г., сс.109-113, ISSN 1999-9429

Прочие публикации:

5. Шифратор речи на основе псевдослучайных последовательностей / Чусов А.А. // Молодежь и научно-технический прогресс: Материалы региональной научно-практической конференции, г. Владивосток, 2008г., Издательство ДВГТУ, 2008г., сс.165-174., ISSN 1996-3238
6. Cryptographic data protection in the GSM cellular system / Statsenko L. G., Chusov A.A. // Pacific Science Review. No. 3. Vol. 10, A publication of Korea and FESTU of Russia, 2008, pp.265-257. ISSN 1229-5450
7. Клиент-серверная система с поддержкой криптографической защиты каналов связи / Чусов А.А. // Вологодские чтения, г. Владивосток, 2008 г., Издательство ДВГТУ, сс. 74-75.
8. Криптобезопасность GSM / Чусов А.А. // Сборник трудов ДВГТУ, г. Владивосток, 2008 г., Издательство ДВГТУ, 2008г., 1с.
9. Криптографический детерминированный генератор случайных битовых последовательностей на основе односторонних хэш-функций / Чусов А.А. // Материалы конференции «Микроэлектроника и информатика-2009», г. Москва, 2009г., сс.142-143, ISBN 978-5-7256-0535-8
10. Обработка аварийной информации при передаче по каналам связи в случае чрезвычайных ситуаций на море / Чусов А.А., Стаценко Л.Г., Миргородская Ю.В. // Молодежь и научно-технический прогресс: Материалы региональной научно-практической конференции, г. Владивосток, 2009г., Издательство ДВГТУ, 2009г., 3 с., ISSN 1996-3238
11. The TCP/IP cryptographic client server system for transmitting voice and text / Chusov A.A. // Труды IX международного форума студентов, аспирантов и молодых ученых стран АТР, г. Владивосток, 2009г., pp.20-24, ISSN 1996-3238
12. Программно-алгоритмическое обеспечение «Acoustic Modeler +» для акустического расчета и озвучения помещений / Чусов А.А., Миргородская Ю.В. // Вологодские чтения, г. Владивосток, 2009г., Издательство ДВГТУ, сс.85-85, 2009г., ISSN 2219-7389
13. Система цифровой беспроводной коммуникации автономного обитаемого подводного аппарата с управляющей станцией / Чусов А.А., Стаценко Л.Г., Миргородская Ю.В. // Материалы третьей всероссийской научно-технической конференции «Технические проблемы освоения мирового океана», г. Владивосток, 2009г., сс.235-238.
14. Поиск базиса кодов с наилучшими автокорреляционными свойствами / Чусов А.А., Родионов А.Ю., Ковылин А.А., Злобин Д.В., Железняков Е.И. // Вологодские чтения, Издательство ДВГТУ, сс.70-70, г. Владивосток, 2009г., ISSN 2219-7389
15. TCP/IP client-server system for transmitting voice and text through secure communication channels / Chusov A.A. // Труды X международного форума студентов, аспирантов и молодых ученых стран АТР, г. Владивосток, 2010г., pp.98-103, ISSN 1996-3238
16. Библиотека переменных произвольной точности для моделирования

физических процессов на ЭВМ архитектур Windows-Intel x86 и x64 с использованием механизмов параллельных вычислений / Чусов А.А. // Волгдинские чтения, Издательство ДВГТУ, сс. 81-82, г. Владивосток, 2010г., ISSN 2219-7389

17. Анализ бинарных последовательностей с наилучшими автокорреляционными свойствами по ЛЧМ-подобным сигнатурам / Чусов А.А., Ковылин А.А. // Молодёжь и научно-технический прогресс: материалы региональной научно-практической конференции, в 3 ч. Ч. 1 – Владивосток: ДВГТУ, апрель-июль 2011., сс.230-231, ISSN 2072-9057

18. Архитектура и методы программной системы обработки речевого сигнала в SMP-среде / Чусов А.А., Стаценко Л.Г., Кулик С.Ю. // Сборник трудов Научной конференции "Сессия Научного совета РАН по акустике и XXIV сессия Российского акустического общества". Т. III. - М.: ГЕОС, 2011. сс.66-68

19. Parallel search for signals with specified cross- and autocorrelation properties on multiprocessor platforms / Chusov A.A., Kovylin A.A., Statsenko L.G., Mirgorodskaya Yu.V. // Radioelectronics and Communications Systems, Vol. 54, No. 8, Allerton Press, New-York, 2011., pp.425-431, ISSN 0735-2727 (print version), ISSN 1934-8061 (electronic version)

20. Использование методов распараллеленных и распределённых вычислений для оперативного прогноза уровня стохастичности поля скорости звука при проведении акустических экспериментов по дальнему распространению / Чусов А.А. // Сборник трудов Научной конференции «Сессия Научного совета РАН по акустике и XXV сессия Российского акустического общества». Т. III. - М.: ГЕОС, 2012г., 4с.

21. Результаты использования методов стохастического моделирования для расчёта зональной структуры акустического поля в двухканальном океаническом волноводе / Сальников Б.А., Сальникова Е.Н., Стаценко Л.Г., Чусов А.А. // Сборник трудов Научной конференции «Сессия Научного совета РАН по акустике и XXV сессия Российского акустического общества». Т. III. - М.: ГЕОС, 2012г., 4с.

22. Об устойчивости лучевых траекторий распространения звука в случайно-неоднородном подводном волноводе / Сальников Б.А., Сальникова Е.Н., Стаценко Л.Г., Чусов А.А. // Сборник трудов Научной конференции «Сессия Научного совета РАН по акустике и XXV сессия Российского акустического общества». Т. III. - М.: ГЕОС, 2012г., 4с.

Личный вклад автора. Все результаты, составляющие содержание диссертации, получены автором самостоятельно. В работах [12, 18, 21, 22] автору принадлежит определение общих принципов построения систем компьютерного анализа физических полей для решения задач архитектурной акустики, а также для анализа поведения поля в стохастическом подводном волноводе. В работах [1, 2, 14, 17-19] автору принадлежат определение алгоритмов, реализующих протоколы синхронизации параллельного доступа к разделяемым ресурсам, а также способов оценки эффективности использования различных методов синхронизации. В работах [5-9, 11] автору принадлежит определение алгоритмов и протоколов криптографической защиты

распределенных информационных систем. В работах [3, 4, 10, 13] автору принадлежит определение способов компрессии и канального кодирования данных, передаваемых между узлами распределенных информационных систем.

Чусов Андрей Александрович

**Гибкая архитектура для параллельного анализа и визуализации
физических полей**

Автореферат

Подписано к печати _____	Усл. печ. л. 1	Уч.-изд. л. 0,9
Формат 60x84/16	Тираж 100	Заказ ____

Издано ДВФУ. Владивосток, ул. Суханова, 8
Отпечатано издательским домом ДВФУ
Владивосток, ул. Октябрьская, 27