

На правах рукописи

Сударева Ольга Юрьевна

**ВСТРЕЧНАЯ ОПТИМИЗАЦИЯ КЛАССА  
ЗАДАЧ ТРЁХМЕРНОГО  
МОДЕЛИРОВАНИЯ ДЛЯ АРХИТЕКТУР  
МНОГОЯДЕРНЫХ ПРОЦЕССОРОВ**

Специальность 05.13.11 —

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

Автореферат

диссертации на соискание учёной степени  
кандидата физико-математических наук

Москва — 2018

Работа выполнена в Федеральном государственном учреждении «Федеральный научный центр Научно-исследовательский институт системных исследований Российской академии наук»

Научный руководитель: кандидат физико-математических наук, доцент  
**Кушниренко Анатолий Георгиевич**

Официальные оппоненты: **Горобец Андрей Владимирович**,  
доктор физико-математических наук,  
Федеральное государственное учреждение  
«Федеральный исследовательский центр  
Институт прикладной математики им.  
М. В. Келдыша Российской академии наук»,  
ведущий научный сотрудник

**Гайсарян Сергей Суменович**,  
кандидат физико-математических наук, доцент,  
Федеральное государственное бюджетное  
учреждение науки «Институт системного  
программирования им. В. П. Иванникова  
Российской академии наук»,  
заведующий отделом

Ведущая организация: Федеральный исследовательский центр  
«Информатика и управление» Российской  
академии наук

Защита состоится 24 мая 2018 г. в 17 часов на заседании диссертационного совета Д 002.087.01 при Федеральном государственном бюджетном учреждении науки «Институт системного программирования им. В. П. Иванникова Российской академии наук» по адресу: 109004, Москва, ул. А. Солженицына, 25.

С диссертацией можно ознакомиться в библиотеке и на сайте Федерального государственного бюджетного учреждения науки «Институт системного программирования им. В. П. Иванникова Российской академии наук».

Автореферат разослан «\_\_\_» \_\_\_\_\_ 2018 г..

Учёный секретарь  
диссертационного совета Д 002.087.01,  
кандидат физико-математических наук

Зеленов С. В.

# Общая характеристика работы

Диссертационное исследование нацелено на анализ ключевых факторов, влияющих на производительность расчётов на гибридной вычислительной системе (ВС), выявление, с учётом этих факторов, достоинств и недостатков отечественных многоядерных микропроцессоров, разработанных в НИИСИ РАН, и подготовку предложений по модернизации программной модели архитектуры этих процессоров.

**Актуальность** темы исследования обусловлена следующими причинами.

В настоящее время для масштабных расчётов — как в науке, так и в промышленности — повсеместно применяются параллельные ВС, собранные из розничных комплектующих мировых производителей. Наряду с ВС классической архитектуры, построенными из универсальных процессоров, в последние годы широкое распространение получили гибридные ВС, включающие универсальные процессоры (CPU) и графические ускорители (GPU). Несмотря на актуальность проблемы импортозамещения и обеспечения информационной безопасности, на данный момент отечественные технологии в этой области практически не представлены.

Имеется, однако, ряд конкурентоспособных и успешно применяемых отечественных разработок в области вычислений специального назначения. Так, линейка КОМДИВ микропроцессоров цифровой обработки сигналов (ЦОС), разработанная в НИИСИ РАН, используется в вычислительных комплексах реального времени для обработки гидроакустических и радиолокационных данных. Альтернативой для ЦОС являются системы на базе процессоров архитектуры Эльбрус разработки АО МЦСТ. Имеется и ряд универсальных процессоров с архитектурой Эльбрус; на их основе в настоящее время выпускаются персональные компьютеры и серверные системы, однако в перспективе сфера применения может быть расширена, что позволит охватить также и высокопроизводительные вычисления.

Одним из возможных путей увеличения доли отечественных разработок на внутреннем, а в перспективе и мировом, рынке, и одновременно обеспечения технической базы для научных исследований, представляется развитие имеющихся аппаратных решений, с целью дальнейшего применения в вычислениях общего назначения. Такое развитие, безусловно, должно учитывать мировой опыт в разработке процессоров и высокопроизводительных ВС.

**Степень разработанности проблемы.** Неослабевающий интерес к изучению различных аспектов высокопроизводительных вычислений (ВПВ) и высокопроизводительных ВС нашёл отражение в многочисленных исследованиях как зарубежных, так и российских авторов.

Ряд работ посвящён особенностям архитектуры конкретных ВС: видам параллелизма, типам памяти, коммуникационным средам — в частности, большое внимание уделяется энергоэффективности систем. В других ра-

ботах описываются различные инструментальные программные средства, в том числе предназначенные для автоматизации разработки и оптимизации кода.

Многие исследования посвящены конкретным классам задач и особенностям их эффективной реализации на ВС той или иной архитектуры, в том числе различным методам распараллеливания и форматам представления данных. В частности, можно отметить:

- исследование С. Уильямса и других учёных, посвящённое эффективной реализации различных процедур, применяемых в научных расчётах, на процессоре STI Cell;
- параллельную реализацию процедур из NAS Parallel Benchmarks для GPU на языке Fortran DVMH, разработанную В. Ф. Алексиным и другими российскими учёными;
- исследование А. В. Монакова, А. И. Лохмотова и А. И. Аветисяна, посвящённое эффективной реализации вычислений с разреженными матрицами на GPU;
- оптимизированную реализацию БПФ для процессоров Эльбрус, разработанную В. Е. Логиновым и П. А. Ишиным.

Что касается процессоров КОМДИВ, для них разработана библиотека цифровой обработки сигналов (БЦОС), включающая низкоуровневые оптимизированные математические процедуры. Эта библиотека и специализированные процедуры ЦОС применяются при решении реальных промышленных задач на многопроцессорных комплексах на платформе КОМДИВ. Имеются работы, в которых рассматривается программирование задач ЦОС под КОМДИВ. Однако исследований влияния специфики архитектуры гибридных процессоров НИИСИ РАН на производительность вычислительных процедур, которые используются при решении задач трёхмерного моделирования, систематически не проводилось.

**Целью диссертационной работы** является разработка методов моделирования и оценки влияния ключевых характеристик гибридной архитектуры вычислительной системы на производительность системы при решении задач трёхмерного моделирования.

Поскольку круг вопросов, связанных с производительностью гибридных ВС, очень обширен, он не может быть охвачен в одной работе. Исследование ограничено рамками нескольких типовых процедур, часто используемых в науке и промышленности, в том числе, для трёхмерного моделирования различных процессов, и гибридной вычислительной системы с сопроцессорами массивно-параллельной архитектуры.

Для достижения поставленной цели необходимо было решить следующие **задачи**.

1. Разработать метод, который позволит получать теоретические оценки ожидаемой производительности той или иной вычислитель-

ной процедуры до начала её реализации на выбранной вычислительной системе.

2. Выбрать тестовый набор процедур и применить к нему разработанный метод. Реализовать выбранные процедуры на доступных системах из рассматриваемого класса: как собранных из импортных комплектующих, так и на базе отечественных процессоров КОМДИВ. Подтвердить применимость метода, сопоставив результаты тестирования этих реализаций с теоретическими оценками, выведенными при помощи разработанного метода.
3. На основании полученных данных подготовить проект доработки программных эмуляторов, используемых в ходе разработки гибридных микропроцессоров ФГУ ФНЦ НИИСИ РАН.

#### **Научная новизна.**

1. Разработана новая модель гибридной вычислительной системы, которая позволяет для заданной вычислительной процедуры вывести теоретическую оценку производительности этой процедуры и оценить сбалансированность вычислительной системы для выполнения этой процедуры.
2. Разработан метод оценки ожидаемой производительности вычислительной процедуры на гибридной системе.
3. Выведен формальный критерий сбалансированности вычислительной системы на заданной вычислительной процедуре.
4. С помощью разработанного метода обобщены и формально обоснованы опубликованные в различных работах наблюдения о влиянии пропускной способности канала доступа к памяти на производительность вычислений на гибридных системах на базе GPU.

**Теоретическая и практическая значимость.** Разработанный метод позволяет для каждой новой вычислительной процедуры оценить ожидаемую производительность на целевой ВС. С учётом этой оценки может быть принято решение о целесообразности использования данной системы, до начала работ по реализации процедуры. С другой стороны, можно определить потенциальные преимущества каждой новой системы при решении имеющихся задач путём подстановки параметров в формулы. В дальнейшем, сопоставив реальную производительность разработанной оптимизированной процедуры с выведенной при помощи метода оценкой, можно выявить, какие особенности архитектуры не позволяют достичь теоретического максимума производительности и требуют доработки.

С помощью разработанного метода впервые проведено исследование производительности набора процедур, применяемых в трёхмерном моделировании, на гибридных процессорах КОМДИВ ВМ7 и ВМ9 оригинальной отечественной архитектуры.

Разработан и обоснован проект доработки программной модели гибридных многоядерных процессоров НИИСИ РАН с целью приближения ре-

альной производительности к ожидаемой производительности, выведенной при помощи разработанного метода. Разработка следующих поколений процессоров с учётом предложенных усовершенствований позволит добиться высокой производительности и эффективности вычислений не только на задачах ЦОС, но и на более широком классе вычислительных задач, включая задачи трёхмерного моделирования.

В ходе исследования получен опыт написания законченной иерархии кодов оптимизированных процедур для процессоров VM7/9: от вычислительных ядер на сопроцессоре CP2 до MPI-программ на управляющих процессорах. В частности, разработаны процедуры БПФ и свёртки, которые вошли в состав программного обеспечения (ПО) обработки сигналов для вычислительных комплексов реального времени и имеют производственные применения. Неулучшаемость этих процедур обоснована при помощи разработанного метода оценки ожидаемой производительности.

Реализация процедуры NPV MG для GPU, выполненная автором в ходе диссертационного исследования, оказалась более производительной, чем аналогичные процедуры, описанные в открытых публикациях.

Результаты исследования представляют интерес для специалистов в области параллельных и высокопроизводительных вычислений, в том числе на вычислительных системах гибридной архитектуры, а также для разработчиков программных эмуляторов таких систем.

**Методология и методы исследования.** Для оценки ожидаемой производительности вычислительной процедуры на гибридной ВС построена обобщённая модель гибридной ВС и разработан метод, позволяющий оценить производительность как величину, зависящую от параметров ВС и от параметров самой процедуры. На основе этого метода выведен критерий сбалансированности ВС на заданной вычислительной процедуре. Математическую основу исследования составляет теория алгоритмов.

Процедуры для гибридных систем на базе GPU были реализованы с использованием открытого стандарта OpenCL для вычислений на GPU, стандарта MPI для обменов данными между ускорителями на узле и между узлами, а также стандарта OpenMP для вспомогательных вычислений на управляющем процессоре. Эти процедуры были протестированы на гибридных узлах и кластере НИИСИ РАН, а также на суперкомпьютере K100 ИПМ им. М. В. Келдыша РАН. Процедуры для процессоров КОМДИВ VM7 и VM9 были реализованы при помощи специализированного интерфейса на языке C для управляющего процессора и языка ассемблера для сопроцессора CP2, а также стандарта MPI для обменов данными между процессорами. Эти процедуры были протестированы на эмуляторах CP2 и контроллера DMA и на существующих процессорных модулях VM7 и VM9.

**Достоверность** исследования обеспечивается тем, что рассмотренные вычислительные процедуры из набора тестов NAS Parallel Benchmarks были реализованы для гибридных систем на базе GPU, с использованием от-

крытого стандарта OpenCL для вычислений на GPU, и были протестированы на гибридных узлах и кластере НИИСИ РАН, а также на суперкомпьютере K100 ИПМ им. М. В. Келдыша РАН. Реализации процедур для процессоров VM7 и VM9 использованы в производственных приложениях.

**Основные положения, выносимые на защиту.**

1. Разработана модель гибридной вычислительной системы, охватывающая широкий класс отечественных и зарубежных архитектур.
2. Разработан метод оценки ожидаемой производительности вычислительной процедуры на гибридных вычислительных системах; выведен формальный критерий сбалансированности вычислительной системы на заданной вычислительной процедуре.
3. Применимость метода подтверждена результатами измерения производительности разработанных автором реализаций нескольких широко применяемых в трёхмерном моделировании вычислительных процедур из набора тестов NAS Parallel Benchmarks на ряде отечественных и импортных вычислительных систем.
4. При помощи разработанного метода обоснована неулучшаемость разработанных оптимизированных библиотечных процедур БПФ и свёртки для отечественных гибридных процессоров VM7 и VM9.
5. Разработан проект оптимизации архитектуры гибридных процессоров НИИСИ РАН, позволяющий за счёт локальных усовершенствований программных моделей подсистем процессора достичь существенного роста производительности на классах вычислительных задач, рассмотренных в диссертации. Проект предполагает расширение функциональных возможностей контроллера DMA и сопроцессора CP2, а также улучшение нескольких количественных характеристик сопроцессора.

**Апробация работы.** Основные результаты работы докладывались на следующих конференциях.

1. Международная конференция «High Performance Computing 2013». Киев, 7-11 октября 2013.
2. XV международная конференция «Супервычисления и математическое моделирование». Саров, 13-17 октября 2014.
3. Международная конференция «The 5th GPU Workshop — The Future of Many-Core Computing in Science 2015». Будапешт, 20-21 мая 2015.
4. 14-й Международный Междисциплинарный Семинар «Математические Модели и Моделирование в Лазерно-Плазменных Процессах и Передовых Научных Технологиях». Москва, 4-9 июля 2016.
5. Научная конференция «Ломоносовские чтения – 2017». Москва, МГУ им. М. В. Ломоносова, апрель 2017.

**Основные результаты по теме диссертации** изложены в 10 печатных изданиях [1–10]: 2 публикации в журнале, рекомендованном ВАК [1; 2],

1 монография [5], 1 свидетельство на программу для ЭВМ [6], 4 публикации в научных журналах [3; 4; 7; 8], 2 публикации в тезисах докладов [9; 10]. В работах [7; 9] вклад автора состоит в построении модели гибридной вычислительной системы, оценке производительности и реализации алгоритмов из NAS Parallel Benchmarks. Вклад автора в работах [1; 2; 10] состоит в оценке производительности алгоритмов на КОМДИВ при помощи разработанного метода, реализации оптимизированных процедур БПФ и МС, сравнительном анализе результатов тестирования и предложениях по дальнейшей оптимизации архитектуры гибридных процессоров НИИСИ РАН.

**Личный вклад.** Все представленные в диссертации результаты получены лично автором.

**Объем и структура диссертации.** Диссертация состоит из введения, четырёх глав, заключения и двух приложений. Полный объём диссертации составляет 175 страниц с 19 рисунками и 24 таблицами. Список литературы содержит 158 наименований.

## Содержание работы

### Глава 1. Производительность процедур трёхмерного моделирования на гибридных системах

В первой главе проводится обзор современных ВС, применяемых при решении задач трёхмерного моделирования, и методов их программирования. Строится модель гибридной вычислительной системы, выбирается набор вычислительных процедур, охватывающих различные классы алгоритмов, используемых в трёхмерном моделировании. Формулируется метод оценки ожидаемой производительности процедуры на гибридной ВС и критерий сбалансированности ВС на заданной процедуре.

#### 1.1 Современные высокопроизводительные вычислительные системы общего назначения и специализированные вычислительные системы разработки НИИСИ РАН

Проанализированы современные тенденции в архитектуре высокопроизводительных ВС, по данным списка Топ500 за последние 10 лет. В частности, рассмотрены такие представители гибридных архитектур, как процессор STI Cell и гибридные узлы на базе GPU, и выявлены их ключевые особенности.

Из отечественных гибридных многоядерных процессоров для рассмотрения выбраны два процессора, разработанных в НИИСИ РАН: КОМДИВ128-РИО (ВМ7) и КОМДИВ128-М (ВМ9). Оба процессора имеют гибридную архитектуру, которая изначально разрабатывалась для эффек-



тивной потоковой обработки сигналов. Важнейшими элементами этой отечественной архитектуры являются математический сопроцессор CP2, реализующий модель параллельных вычислений SIMD, и канал DMA для обменов данными между CP2 и системной памятью. Анализируются особенности архитектуры ЦОС-процессоров BM7 и BM9, находятся общие черты этой архитектуры с архитектурами гибридных систем мировых производителей, используемыми в ВПВ.

## 1.2 Методы программирования

Проведено сравнение методов программирования параллельных вычислений, которые используются для гибридных систем на базе GPU и для гибридных процессоров НИИСИ РАН. Программирование таких систем в целом подчиняется общим принципам и требует архитектурно-зависимой оптимизации на каждом этапе. В архитектуре Эльбрус применяется принципиально иной подход: процессоры являются универсальными, высокая производительность обеспечивается за счёт оптимизирующего компилятора.

## 1.3 Модель гибридной вычислительной системы

С учётом общности архитектурных решений, обоснованной в предыдущих разделах, построена модель гибридной вычислительной системы. Она описывает системы следующей конфигурации:

- системная память;
- управляющий процессор (CPU);
- один или несколько сопроцессоров одинаковой архитектуры;
- канал для обменов данными между памятью сопроцессоров и системной памятью, разделяемый между всеми сопроцессорами.

Гибридная ВС такой конфигурации характеризуется рядом параметров, среди которых наиболее важными в рамках данного исследования являются  $K$  — количество сопроцессоров — и  $BW$  — пропускная способность канала обменов данными.

Схематичное изображение системы, соответствующей модели, представлено на рисунке 1.

Система поддерживает выполнение операций четырёх типов:

- вычислительные «ядра» — вычислительные процедуры для сопроцессоров; все сопроцессоры параллельно исполняют одно и то же ядро, каждый — над данными в собственной памяти;
- загрузки данных из системной памяти в память сопроцессоров, по разделяемому каналу;
- выгрузки данных из памяти сопроцессоров в системную память, по разделяемому каналу;

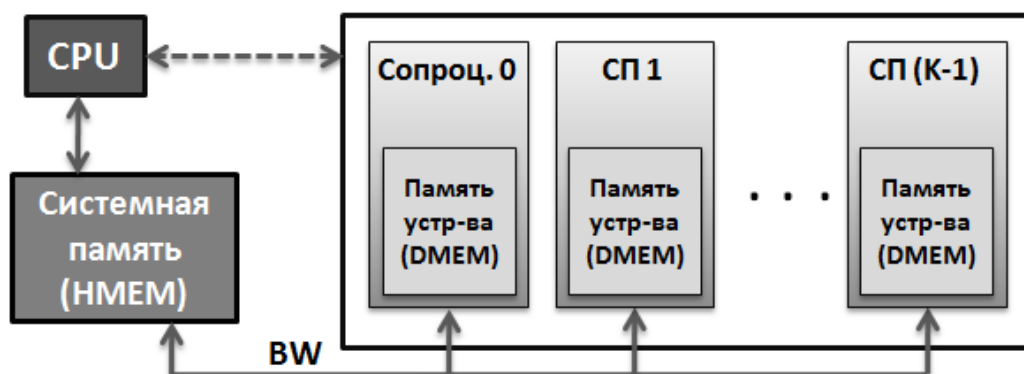


Рис. 1 — Модель гибридной вычислительной системы

– вычислительные процедуры на управляющем процессоре, над данными в системной памяти.

Как следствие, обмен данными между сопроцессорами возможен только через системную память.

При этом предполагается следующее:

- для каждой процедуры управляющий процессор устанавливает зависимости между операциями четырёх типов и запускает их на исполнение в необходимой последовательности;
- поддерживается параллельное исполнение вычислительного ядра на сопроцессорах и операции пересылки данных (загрузки или выгрузки), при условии, что ядро и операция пересылки обращаются к непересекающимся областям памяти каждого сопроцессора.

Каждой процедуре, которая выполняется на такой гибридной системе и задействует вычисления на сопроцессорах, соответствует собственный набор вычислительных ядер. Каждое ядро, в свою очередь, характеризуется параметром  $Perf_{kernel}$  — производительностью на одном сопроцессоре.

Построенная модель предназначена для теоретической оценки ожидаемой производительности заданной вычислительной процедуры на гибридной ВС. Входными параметрами при использовании модели являются:

- вычислительная процедура;
- набор значений параметров модели.

Данная модель описывает как гибридный узел с GPU, так и гибридные процессоры НИИСИ РАН с математическим сопроцессором CP2 — в частности, существующие процессоры VM7 и VM9. В первом случае сопроцессором в терминах модели является GPU, во втором — одна вычислительная секция CP2. В рамках модели может быть также рассмотрен процессор STI Cell.

Построенная модель имеет потенциал для расширения на вычислительные кластеры на базе GPU и многопроцессорные комплексы на базе гибридных процессоров НИИСИ РАН. Для этого, в частности, к введённым параметрам необходимо добавить количество узлов/процессоров, пропускную

способность коммуникационных каналов, а также учесть топологию соединения узлов.

## 1.4 Целевые вычислительные процедуры

Для рассмотрения в диссертации были выбраны три алгоритма (процедуры), каждый из которых может служить представителем одного из трёх важных широких классов: 1) алгоритмы ЦОС; 2) сеточные алгоритмы; 3) алгоритмы работы с разреженными матрицами. Все три представителя выбраны из состава распространённого тестового пакета NAS Parallel Benchmarks. Алгоритмы всех трёх классов широко применяются в самых разных предметных областях, в том числе в задачах трёхмерного моделирования. Об этом свидетельствует наличие большого количества публикаций, посвящённых как использованию алгоритмов для решения конкретных задач, так и эффективной программной реализации.

В качестве представителя класса алгоритмов ЦОС выбран алгоритм FT. Этот алгоритм сводится к трёхмерному комплексному быстрому преобразованию Фурье (БПФ). По схеме программной реализации трёхмерное БПФ аналогично двумерному, а также одномерному БПФ вектора, который не помещается целиком в память одного сопроцессора.

В качестве представителя сеточных алгоритмов выбран алгоритм MG — геометрический многосеточный метод, который сводится к последовательности разностных операторов (РО) на трёхмерных сетках различного размера.

В качестве представителя алгоритмов работы с разреженными матрицами выбран алгоритм CG — метод сопряжённых градиентов для разреженной матрицы. Он сводится к умножению разреженной матрицы на вектор (SpMV).

## 1.5 Метод оценки ожидаемой производительности

Под оптимизацией вычислительной процедуры под определённую ВС в диссертации подразумевается оптимизация процесса выполнения процедуры по числу выполняемых вычислительных операций, достижение как можно более высокой производительности вычислений, т.е. количества выполняемых арифметических операций в единицу времени. В диссертации предложен подход к исследованию влияния характеристик архитектуры на производительность оптимизированной процедуры.

Ключевым этапом исследования является построение схемы вычислений в виде последовательности элементарных вычислительных процедур, выполняемых на сопроцессорах («вычислительных ядер»), пересылок данных из системной памяти на сопроцессоры и обратно и вспомогательных операций на управляющем процессоре.

Предлагается следующий метод оценки производительности вычислительной процедуры на гибридной ВС, которая соответствует построенной модели:

1. выделить в процедуре наиболее нагруженный вычислениями этап, т.е. подпроцедуру, которая занимает основное время работы алгоритма — базовую операцию;
2. представить вычисление базовой операции как последовательность операций трёх типов:
  - загрузка данных из системной памяти на сопроцессоры,
  - запуск вычислительного ядра на сопроцессорах,
  - выгрузка данных с сопроцессоров в системную память;
3. установить зависимости между этими операциями, определить, на каких этапах может быть использовано совмещение операций пересылки данных (загрузки или выгрузки) и запуска ядра;
4. с учётом совмещения, выделить в последовательности операций основной цикл;
5. оценить производительность ядра ( $Perf_{kern}$ ) — с учётом набора инструкций и особенностей архитектуры сопроцессора (в случае, если имеются готовые реализации аналогичных ядер, вместо оценок можно воспользоваться данными об их производительности при исполнении на аппаратном стенде или потактовом программном эмуляторе);
6. оценить время работы ядра в основном цикле как количество арифметических операций на одном сопроцессоре, делённое на производительность ядра;
7. оценить время пересылок в основном цикле как общий объём пересылаемых данных для всех сопроцессоров, делённый на пропускную способность канала ( $BW$ );
8. оценить время на одну итерацию цикла: сравнить время вычислений в ядре и время пересылок, которые выполняются с совмещением, и к большему из двух прибавить время на оставшиеся операции, которые выполняются без совмещения;
9. оценить общее время работы процедуры как время на одну итерацию цикла, умноженное на количество итераций;
10. оценить производительность процедуры как общее количество арифметических операций, делённое на оценку времени работы.

Данный метод позволяет до начала реализации процедуры оценить целесообразность использования той или иной ВС для задач выбранного класса. Ключевым является этап 8, на котором в общие формулы подставляются параметры конкретной ВС и определяется, чем в действительности ограничена производительность процедуры на этой ВС: производительностью ядра на сопроцессоре или пропускной способностью канала между сопроцессорами и системной памятью.

В дальнейшем, в ходе работы над реализацией, оценки могут быть уточнены. В частности, в формулы может быть подставлена производительность реализованных ядер и скорость пересылок на конкретных схемах доступа к памяти, вместо теоретических максимумов. Сравнение таких уточнённых оценок с результатами замеров для разработанной процедуры позволяет оценить качество программного кода и выявить те существенные особенности архитектуры, которые ограничивают производительность данной процедуры и не позволяют достичь теоретического максимума производительности.

С другой стороны, формулы, отражающие зависимость между производительностью процедуры и параметрами системы, позволяют оценить ожидаемый прирост производительности за счёт изменения того или иного параметра программной модели системы и таким образом определить приоритетные направления развития архитектуры. Важно, что результаты предсказания производительности могут быть уточнены и подтверждены на раннем этапе разработки ВС, в момент завершения реализации программного эмулятора.

В случае, если при проектировании новой ВС будут учтены соотношения между её параметрами, а также выявленные ограничения в существующих архитектурах, это позволит добиться высокой производительности и эффективности вычислений на целевых классах задач. В этом состоит оптимизация архитектуры под тот или иной класс задач, как она рассматривается в данной работе.

## **1.6 Формальный критерий сбалансированности вычислительной системы на заданной вычислительной процедуре**

На основе метода оценки производительности, описанного в предыдущем разделе, введено понятие ВС, сбалансированной на заданной вычислительной процедуре, и сформулирован критерий сбалансированности.

## **Глава 2. Исследование реализаций процедур на GPU**

Во второй главе на примере современных гибридных систем с графическими ускорителями и трёх выбранных процедур обосновывается применимость предложенного метода оценки производительности.

Исследование каждой процедуры подразделяется на несколько этапов:

1. *Схема вычислений.* Для базовой операции, в соответствии с этапами 1-4 предложенного метода, построена схема вычислений на гибридной системе, в предположении, что суммарного объёма памяти сопроцессоров достаточно для хранения всех входных и выходных данных.
2. *Оценки производительности.* Выведены оценки времени работы ядра, времени пересылок и производительности всей процедуры, в

соответствии с этапами метода 5-10. Оценки рассмотрены на примере типичного гибридного узла с GPU и сопоставлены с производительностью реализаций той же процедуры на универсальных процессорах, по результатам проведённых автором диссертации замеров и по данным из открытых источников. Сделан прогноз о целесообразности реализации процедуры с использованием GPU.

3. *Практическая реализация.* Для тех процедур, для которых прогноз оказался благоприятным, автором разработаны реализации для GPU в стандарте OpenCL. Реализованы сами вычислительные ядра, а также запуск ядер и пересылок данных в необходимой последовательности, в соответствии с построенной схемой вычислений.
4. *Результаты тестирования.* Для обоснования применимости предложенного метода проведены замеры производительности разработанных процедур на доступной аппаратуре: на узлах и кластере НИИСИ РАН, включающих различные GPU, а также на суперкомпьютере K100. Результаты сопоставлены с выведенными теоретическими оценками, а также с производительностью аналогичных процедур на универсальных процессорах и с производительностью процедур, описанных в открытых публикациях.

## 2.1 Процедура NPВ FT

Проведённые оценки показали, что производительность трёхмерного БПФ больших массивов определяется пропускной способностью канала обменов данными. Сравнение ожидаемой производительности с результатами проведённого тестирования процедуры из библиотеки FFTW на универсальных процессорах показало, что использование GPU для расчёта многомерного БПФ больших массивов не позволяет получить существенный выигрыш в производительности. Справедливость этих выводов подтверждается опубликованными наблюдениями других исследователей.

## 2.2 Процедура NPВ MG

Согласно проведённым теоретическим оценкам, производительность процедуры MG определяется производительностью ядра, и можно ожидать многократного прироста производительности относительно референсной реализации на CPU. Этот вывод подтверждается результатами замеров, проведённых для разработанной автором процедуры.

Разработка ядер PO для GPU потребовала архитектурно-зависимой оптимизации, в частности, использования локальной памяти вычислителей на GPU. Реальная производительность разработанной в целях исследования процедуры MG составила 62% от теоретического максимума на одном ускорителе и 45% на четырёх ускорителях. При достаточно большом объёме

обрабатываемых данных эта процедура показывает более высокую эффективность, чем аналогичные процедуры, описанные в открытых публикациях. Ключевым фактором, который позволил добиться высокой эффективности, является использование схемы вычислений с совмещением пересылок и вычислений на сопроцессорах, которая была построена в ходе анализа производительности по предложенному методу.

### 2.3 Процедура NPВ CG

Рассмотрим подробно исследование процедуры CG на гибридном узле на базе GPU при помощи разработанного метода оценки производительности.

Базовой операцией алгоритма является вычисление SpMV — умножения разреженной матрицы на вектор. Все такие операции в алгоритме выполняются с одной и той же матрицей, причём входной вектор очередного SpMV вычисляется по выходному вектору предыдущего.

Исследование процедуры SpMV проводилось на примере формата упаковки Sliced ELLpack. В этом формате матрица нарезается на «слои» по  $H$  рядов, и каждый слой упаковывается в формате ELL; при этом все ряды в каждом слое дополняются нулями до равной длины. В упакованном виде матрица представляется в виде трёх массивов:

- $val$  — вещественный массив значений — ненулевых элементов и дополнительных нулей;
- $col$  — массив соответствующих номеров столбцов;
- $ptr$  — массив индексов начал слоёв в массивах  $val$  и  $col$ .

Процедура SpMV характеризуется следующими параметрами:

- $N$  — порядок матрицы;
- $NZ$  — количество ненулевых элементов;
- $H$  — количество рядов в слое;
- $Perf_{kern}$  — производительность ядра SpMV (вообще говоря, она зависит от портрета разреженной матрицы, т.е. расположения ненулевых элементов).

Рассмотрим вычисление SpMV на гибридной ВС, соответствующей построенной модели. Будем предполагать, что памяти каждого сопроцессора ( $DMEM$ ) достаточно для хранения входного вектора целиком, и кроме того, суммарного объёма памяти всех сопроцессоров ( $K \cdot DMEM$ ) достаточно для хранения всей матрицы, упакованной в формате Sliced ELLpack (если эти ограничения не выполняются, может использоваться постраничная схема обработки матрицы, аналогичная описанной в разделе 3.3.1 диссертации).

Схема вычисления состоит в следующем. Каждый сопроцессор обрабатывает один или несколько слоёв входной матрицы. Соответствующий набор слоёв загружается в память сопроцессора один раз перед началом вычислений. Вычисление каждого SpMV состоит из трёх шагов: загрузка входного

вектора в память каждого сопроцессора, умножение на него каждого слоя матрицы и выгрузка полученных фрагментов выходного вектора.

Для умножения любого слоя требуется весь входной вектор, поскольку индексы, по которым будут происходить обращения к нему, заранее не известны и считываются из массива *col*. Совместить эту загрузку с вычислениями невозможно. Однако можно организовать совмещение при выгрузке: слои обрабатываются по одному, и вычисления в ядре над очередным слоем совмещаются с выгрузкой предыдущего фрагмента результата.

Общее количество слоёв матрицы составляет  $S = N/H$ . Теоретические оценки производительности выводятся в предположениях, что  $S$  кратно  $K$  (количеству сопроцессоров в системе), длина массива значений (*val*) равна  $NZ$ , и на каждый слой приходится по  $NZ/S$  значений.

В описанной схеме вычислений время обработки  $K$  слоёв, по одному на каждом сопроцессоре, — большее из двух: время вычисления SpMV на одном сопроцессоре для одного слоя ( $T_C$ ) или время выгрузки  $K$  фрагментов выходного вектора со всех сопроцессоров ( $T_S$ ). Общее время вычисления SpMV — это время обработки одного слоя, умноженное на количество слоёв на каждом сопроцессоре, плюс время на загрузку  $K$  экземпляров входного вектора на сопроцессоры ( $T_L$ ).

Время на загрузку оценивается по формуле

$$T_L = \frac{KN}{BW}. \quad (1)$$

Количество ненулевых элементов в одном слое составляет  $NZ/S$ , поэтому число арифметических операций на вычисление SpMV для слоя оценивается как

$$2 \frac{NZ}{S} = \frac{2NZ \cdot H}{N}. \quad (2)$$

Отсюда,

$$T_C = \frac{2NZ \cdot H}{N \cdot Perf_{kern}}. \quad (3)$$

Один фрагмент выходного вектора содержит  $H$  элементов, поэтому время выгрузки со всех сопроцессоров оценивается как

$$T_S = \frac{KH}{BW}. \quad (4)$$

С учётом единиц измерения входящих в формулы величин (все элементы — вещественные двойной точности,  $Perf_{kern}$  — в МОП/с,  $BW$  — в МБ/с), сравнение между временами (3) и (4) приводится к виду:

$$NZ \cdot BW \quad \vee \quad 4KN \cdot Perf_{kern}. \quad (5)$$



Всего каждый сопроцессор обрабатывает  $S/K$  слоёв. Тогда, если больше значение выражения в левой части сравнения, то время обработки слоёв определяется временем вычислений, и общее время на одно SpMV оценивается по формуле:

$$T_{\text{SpMV}} \approx \frac{KN \cdot 8}{BW} + \frac{S}{K} \cdot \frac{2NZ}{S \cdot \text{Perf}_{\text{kern}}} = \frac{8KN}{BW} + \frac{2NZ}{K \cdot \text{Perf}_{\text{kern}}}. \quad (6)$$

Подстановка в формулы параметров реальных гибридных узлов с GPU и производительности существующих реализаций ядра SpMV показала, что имеет место именно эта ситуация: время на пересылку вектора пренебрежимо мало по сравнению со временем вычисления SpMV. Более того, первое слагаемое в формуле (6) вносит незначительный вклад в общую сумму, поэтому приближённо можно считать, что производительность вычисления SpMV определяется производительностью ядра на сопроцессоре и линейно масштабируется с ростом количества сопроцессоров:

$$\text{PERF}_{\text{SpMV}} \approx K \cdot \text{Perf}_{\text{kern}}. \quad (7)$$

Оценка производительности SpMV является также оценкой и для всего алгоритма CG.

Для проверки вывода о том, что производительность процедуры CG определяется именно ядром SpMV, а не пересылками данных, автором была разработана процедура SpMV в стандарте OpenCL, и на её основе — процедура CG. Разработанное ядро по производительности сопоставимо с существующими реализациями для GPU и для универсальных процессоров мировых производителей. Производительность ядра зависит от скорости нерегулярного доступа к памяти на сопроцессоре — на GPU он осуществляется через кэш-память.

Результаты замеров, проведённых для разработанной процедуры, подтверждают вывод, сделанный на основе разработанного метода оценки производительности.

## 2.4 Выводы

Разработанный метод оценки производительности позволил обобщить и дать формальное обоснование опубликованным в различных работах наблюдениям о том, что производительность большинства расчётных задач на гибридных системах на базе GPU определяется скоростью доступа к памяти на том или ином уровне: либо пересылок данных по PCI Express, либо доступа к памяти в ядре на самом GPU.

## Глава 3. Исследование реализаций процедур на гибридных процессорах КОМДИВ

В третьей главе применимость разработанного метода оценки производительности обосновывается на примере процессоров VM7 и VM9 и трёх процедур: множественного БПФ длинных векторов, MG и SpMV. Описываются реализации вычислительных ядер для CP2 и выводятся оценки их производительности, на основе программной модели CP2.

Исследование включает следующие этапы:

1. *Схема вычислений.* В связи с относительно небольшим объёмом памяти секций CP2, используются иные схемы вычислений, чем на GPU: весь объём данных постранично пересылается на CP2 и обратно.
2. *Вычислительные ядра для CP2.* Выведены оценки производительности ядер, по минимальному необходимому с точки зрения алгоритма количеству VLIW-инструкций, что соответствует этапу 5 метода оценки производительности. Здесь также учитывается совмещение вычислений и обменов данными, но на более низком уровне: в одной VLIW-инструкции совмещается арифметическая команда и команда обменов данными между памятью CP2 и регистрами. Описаны реальные ядра, разработанные автором, использованные приёмы оптимизации и возникающие накладные расходы. Производительность ядер сопоставлена с теоретической оценкой.
3. *Оценки производительности.* С использованием метода выведены оценки производительности, для предложенных схем вычислений и разработанных автором ядер.
4. *Результаты тестирования.* Разработаны процедуры, которые протестированы на реальных модулях с VM7 и VM9. Производительность по результатам замеров сопоставлена с теоретическими оценками; во внимание принимается также разница в производительности и эффективности между VM7 и VM9. Установлено, какие накладные расходы и особенности архитектуры оказывают отрицательное влияние на производительность и не позволяют достичь выведенного при помощи предложенного метода теоретического максимума производительности. В частности, проанализировано влияние работы контроллера системной памяти на скорость пересылок по DMA.
5. *Сравнение с производительностью на других процессорах.* Для оценки потенциала использования гибридных процессоров НИИСИ РАН для ВПВ проведено сравнение производительности разработанных реализаций с производительностью аналогичных процедур на универсальных процессорах Intel Xeon, по результатам замеров,

а также на ряде других процессоров, в том числе STI Cell и Эльбрус, по данным из соответствующих открытых публикаций.

### 3.1 Процедура БПФ

Разработанная библиотечная процедура БПФ длинных векторов, которые не помещаются целиком в память CP2, показывает производительность существенно ниже пиковой производительности сопроцессора CP2 — на VM9 эффективность составляет порядка 20% от пика. При этом эффективность разработанных ядер для CP2, по результатам тестирования на эмуляторе, составляет порядка 80% от пика. При помощи предложенного метода оценки производительности обосновано, что производительность процедуры БПФ ограничена пропускной способностью канала DMA. Те же соображения справедливы и для процедуры свёртки длинных векторов, которая вычисляется через БПФ.

Показано также, что на VM9 преобладание пересылок над вычислениями более значительно, чем на VM7. Это объясняется разницей в соотношениях частот процессорного ядра и системной памяти, а так же в различной архитектуре контроллеров системной памяти; от этих факторов напрямую зависит значение параметра  $BW$  (пропускной способности канала DMA) и его соотношение с производительностью ядра. Тем не менее, пропускной способности оказывается достаточно, чтобы достичь высоких показателей производительности, в сравнении с другими процессорами. Из всех рассмотренных процессоров только ЦОС-процессор TMS320C6678 демонстрирует, в пересчёте на одно процессорное ядро и один канал доступа к памяти, более высокую производительность на процедуре БПФ, чем VM9.

Дальнейшее повышение производительности разработанной процедуры БПФ возможно за счёт увеличения пропускной способности канала пересылки данных. Кроме того, увеличение объёма локальной памяти CP2 позволит обрабатывать более длинные векторы по более эффективной схеме, в которой вычисления в ядре преобладают по времени над пересылками.

### 3.2 Процедура NPВ MG

Оценки производительности в соответствии с предложенным методом показали, что время работы разработанных ядер разностных операторов в процедуре MG близко ко времени соответствующих пересылок: например, на VM7 на операторе невязки преобладают вычисления в ядре, а на VM9 — пересылки. Справедливость этого вывода подтверждается результатами проведённых замеров: эффективность вычисления невязки на VM7 составляет 82% от теоретического максимума, а на VM9 — 58%.

В нынешней версии процессоров производительность ядер РО ограничена особенностями архитектуры CP2, главным образом, отсутствием под-

держки 32-разрядной адресации, что влечёт различные накладные расходы. При пересылках данных также возникают накладные расходы, связанные с требованиями контроллера DMA к выравниванию и, как следствие, пересылкой избыточного объёма данных. Дальнейшее повышение производительности MG требует расширения функциональных возможностей как CP2, так и контроллера DMA. На данный момент производительность MG на VM9 составляет порядка 50% от производительности на одном ядре Intel Xeon E5-2670v1 и порядка 93% от производительности на одном ядре Эльбрус.

Для процедуры MG была также разработана реализация с использованием MPI для многопроцессорного комплекса. В настоящий момент поддержка стандарта MPI для коммуникаций в многопроцессорных комплексах на базе КОМДИВ находится на стадии разработки. Для оценки производительности разработанной процедуры MG построенная модель гибридной системы и метод оценки производительности были расширены: рассматривалось также количество процессоров в комплексе и пропускная способность коммуникационных каналов. Проведённые таким образом оценки показали, что пропускной способности сети RapidIO достаточно для эффективных расчётов по алгоритму MG.

### 3.3 Процедура SpMV

В реализации для CP2 использовался формат упаковки на основе CSR («compressed sparse row»), который более удобен для вычислений на CP2, чем формат Sliced ELLpack, использованный в реализации для GPU. Производительность ядра, как и на GPU, ограничена скоростью чтения данных на сопроцессоре в не известном заранее порядке: архитектура CP2 в нынешней версии не позволяет реализовать такой доступ эффективно. Помимо отсутствия поддержки 32-разрядной адресации, сопроцессор CP2 не позволяет осуществлять доступ к данным четырёх вычислительных секциях одновременно по разным адресам; основные накладные расходы в ядре связаны с этими архитектурными ограничениями.

Проведённые оценки показали, что время вычислений в ядре превышает время пересылок, что подтверждается и результатами замеров. Таким образом, повышение производительности на процедуре SpMV требует, в первую очередь, доработки архитектуры CP2. В настоящий момент на процедуре SpMV процессор VM9 уступает одному ядру Intel Xeon в несколько раз.

### 3.4 Выводы

Рассмотрение трёх вычислительных процедур, применяемых в задачах трёхмерного моделирования, на гибридных процессорах НИИСИ РАН подтвердило применимость предложенного метода оценки производительности. Сопоставление реальной производительности с теоретическими оценка-

ми в ходе реализации процедур позволило выявить особенности архитектуры процессоров, которые ограничивают реальную производительность этих процедур.

## Глава 4. Рекомендации по дальнейшему развитию архитектуры гибридных многоядерных процессоров НИИСИ РАН

В четвёртой главе формальный критерий сбалансированности ВС на заданной процедуре, сформулированный в разделе 1.6, конкретизируется для гибридных процессоров НИИСИ РАН и рассмотренных в диссертации процедур. Проводится анализ особенностей архитектуры гибридных процессоров, которые не позволяют достичь теоретического максимума производительности. Предлагается проект доработки архитектуры процессоров.

### 4.1 Критерий сбалансированности процессоров НИИСИ РАН на выбранном классе задач

В последней главе рассматривается частный случай гибридной ВС — перспективный гибридный микропроцессор НИИСИ РАН, архитектура которого является развитием архитектуры VM7/9 и имеет аналогичную программную модель. Соответственно, предполагаются аналогичные ядра для сопроцессора CP2 и общие схемы вычислений. Аналогичные рассуждения можно провести и для других гибридных систем, соответствующих построенной модели.

Сформулируем критерий сбалансированности гибридных процессоров НИИСИ РАН на каждой из рассмотренных процедур: БПФ, MG и SpMV, — они представляют классы задач ЦОС, вычислений на сетках и вычислений с разреженными матрицами. В соотношения входят три ключевых параметра:

- $F$  — тактовая частота сопроцессора, которая совпадает с частотой процессорного ядра и выражается в МГц;
- $K$  — количество секций сопроцессора;
- $BW$  — пропускная способность канала доступа к системной памяти через DMA, в МБ/с.

На основе формул для оценки производительности, выведенных в главе 3, в том числе формул для оценки производительности ядра, получены следующие соотношения на параметры сбалансированной системы:

1. для множественного БПФ длины  $N \leq 4096$  ( $N = 2^n$ ):

$$BW \log N = 32KF; \quad (8)$$

2. для процедуры MG:

$$31 BW = 64KF; \quad (9)$$

3. для процедуры SpMV:

$$BW = \frac{16F}{2K + 1}. \quad (10)$$

В этих формулах  $BW$  — максимальная пропускная способность, достижимая с учётом всех характеристик подсистемы памяти на пересылках того типа, которые используются в процедуре.

Все формулы были выведены в предположении, что вычисления выполняются с одинарной точностью. Если рассматривать поддержку вычислений с двойной точностью в архитектуре сопроцессора, в формулы необходимо внести соответствующие поправки.

## 4.2 Проект оптимизации архитектуры гибридных процессоров НИИСИ РАН

В предыдущих главах были выведены теоретические оценки производительности процедур, которые, в частности, зависят от сбалансированности системы. Вместе с тем, реальная производительность разработанных процедур не достигает теоретического максимума, хотя и соответствует ему по порядку — это объясняется не только наличием накладных расходов в программных реализациях процедур, но и качественными особенностями архитектуры процессоров. Существенного повышения производительности на рассмотренных классах задач можно добиться за счёт усовершенствований программных моделей отдельных подсистем процессора.

В данном разделе проанализированы реализации алгоритмов, описанные в главе 3, и выявлены архитектурные особенности гибридных процессоров НИИСИ РАН, которые не позволяют достичь теоретического максимума производительности. Описаны перспективы оптимизации процедур при условии внесения соответствующих изменений в архитектуру. Кроме того, проведено сравнение функциональных возможностей, предоставляемых графическими ускорителями и использованных в реализациях, описанных в главе 2, с возможностями сопроцессора CP2.

По итогам проведённого анализа предлагается проект оптимизации архитектуры процессоров, который предполагает следующие усовершенствования:

- возможность обменов между локальной памятью и регистрами CP2 32-разрядными словами;
- отдельные наборы адресных регистров в секциях CP2;
- инструкции условной записи/чтения из локальной памяти CP2;
- отдельные счётчики циклов в секциях;
- общая память на CP2, разделяемая между секциями;
- доступ на запись к биту, определяющему направление БПФ, со стороны CP2;

- возможность пересылки по DMA произвольного набора 32-разрядных слов;
- увеличение количества вещественных регистров FPR;
- увеличение объёма локальной памяти CP2.

Настоящие предложения могут также рассматриваться как раздел технического задания на разработку потактового эмулятора подсистемы CP2 и контроллера DMA перспективных гибридных микропроцессоров НИИСИ РАН.

Предварительные оценки показывают, что в совокупности эти усовершенствования способны обеспечить прирост производительности на процедурах БПФ и MG в 3 раза, а на процедуре SpMV — в 6 раз. Эти оценки могут быть уточнены ещё до разработки реального процессора, на этапе разработки эмулятора.

#### **4.3 Достоинства и недостатки гибридных процессоров НИИСИ РАН в контексте высокопроизводительных вычислений**

В данном разделе отмечены различия в архитектурах гибридных узлов с GPU и гибридных процессоров с сопроцессором архитектуры CP2, которыми обусловлена разница в используемых схемах вычислений и в производительности готовых процедур.

Проведённое исследование наглядно продемонстрировало, что одним из ключевых достоинств гибридных процессоров НИИСИ РАН является высокая пропускная способность канала DMA — это позволяет эффективно реализовать даже схемы вычислений, связанные с пересылками больших объёмов данных. Вместе с тем, сопроцессор CP2 изначально спроектирован для задач потоковой обработки сигналов, и его архитектура имеет ряд особенностей, которые затрудняют реализацию задач других классов. Многие из этих особенностей, в том числе, замена кэш-памяти контролируемой программно локальной памятью, присущи и другим архитектурам, которые успешно применяются в ВПВ.

Таким образом, обоснован вывод, что при условии дальнейшего развития архитектуры, гибридные процессоры НИИСИ РАН имеют хороший потенциал для использования в ВПВ.

## **Основные результаты диссертационной работы**

1. Разработана модель гибридной вычислительной системы, охватывающая широкий класс отечественных и зарубежных архитектур.
2. Разработан метод оценки ожидаемой производительности вычислительной процедуры на гибридных вычислительных системах; вы-

веден формальный критерий сбалансированности вычислительной системы на заданной вычислительной процедуре.

3. Применимость метода подтверждена результатами измерения производительности разработанных автором реализаций нескольких широко применяемых в трёхмерном моделировании вычислительных процедур из набора тестов NAS Parallel Benchmarks на ряде отечественных и импортных вычислительных систем.
4. При помощи разработанного метода обоснована неулучшаемость разработанных оптимизированных библиотечных процедур БПФ и свёртки для отечественных гибридных процессоров VM7 и VM9.
5. Разработан проект оптимизации архитектуры гибридных процессоров НИИСИ РАН, позволяющий за счёт локальных усовершенствований программных моделей подсистем процессора достичь существенного роста производительности на классах вычислительных задач, рассмотренных в диссертации. Проект предполагает расширение функциональных возможностей контроллера DMA и сопроцессора CP2, а также улучшение нескольких количественных характеристик сопроцессора.

## Публикации автора по теме исследования

1. *Богданов П. Б., Сударева О. Ю.* Применение отечественных специализированных процессоров семейства КОМДИВ в научных расчётах // Информационные Технологии и Вычислительные Системы. — 2016. — Т. 3. — С. 45–65.
2. *Богданов П. Б., Сударева О. Ю.* Производительность процессоров КОМДИВ на ряде типовых расчётных задач // Информационные Технологии и Вычислительные Системы. — 2017. — Т. 4. — С. 104–111.
3. *Сударева О. Ю.* Реализация алгоритма MG из пакета NPВ для многопроцессорного вычислительного комплекса на базе микропроцессора КОМДИВ128-РИО // Труды НИИСИ РАН. — 2015. — Т. 5, № 1. — С. 75–87.
4. *Сударева О. Ю.* Развитие микропроцессоров линейки КОМДИВ для применений в научных расчётах: предложения по оптимизации архитектуры // Современные научные исследования и разработки. — 2018. — № 2(19). — С. 295–301. — URL: [http://olimpiks.ru/d/1340546/d/zhurnal\\_219.pdf](http://olimpiks.ru/d/1340546/d/zhurnal_219.pdf) (дата обращения: 15.03.2018).
5. *Сударева О. Ю.* Эффективная реализация алгоритмов быстрого преобразования Фурье и свёртки на микропроцессоре КОМДИВ128-РИО. — М.: НИИСИ РАН, 2014. — 266 с.



6. Свидетельство о государственной регистрации программы для ЭВМ № 2017617058. Библиотека цифровой обработки сигналов для микропроцессора КОМДИВ128-РИО для ОС RV Багет 3.5 (БЦОС 3.5) / Г. О. Райко, О. Ю. Сударева, М. С. Хропов, М. С. Аристов / Правообладатель: Федеральное государственное учреждение «Федеральный научный центр Научно-исследовательский институт системных исследований Российской академии наук» (ФГУ ФНЦ НИИСИ РАН). — 22 июня 2017.
7. *Богданов П. Б., Сударева О. Ю.* Гетерогенное программирование в рамках стандарта OpenCL // Супервычисления и математическое моделирование: труды XV Международной конференции, 13-17 октября 2014 г. / под ред. Р. М. Шагалиева. — Саров: ФГУП «РФЯЦ ВНИИЭФ», 2015. — С. 123–137.
8. *Сударева О. Ю.* Распределённые вычисления на процессорах КОМДИВ на примере алгоритма NPV MG // Наука нового времени: сохраняя прошлое – создаём будущее. Сборник научных статей по итогам международной научно-практической конференции, г. Санкт-Петербург, 22-23 декабря 2017. — СПб: Изд-во «КультИнформПресс», 2017. — С. 60–63.
9. *Bogdanov P., Efremov A., Sudareva O.* Heterogeneous programming methodology based on OpenCL framework // Proceedings of High Performance Computing 2013, Kyiv, October 7-11. — 2013. — Pp. 391–392. — URL: <http://hpc-ua.org/hpc-ua-13/files/proceedings/74.pdf> (access date: 23.08.2017).
10. *Богданов П., Сударева О.* Применение отечественных специализированных процессоров семейства «КОМДИВ» в научных расчётах // Тезисы 14го Международного Междисциплинарного Семинара «Математические Модели и Моделирование в Лазерно-Плазменных Процессах и Передовых Научных Технологиях», г. Москва, 4-9 июля. — 2016. — URL: <http://lppm3.ru/files/histofprog/LPpM3-2016-1-Programme.pdf> (дата обращения: 23.08.2017).

*Сударева Ольга Юрьевна*

Встречная оптимизация класса задач трёхмерного моделирования для архитектур многоядерных процессоров

Автореф. дис. на соискание учёной степени канд. физ.-мат. наук

Подписано в печать \_\_\_\_ . \_\_\_\_ . 2018. Заказ № \_\_\_\_\_

Формат 60×90/16. Усл. печ. л. 1. Тираж 100 экз.

Типография \_\_\_\_\_