

На правах рукописи



Винс Дмитрий Владимирович

**ИМИТАЦИОННОЕ МОДЕЛИРОВАНИЕ
УПРАВЛЯЮЩИХ И ВЫЧИСЛИТЕЛЬНЫХ ПРОЦЕССОВ В
ВЫСОКОПРОИЗВОДИТЕЛЬНЫХ ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМАХ**

Специальность 05.13.15 – Вычислительные машины, комплексы
и компьютерные сети

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата технических наук

Новосибирск – 2016

Работа выполнена в Федеральном государственном бюджетном учреждении науки Институте вычислительной математики и математической геофизики Сибирского отделения Российской академии наук (ИВМиМГ СО РАН).

Научный руководитель: доктор технических наук, профессор

Глинский Борис Михайлович

Официальные оппоненты:

Окольнишников Виктор Васильевич,

доктор технических наук, Конструкторско-технологический институт вычислительной техники СО РАН (г. Новосибирск), лаборатория информационных систем, ведущий научный сотрудник.

Ефимов Александр Владимирович,

кандидат технических наук, доцент по кафедре вычислительных систем, ФГБОУ ВО СибГУТИ, ФГБУН Институт физики полупроводников СО РАН (г. Новосибирск), лаборатория вычислительных систем, ведущий инженер-электроник.

Ведущая организация

Федеральное государственное бюджетное образовательное учреждение высшего образования «Московский государственный университет имени М.В. Ломоносова»

Защита состоится «1» декабря 2016 г. в 14⁰⁰ часов на заседании диссертационного совета Д 219.005.02 при федеральном государственном бюджетном образовательном учреждении высшего образования «Сибирский государственный университет телекоммуникаций и информатики».

С диссертацией можно ознакомиться в библиотеке и на сайте федерального государственного бюджетного образовательного учреждения высшего образования «Сибирский государственный университет телекоммуникаций и информатики» <http://www.sibsutis.ru>.

Автореферат разослан « ____ » _____ 2016 г.

Ученый секретарь

диссертационного совета Д 219.005.02

кандидат технических наук, доцент

Резван И.И.

Общая характеристика работы

Актуальность работы

В настоящее время существует большое количество задач, требующих применения мощных вычислительных систем, среди которых задачи моделирования работы систем массового обслуживания и управления, распределённых систем, сложные вычислительные задачи и др. Для решения таких задач всё шире используются гетерогенные кластеры, которые становятся общедоступными и относительно недорогими аппаратными платформами для высокопроизводительных вычислений.

Мощности и возможности вычислительных систем также с каждым годом возрастают. По расчетам, приведенным Д. Донгарра¹, производительность в эксафлопс будет достигнута к 2018 – 2020 году. Суперкомпьютер к тому времени сможет обслуживать одновременно 1 млрд потоков вычисления. Количество ядер у него будет 10-100 млн.

Создание эксафлопсных суперкомпьютеров идет по двум направлениям: эволюционному и инновационному. Первое предполагает постепенное улучшение существующих суперкомпьютеров на легких и тяжелых суперскалярных процессорных ядрах. Подход с тяжелыми ядрами более доступен и популярен в сочетании с применением графических сопроцессоров (GPU), сопроцессоров типа Xeon Phi и ускорителей на программируемых логических матрицах (FPGA).

Задача инновационных проектов — поиск новых решений и технологий для создания эффективных и отказоустойчивых суперкомпьютеров с реальной эксафлопсной производительностью на широком круге задач. По ряду причин такие технологии называют экзамасштабными, и одно из их главных свойств — обеспечение удельной энергетической эффективности в 50 GFLOPS/Вт, что в десятки раз ниже, чем у нынешних суперкомпьютеров².

До недавнего времени наиболее заметны были результаты работ эволюционного направления, но сегодня в новейших образцах суперкомпьютеров уже просматриваются инновационные элементы будущих эксафлопсных гигантов, стали заметны некоторые общие черты новых систем, осознана сложность инновационного направления по преодолению физических ограничений имеющейся элементно-конструкторской базы и созданию новой, пост-Муровской эпохи, проведены мероприятия по расширению и улучшению организации соответствующих работ.

Если с аппаратной точки зрения основные направления и перспективы как-то обозначены, то вопрос программного обеспечения таких огромных систем все еще не нашел отражения. Предполагается, что для таких суперкомпьютеров программное обеспечение будет разрабатываться уже после их создания, а на первых этапах будет применяться уже существующее параллельное системное и прикладное программное обеспечение, но какую эффективность оно покажет достоверно предугадать сложно.

Проверить эффективность существующего программного обеспечения при выполнении на миллионах вычислительных ядер представляется возможным только посредством проведения имитационного моделирования. Именно имитационное моделирование выполнения алгоритмов и программ позволяет дать качественную оценку эффективности управляющих и вычислительных алгоритмов и программ при выполнении на будущих эксафлопсных суперкомпьютерах, что и является предметом исследования настоящей диссертации.

¹ HPC Challenge: Design, History, and Implementation Highlights, J. Dongarra and P. Luszczek, Ed. Jeffrey S. Vetter, Contemporary High Performance Computing: From Petascale Toward Exascale, Taylor and Francis, Boca Raton, CRC Computational Science Series, 2013

² В.С.Горбунов, Л.К.Эйсмонт, А.В.Речинский, В.С.Заборовский, П.В.Забеднов. Суперкомпьютеры для промышленности – вопросы тестирования, анализа и разработки. Материалы 2-й Всесоюзной конференции “Суперкомпьютерные технологии” (СКТ-2012), стр.360-364.

Цель работы разработка методов исследования эффективности использования перспективных и существующих управляющих и вычислительных алгоритмов для высокопроизводительных суперкомпьютеров на основе имитационных моделей.

Задачи исследования

1. Анализ и разработка подходов к имитационному моделированию организации функционирования ВС в мультипрограммных режимах.
2. Экспериментальные исследования по определению эффективности существующих алгоритмов управления с использованием разработанного имитационного подхода.
3. Анализ и разработка методов исследования эффективности вычислительных алгоритмов и программ при их отображении на архитектуру высокопроизводительных суперкомпьютеров методами имитационного моделирования.
4. Экспериментальные исследования по определению эффективности реализаций некоторых существующих параллельных алгоритмов и программ с использованием разработанного метода.

Методы исследования

1. Имитационное моделирование с использованием системы мультиагентного моделирования AGNES³ процесса оперативного управления ВС для исследования управляющих алгоритмов и процесса выполнения параллельных программ на ВС для исследования вычислительных алгоритмов.
2. Статистическая обработка данных реальной работы алгоритмов и модельного эксперимента с использованием электронных таблиц Microsoft Excel.

Научная новизна работы заключается в том, что в ней впервые

1. Разработана и исследована имитационная модель *оперативного* управления ресурсами распределенной вычислительной системы, базирующаяся на агентно-ориентированном подходе.
2. Экспериментально обоснован набор алгоритмов управления, позволяющий уменьшить время ожидания в очереди до 10%.
3. Разработан метод исследования эффективности выполнения параллельных алгоритмов и программ на современных и перспективных высокопроизводительных суперкомпьютерах с использованием имитационного моделирования.
4. Произведена оценка эффективности использования нескольких вычислительных алгоритмов на большом числе ядер высокопроизводительных суперкомпьютеров. Для некоторых предложены способы ее повышения.

Личный вклад автора состоит

1. В разработке формального описания, алгоритма и программной реализации программных агентов моделей оперативного управления высокопроизводительными ВС и моделей исполнения параллельных программ на высокопроизводительных ВС.
2. В выполнении модельного эксперимента, статистической обработке и анализе полученных результатов.
3. В совместных публикациях конфликта интересов нет.

Научное и практическое значение работы

Разработанная в диссертации имитационная модель и ее программная реализация развивают перспективные методы исследования организации функционирования

³ Podkorytov, D., Rodionov, A., Choo, H.: Agent-based simulation system AGNES for networks modeling: review and researching. Proc. of the 6th Int. Conference on Ubiquitous Information Management and Communication (ACM ICUIMC 2012), ISBN 978-1-4503-1172-4, Paper 115, 4 pages, ACM (2012).

высокопроизводительных вычислительных систем в мультипрограммных режимах при решении параллельных вычислительных задач.

Разработанные в диссертации модели предназначены для исследования эффективности использования управляющих и вычислительных алгоритмов на существующих и перспективных высокопроизводительных ВС. Разработанная модель оперативного управления ВС позволяет исследовать эффективность различных наборов управляющих алгоритмов при большом количестве вычислительных ресурсов. Разработанный метод исследования вычислительных параллельных алгоритмов позволяет обозначить границы эффективного применения того или иного алгоритма, а также дать рекомендации по улучшению масштабируемости рассматриваемых алгоритмов.

Реализация и внедрение результатов работы.

Результаты диссертации нашли применение в работах по развитию ЦКП ССКЦ СО РАН.

Результаты исследований были применены при выполнении Федеральной целевой программы «Исследования и разработки по приоритетным направлениям развития научно-технологического комплекса России на 2007 – 2013 гг.» (ГК № 07.514.11.4016 «Исследования и разработка методов имитационного моделирования функционирования гибридных экзафлопсных вычислительных систем»). Исследования проводились в рамках грантов Российского фонда фундаментальных исследований № 13-07-00589, 16-07-00434.

Внедрение результатов диссертационного исследования подтверждается соответствующими актами.

Научные положения, выносимые на защиту, соответствуют пунктам 2, 4, 6 паспорта специальности 05.13.15 – «Вычислительные машины, комплексы и компьютерные сети».

1. Создана мультиагентная модель оперативного управления для вычислительных систем высокой производительности, позволяющая исследовать алгоритмы управления. В данной модели могут быть исследованы алгоритмы статической и динамической балансировки нагрузки, составления очередей расписаний, обеспечения энергоэффективности и отказоустойчивости.
2. Экспериментально обоснован набор управляющих алгоритмов, который на реальной статистике ЦКП ССКЦ СО РАН показал лучшие параметры распределения ресурсов, возможности корректной обработки отказов и возможности обеспечения энергоэффективности.
3. Создан метод исследования эффективности параллельных алгоритмов и программ при их исполнении на суперкомпьютерах высокой производительности. Метод основывается на создании для каждого типа вычислительного алгоритма мультиагентной модели его исполнения на разном числе вычислительных узлов ВС.
4. С помощью модельных экспериментов показана хорошая согласованность получаемых данных с результатами действительного запуска рассматриваемых параллельных вычислительных алгоритмов и хорошая масштабируемость для некоторых из них. Для остальных алгоритмов предложены способы повышения масштабируемости при их исполнении на большом числе вычислительных ресурсов.

Степень достоверности научных положений достигается

1. Использованием при построении модели общепризнанных данных по алгоритмам и принципам функционирования высокопроизводительных вычислительных систем, имеющихся в отечественной и зарубежной литературе.
2. Объемом информации получаемой в модельном эксперименте, достаточной для корректной статистической обработки.

3. Хорошим качественным согласием результатов модельного эксперимента с соответствующими данными реальной статистики ЦКП ССКЦ СО РАН и данными реального исполнения исследуемых вычислительных алгоритмов.
4. Экспертизами работы, прошедшими при получении грантов.

Апробация

Основные положения диссертационной работы докладывались и обсуждались на всероссийских и международных конференциях:

– Международная конференция "Суперкомпьютерные дни в России" (RuSCDays 2015), 28-29 сентября 2015, Москва, Россия.

– 4th International Young Scientists Conference. (YSC -2015). June 25 - July 03 2015. Athens, Greece.

– 15th IEEE/ACM International Symposium Cluster, Cloud and Grid Computing (CCGrid). May 4-7 2015. Shenzhen, China.

– VII Международная научная конференция Параллельные вычислительные технологии (ПавТ), 1-5 апреля 2013, Челябинск, Россия.

– «Научный сервис в сети интернет» Международная суперкомпьютерная конференция и конференции молодых ученых. 17- 22 сентября 2012. Новороссийск (Абрау-Дюрсо), Россия.

– 14th IEEE International Conference on High Performance Computing and communications (HPCC-2012), June 25-27, 2012, Liverpool, UK.

– VI Международная конференция «Параллельные вычисления и задачи управления» (РАСО'2012). 24 - 26 октября 2012 г. Москва, Россия.

– Национальный Суперкомпьютерный Форум (НСКФ-2015). 24-27 ноября 2015 года. Переславль-Залесский, Россия.

– 8-я Всероссийская Мультиконференция по проблемам управления (МКПУ-2015). 28 сентября – 3 октября 2015. Геленджик, Россия.

– X Всероссийская школа-конференция молодых ученых "Управление большими системами" 5-7 июня 2013г., Уфа, Россия, работа награждена дипломом третьей степени.

Публикации. Автором по теме диссертации опубликовано 14 работ. Основные результаты диссертации полностью опубликованы в 6 статьях рецензируемых изданий, рекомендованных ВАК, 3 статьи в других рецензируемых журналах, 5 статей – в трудах конференций.

Структура и объем диссертации Диссертация состоит из введения, трех глав, заключения, списка литературы из 159 наименований. Общий объем диссертации – 137 страниц. Работа содержит 28 рисунков и 2 таблицы.

Автор выражает глубокую благодарность научному руководителю Б.М. Глинскому за помощь в проведении научных исследований по теме диссертационной работы.

Автор выражает также благодарность А.С. Родионову за консультации в области имитационного моделирования.

Автор выражает также благодарность коллективу лаборатории ССКЦ ИВМиМГ СО РАН за предоставленную возможность использовать статистику и ресурсы вычислительных кластеров.

Основное содержание работы

В первой главе рассматриваются вопросы функционирования высокопроизводительных ВС (ВПВС), рассматриваются основные архитектуры, режимы функционирования ВС, а также иерархия средств управления. Проводится анализ современных алгоритмов, подходов и средств управления на ВПВС. Рассматриваются

современные подходы к исследованию высокопроизводительных ВС, обоснован мультиагентный подход к моделированию функционирования ВПВС.

Вторая глава посвящена описанию разработанной мультиагентной модели оперативного управления высокопроизводительной вычислительной системой. Рассматриваемые вычислительные системы – распределенные, гетерогенные, мультикластерные, с архитектурой MPP. Система управления представлена несколькими наборами однотипных программных агентов, реализующие свои конкретные задачи и обменивающиеся между собой посредством специализированных сообщений. Все агенты системы разделены на несколько классов: агенты коммутаторы, агенты анализа, агенты коллекторы, агенты контроллеры и агенты вычислительных узлов. Общая схема модели разработанной системы представлена на рисунке 1.

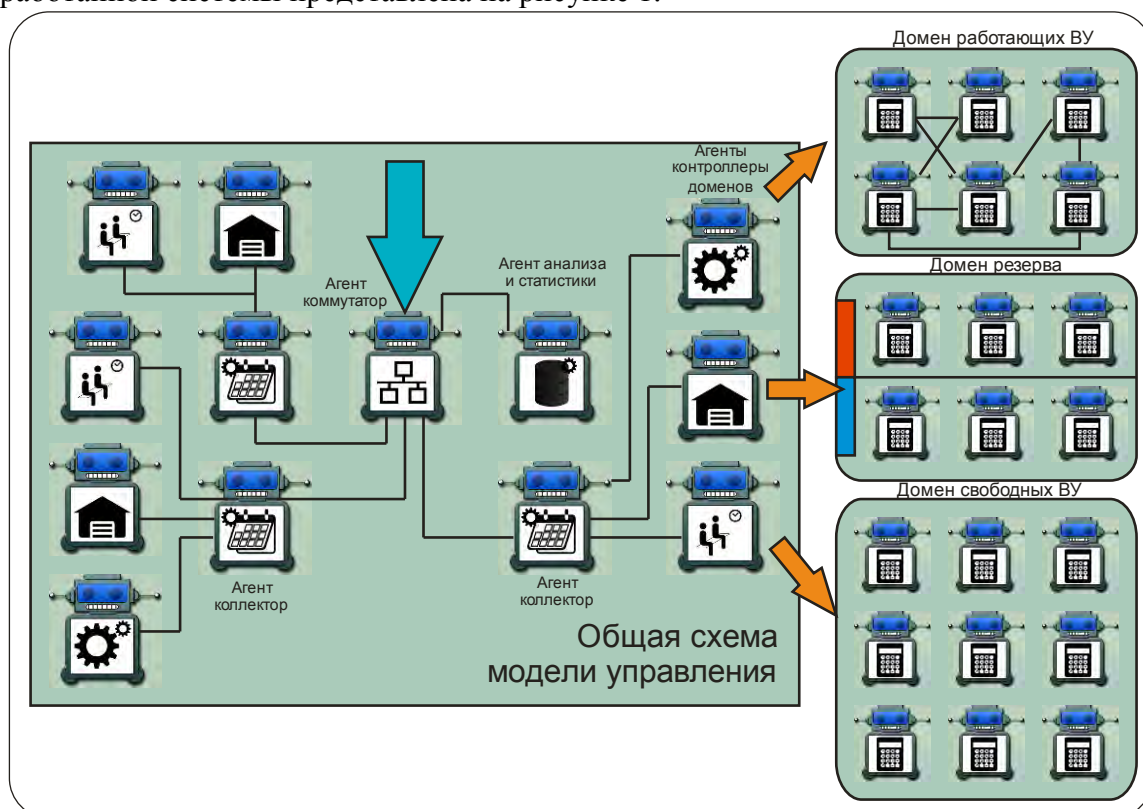


Рис. 1. Общая схема модели системы оперативного управления ВПВС, классы агентов и логические связи между ними.

Агенты сбора и анализа статистики собирают и хранят обобщенную информацию о заданиях уже решенных на ВПВС для каждого конкретного пользователя. По запросу от агентов коммутаторов предоставляет взвешенный список агентов коллекторов, которым предпочтительнее отправить поступившее параллельное задание пользователя.

Агенты коммутаторы ВПВС распределяют поток пришедших для решения заданий между входящими в его состав кластерами, управление которыми делегировано агентам коллекторам. При распределении учитывается их состояние на текущий момент времени и статистика считающего пользователя. Агенты коммутаторы также осуществляют поиск доступных вычислительных ресурсов в контролируемой ВС при запросах от подчиненных агентов коллекторов.

Агенты коллекторы составляют расписания выполнения поступившего на выполнение набора задач, организывают выполнение этого расписания, занимаются поиском свободных вычислительных ресурсов при отказах или необходимости балансировки.

Агенты контроллеры доменов вычислительных узлов (ВУ) организуют корректную работу ВУ, находящихся под их управлением. Происходит контроль как физических

параметров узлов, так и процесса выполнения параллельного задания, во избежание ситуаций отказа оборудования, заикливания программы и прочее.

Агенты вычислительных узлов по заданной программе (в зависимости от характеристик узла) имитирует действия реальных вычислительных узлов, выполняя свою ветвь параллельной программы или находясь в режиме ожидания.

В рассматриваемой системе введено понятие домена вычислительных узлов – набор агентов вычислительных узлов, находящийся под управлением агента контроллера и в каждый конкретный момент выполняющий одно и то же действие. Виды доменов ВУ:

– домен резерва ВУ содержит в себе вычислительные узлы, отправленные в «холодный» или «горячий» резерв, а также узлы, выведенные из эксплуатации для проведения профилактики или ремонта. Логические линии связи в данной области не строятся. Агент контроллер осуществляет своевременный перевод ВУ в соответствующий тип резерва;

– домен свободных ВУ содержит в себе все вычислительные узлы кластера, которые в данный момент времени не заняты решением пользовательской задачи или не находятся в резерве. Логические линии связи между ними полностью повторяют физические линии. Агент контроллер в таком случае просто отслеживает физические характеристики всех ВУ (температура, исправность линии связи) и ведет их учет;

– домен работающих ВУ содержит в себе вычислительные узлы, занятые в данный момент времени решением одной и той же пользовательской задачи. Логическая связь между ВУ организована по типу «каждый с каждым». Агент контроллер в данном случае наблюдает не только физические характеристики узлов, но и ход исполнения параллельной программы.

Исследование алгоритмов управления с помощью предложенной модели на данных статистики работы ЦКП ССКЦ СО РАН

Для исследования алгоритмов управления на статистике использования реальной системы в модели была воспроизведена коммуникационная среда кластера НКС-30Т+GPU (ЦКП ССКЦ СО РАН), который состоит из неоднородных узлов: G5; G6; G7; SL390, SMP.

Для исследования была собрана статистика работы ЦКП ССКЦ СО РАН за 2011 – 2015 гг. Данные были выбраны из лог-файлов системы управления PBS Pro, занесены в базу данных MySQL и, с ее помощью, обобщены и приведены к виду, пригодному для запуска процесса моделирования.

Для проведения модельного эксперимента по данной статистике использовались следующие алгоритмы управления:

1. Алгоритмы поиска очередей заданий (агентов коллекторов):
 - 1) экспертные оценки и основанные на знаниях системы [M.S. Fox 1984, D. Hildum 1994].
 - 2) семейство стратегий, состоящих из опорных планов, и прогнозирования освобождения ресурсов на основе моделирования действий локальных систем планирования [В.В. Топорков 2004].
2. Алгоритмы формирования расписания:
 - 1) простейшие алгоритмы, такие как FCFS, SJF, LJF [U. Schwiegelshohn 1998]
 - 2) Advance Reservation (AR) и Backfilling [S. Srinivasan 2002].
 - 3) генетические алгоритмы [С.И. Смагин 2010, А.В. Ефимов 2012].
3. Алгоритмы и методы, используемые при балансировке:
 - 1) динамическая система балансировки TriadBalance [А.И. Миков 2010].
 - 2) иерархический метод вложения параллельных программ в ВС [М.Г. Курносков 2007].
4. Методы и алгоритмы для обеспечения отказоустойчивости:
 - 1) динамическая миграция вычислительной нагрузки при достижении критического значения температуры [К. Skadron 2004].

- 2) модель функционирования распределенной вычислительной системы со структурной избыточностью [К.В. Павский, 2013]
5. Алгоритмы обеспечения энергоэффективности:
 - 1) методы аппаратного резервирования [Г.Н. Черкесов 2005, Л.Н. Александровская 2001].
 - 2) метод распределения потоков по выч. ядрам в зависимости от трендов физических параметров [А.К. Coskun 2007].
 - 3) динамическая миграция вычислительной нагрузки при достижении критического значения температуры [К. Skadron 2004].

Был проведен модельный эксперимент по эффективности алгоритмов распределения вычислительных ресурсов: формирования расписаний и поиска очередей заданий. Остальные алгоритмы управления, приведенные в списке выше, включены по умолчанию. Исследуемый параметр для данных алгоритмов – среднее время ожидания выполнения (разница между временем прихода задания в систему и фактическим временем начала ее исполнения). Исследовались следующие комбинации алгоритмов составления расписаний:

- простейший алгоритм FCFS, AR и Backfilling (без алгоритмов поиска очереди);
- генетический алгоритм, AR и Backfilling (без алгоритмов поиска очереди);
- генетический алгоритм, AR и Backfilling с использованием алгоритмов поиска очереди (агент коммутатор включен).

В зависимости от выбранного набора алгоритмов распределения ресурсов среднее время выполнения задания изменяется по разному (рис. 2). В случае использования простейшего алгоритма оно всегда существенно больше реальных показателей. С использованием генетического алгоритма изменение среднего времени ожидания сопоставимо с реальными данными. А использование алгоритмов распределения заданий по очередям в разные месяцы может сократить это время до 11%. Это связано с перераспределением заданий по различным очередям (рис. 3), что позволяет разгрузить более нагруженные очереди заданий (агенты коллекторы).

Для проверки работы алгоритмов обеспечения отказоустойчивости и обеспечения энергоэффективности был проведен модельный эксперимент над данными ЦКП ССКЦ СО РАН за 2014г. В модель вводилась вероятность отказа от 0 – 15%. Результаты по контролируемым параметрам – изменение среднего времени ожидания выполнения и средний объем резерва ВУ приведены на рисунке 4. Эксперимент показал, что повышение вероятности отказа до 15% не влияет на количество выполненных заданий, однако существенно снижает объем резерва и увеличивает время ожидания выполнения задания.

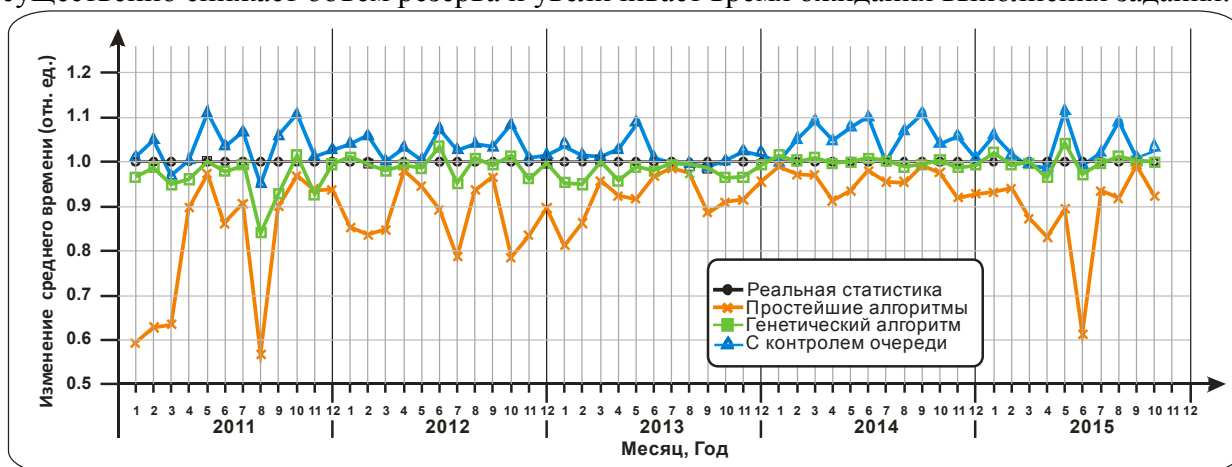


Рис. 2. Изменение среднего времени ожидания выполнения задания в зависимости от исследуемого набора алгоритмов распределения вычислительных ресурсов за 2011–2015г.

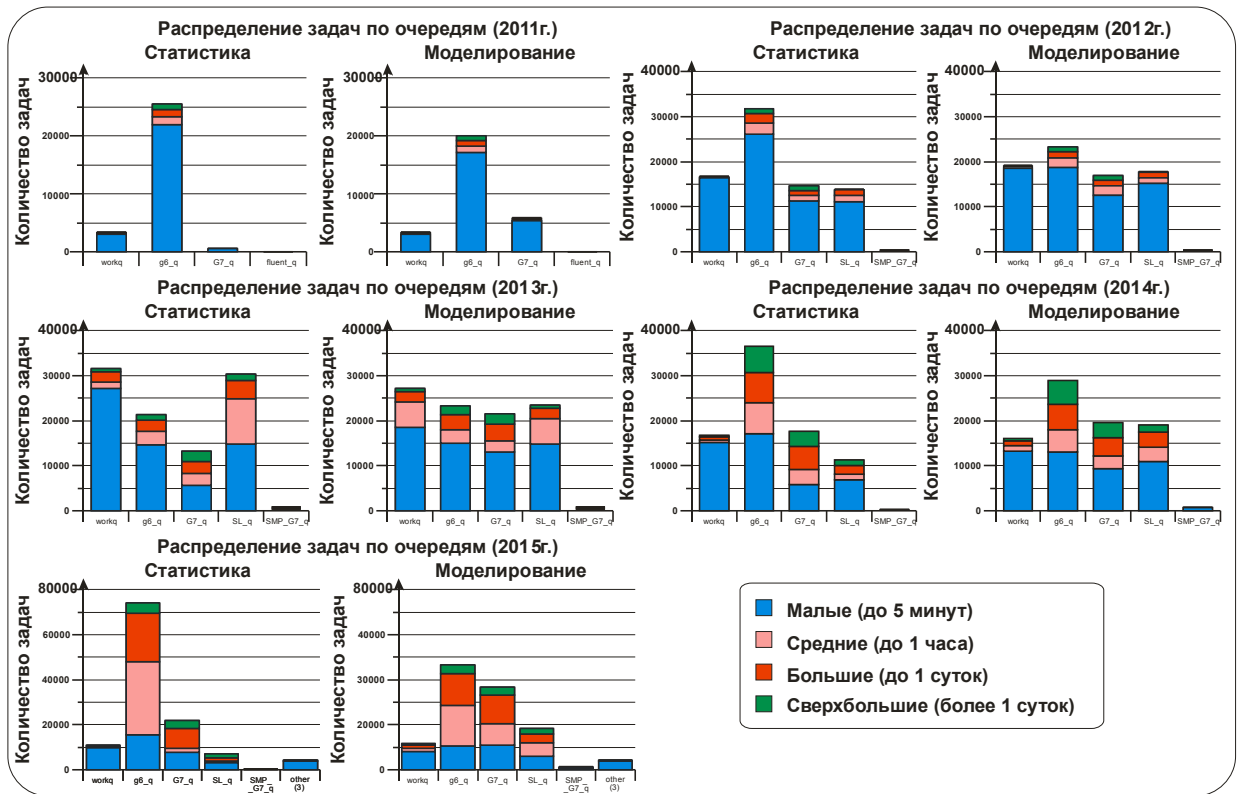


Рис. 3. Изменение распределения задач по очередям вследствие работы алгоритмов поиска (контроля) очереди за 2011 – 2015г.

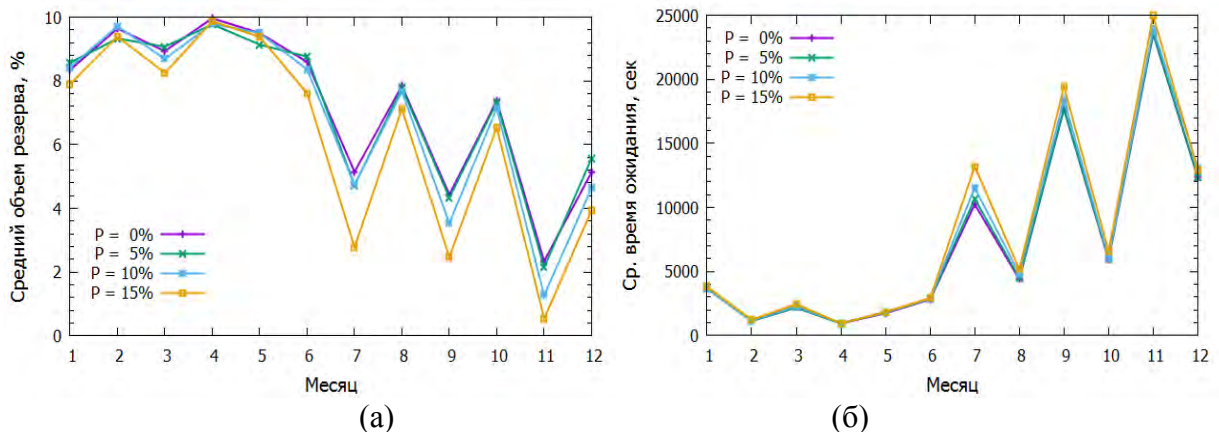


Рис. 4. Изменение среднего времени ожидания выполнения задания (а) и среднего объема резерва ВУ (б) в зависимости от вероятности отказа за 2014г.

В третьей главе представлен метод исследования масштабируемости параллельных вычислительных алгоритмов при их исполнении на ВПВС. Суть представленного метода заключается в создании мультиагентной модели процесса исполнения программы на различном числе вычислительных узлов. В общем виде вычислительный процесс представляет собой множество параллельно выполняемых на индивидуальном вычислительном узле нитей, которые взаимодействуют между собой посредством обмена сообщениями.

Для мультиагентного моделирования вычислительного процесса исследуется модель взаимодействия нитей и выделяются нити, имеющие схожее поведение – выполняющие одинаковые вычисления и обмены. Для каждой группы однотипных нитей создается класс программных агентов, который имитирует выполнение блоков расчета и обмен сообщениями с другими агентами (нитьями). Для каждого класса программных агентов создается алгоритм имитации поведения вычислительной нити и передаются основные параметры выполнения нити, такие как массив времен вычислений каждого

блока расчета, массивы числа передаваемых значений и времени передачи каждого значения для всех блоков коммуникаций. Для имитации выполнения всего вычислительного процесса, создается необходимое количество экземпляров каждого класса программных агентов. Когда все программные агенты созданы и зарегистрированы в системе AGNES, начинается процесс инициализации мультиагентной модели выполнения вычислительного процесса. Каждый программный агент начинает поиск своих «соседей» - агентов с которыми в процессе имитации выполнения нити он будет совершать коммуникации. Соседи легко отыскиваются в сервисе «желтых» страниц, по имени и порядковому номеру. В соответствии с алгоритмом выполнения нити и переданными параметрами агент настраивает временные счетчики, для имитации блоков расчетов и ожидания приема сообщений.

Исследуемая величина – изменение времени выполнения вычислительного процесса в зависимости от числа вычислительных ядер, на котором выполняется процесс:

$$S_L(M) = \frac{T_L(M_{\min})}{T_L(M)}$$

где M_{\min} - наименьшее число ядер, использованных при расчетах.

Исследование возможности масштабирования алгоритмов распределенного статистического моделирования.

По описанной схеме исследования исполнения вычислительных алгоритмов была исследована возможность масштабирования распределенного статистического моделирования на большое количество вычислительных ядер. Это задачи требующие моделирования экстремально большого количества независимых реализаций. К числу таких проблем относятся задачи моделирования с использованием прямого статистического моделирования (ПСМ) течений разреженного газа с учетом химических реакций, задачи переноса излучения и теории дисперсных систем. Реализация решения данной задачи основана на методе Монте-Карло.

Для имитации исполнительного алгоритма расчета по методу Монте-Карло создано два класса функциональных агентов – MCCollector и MCCalculator. MCCollector – сборщик, собирает и усредняет данные от вычислителей. MCCalculator – вычислитель, производит циклы расчета и отправляет данные сборщику. Схема взаимодействия приведена на рисунке 5(a).

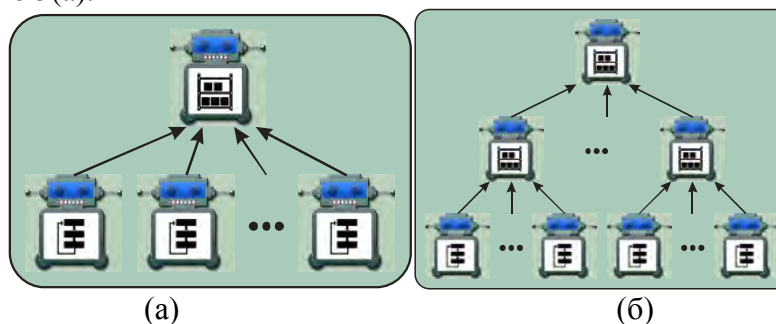


Рис. 5. Схема модели взаимодействия процессов. До модернизации (а) и после (б).

Проведенное имитационное моделирование с числом ядер от 48 до 10^5 (рис. 6(a)) показало, что при большом числе используемых вычислительных ядер (больше 10^4) реальное ускорение от распараллеливания существенно отличается от теоретического, что связано с большой загрузкой выделенных агентов-сборщиков, которые обрабатывают поступающие пакеты данных с агентов-вычислителей.

С целью повышения эффективности распараллеливания исследовались различные варианты организации обмена данными между агентами. А именно, целесообразно осуществлять периодическую пересылку результатов промежуточного осреднения реализаций, независимо полученных на агентах-вычислителях, на агенты-сборщики, объединенные в многоуровневую структуру. Агенты-сборщики будут периодически

получать переданные им данные и усреднять их, передавая затем результаты на агент, который соответствует вершине многоуровневой структуры (рис. 5(б)).

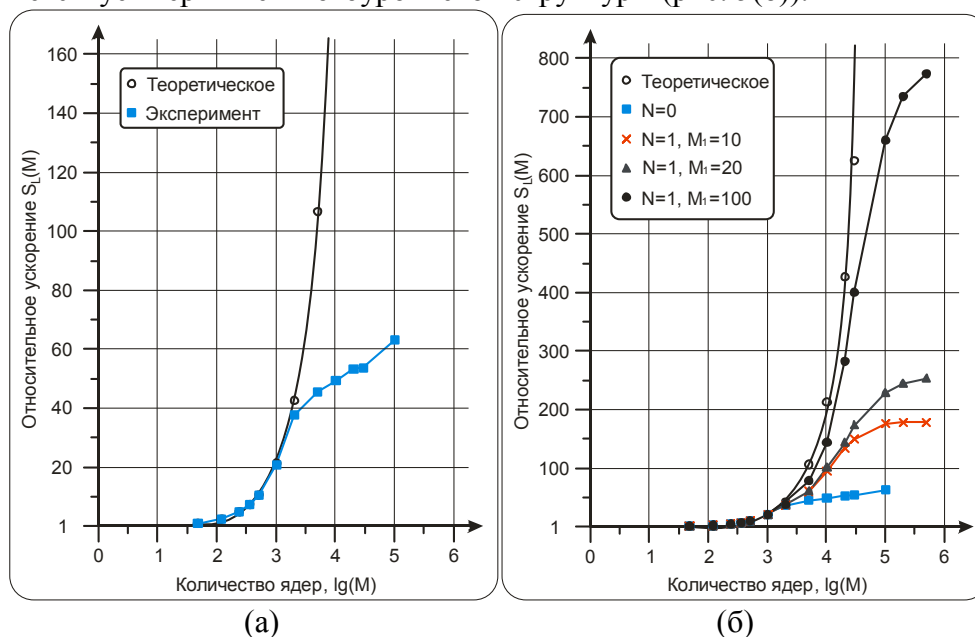


Рис. 6. Относительное ускорение алгоритма при увеличении числа ядер M . До модернизации (а) и после (б).

Результаты модельного эксперимента с многоуровневой иерархией (промежуточный уровень $N=1$) приведены на рисунке 6(б). Агенты-вычислители были поделены на M_1 равных частей ($M_1 = 10, 20, 100$), для каждой из которых данные с агентов-вычислителей всегда отправлялись на «свой» агент-сборщик первого уровня. На рисунке ясно видна закономерность: увеличение числа ядер-сборщиков в группах приводит к увеличению относительного ускорения.

В этой связи был исследован вопрос об оптимальном (в смысле максимального значения относительного ускорения) числе ядер-сборщиков при фиксированном общем числе моделируемых ядер ($M=10^6$). Экспериментально установлено, что число агентов-вычислителей для одного агента-сборщика при $N=1$ равно $M_1=1000$ штук.

Исследование возможности масштабирования алгоритмов численного моделирования. Сеточные методы.

Особенностью параллельных сеточных методов является возможность геометрической декомпозиции расчетной области и последующим обменом граничных значений между только соседними вычислительными узлами. При этом, если в подобласти содержится N^3 ячеек, то число передаваемых элементов соседнему вычислительному узлу составляет N^2 элементов. При выполнении вычислений используются однотипные нити-вычислители, совершающие циклы расчета в своей подобласти и обменов граничными значениями с «соседними» нитями. В каждой итерации цикла может происходить несколько блоков вычислений и обменов, в зависимости от решаемой задачи. Процесс обменов может происходить как в синхронном, так и в асинхронном режиме.

Критерием масштабируемости алгоритмов, основанных на сеточных методах, будем считать сохранение постоянным время расчета при увеличении расчетной области пропорционально увеличению количества вычислительных узлов. То есть, если на одном вычислительном узле в подобласти содержится N ячеек, то при увеличении количества вычислительных узлов (M) вся расчетная область уже содержит $M \times N$ ячеек.

Каждой нити вычислителю соответствует программный агент, имитирующий вычисления. По данной схеме проводились исследования нескольких задач численного моделирования.

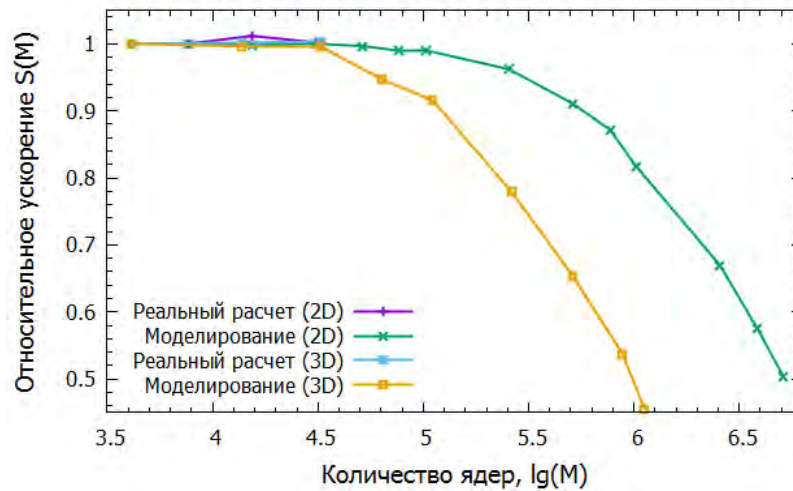


Рис. 7. Результаты исследования алгоритма моделирования 3D сейсмических полей для 2-х мерной и 3-х мерной декомпозиции.

Результаты исследования процесса численного моделирования 3D сейсмических полей в изотропной неоднородной упругой среде⁴ приведены на рисунке 7. Исследованы 2 вида декомпозиции области – 2D и 3D, соответственно у агентов может быть либо от 1 до 2, либо от 3 до 6 соседей. Коммуникации организованны в асинхронном режиме. Как видно из графиков, эффективность алгоритмов при увеличении числа ядер постепенно снижается, это объясняется увеличением числа обменов каждого узла с соседями на каждой итерации, для 3-х мерного варианта это число растет более стремительно. Эффективное использование этого алгоритма на суперкомпьютерах требует модификации его в сторону снижения количества производимых обменов.

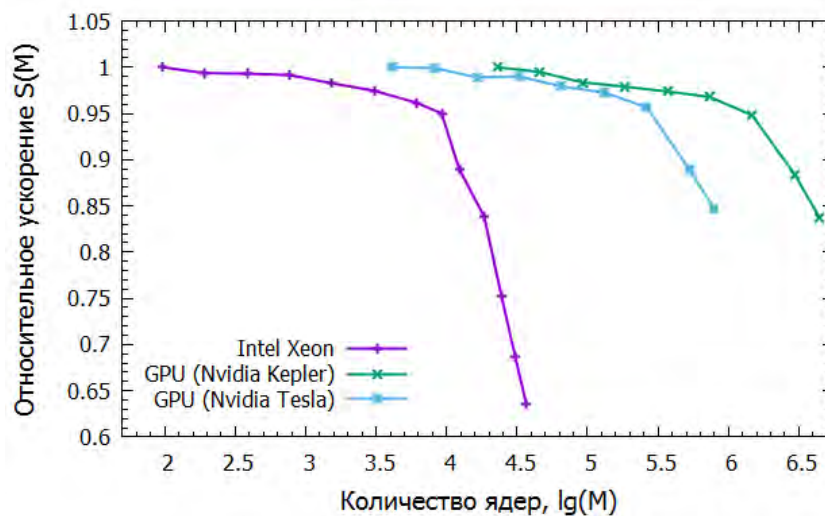


Рис. 8. Результаты исследования алгоритма моделирования взаимодействия электронного пучка с плазмой.

Результаты исследования процесса численного моделирования взаимодействия электронного пучка с плазмой⁵ приведены на рисунке 8. Исследовано решение данной задачи на различных типах вычислительных узлов. Коммуникации организованны в синхронном режиме. На этапе формирования полной матрицы текущей плотности тока и заряда необходим обмен значениями «все-со-всеми». Заметно, что эти обмены

⁴ Б.М.Глинский, Д.А.Караваев, В.В.Ковалевский, В.Н.Мартынов Численное моделирование и экспериментальные исследования грязевого вулкана «Гора Карабетова» вибросейсмическими методами. //Вычислительные методы и программирование. М.: Изд-во МГУ, 2010, Том 11, No1, С. 99-108

⁵ Вшивков В.А., Снытников А.В. Особенности проведения экзафлопс-расчетов в физике плазмы //Вычислительные методы и программирование: новые вычислительные технологии. 2012. Т. 13. № 1 (25). С. 44-48.

существенно снижают эффективность алгоритма. Решение данной задачи исследовано на различных типах вычислительных узлов. Как видно хорошую масштабируемость показывают вычисления на узлах с ускорителями Nvidia Kepler K40.

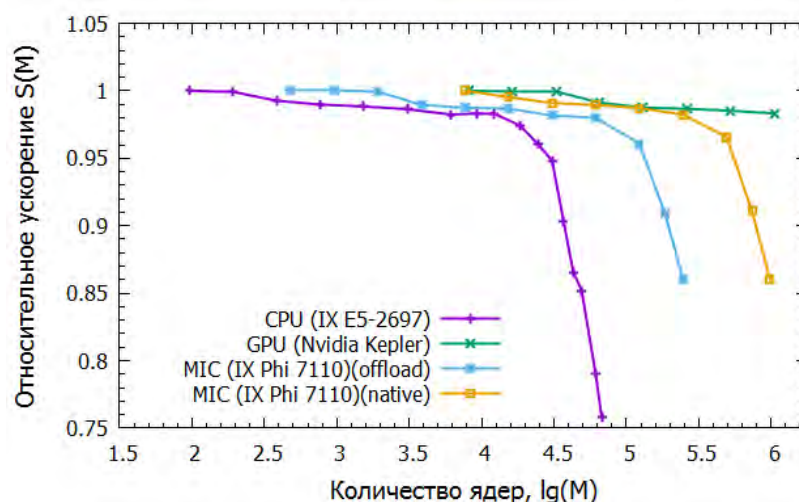


Рис. 9. Результаты исследования алгоритма моделирования задачи столкновения галактик.

Результаты исследования процесса численного моделирования задачи столкновения галактик⁶ приведены на рисунке 9. Исследовано решение данной задачи на различных типах вычислительных узлов. Коммуникации организованы в синхронном режиме. Каждая нить обменивается граничными значениями только с 2 соседями. Как видно из графиков алгоритм показывает хорошую масштабируемость. Лучшую масштабируемость показывают вычисления на узлах с ускорителями Nvidia Kepler K40 и Intel Xeon Phi.

Заключение

В диссертации реализованы методы исследования вычислительных и управляющих процессов в высокопроизводительных ВС с использованием мультиагентного имитационного моделирования, показана эффективность модели оперативного управления ВС для предварительного исследования наборов алгоритмов управления, реализованы модели исполнения вычислительных алгоритмов на ВС заданной архитектуры, показана хорошая масштабируемость для нескольких вычислительных алгоритмов и хорошее качественное согласие результатов модельного эксперимента и действительных запусков их на реальной системе.

Диссертация является завершенной научно-квалификационной работой, в которой содержится новое решение актуальной задачи, имеющая существенное значение для развития методов исследования

Дальнейшие исследования по данной тематике будут вестись в следующих направлениях: исследование и анализ алгоритмов, повышающих отказоустойчивость системы управления; анализ и развитие методов и алгоритмов повышающих энергоэффективность ВПВС; исследование процессов функционирования ВПВС с другими перспективными архитектурами.

Основные выводы:

1. Разработана имитационная мультиагентная модель оперативного управления высокопроизводительными вычислительными системами для исследования управляющих процессов, имеющая следующие отличительные характеристики:
 - 1) модель системы оперативно реагирует на изменения внешней среды и внутреннего состояния всех входящих в нее агентов.

⁶ Kulikov, I.: GPUPEGAS: A New GPU-accelerated Hydrodynamic Code for Numerical Simulations of Interacting Galaxies. The Astrophysical Journal Supplement Series. 214, 12, 1-12 (2014)

- 2) с помощью данной модели, можно воссоздать вычислительные ресурсы и коммуникационную среду распределенных гетерогенных мультикластерных вычислительных систем. В частности, автором воспроизведена коммуникационная среда кластера НКС-30Г+GPU (ЦКП ССКЦ СО РАН).
 - 3) в модель системы возможно как вложить различные наборы алгоритмов управления, так и заменить в наборе какой-то отдельный алгоритм, и в модельном эксперименте проверить их влияние на функционирование всей ВПВС.
 - 4) так как имитационное моделирование проводилось с помощью мультиагентного подхода, легко осуществим переход от программных агентов к программным модулям реальной системы помогающим в управлении или же полностью управляющим системой.
2. Модельными экспериментами функционирования ВПВС с применением различных алгоритмов управления по реальной статистике 2011 – 2015 гг. суперкомпьютера ЦКП ССКЦ СО РАН подобран набор управляющих алгоритмов, который позволяет снизить среднее время ожидания выполнения задания до 11% и обеспечивает отказоустойчивость и энергоэффективность.
 3. Разработан метод анализа эффективности вычислительных процессов при их исполнении на большом числе вычислительных узлов ВПВС, основанный на имитационном моделировании поведения параллельных вычислительных процессов, который позволяет провести имитацию вычислений исследуемого параллельного алгоритма на таком количестве узлов, который в данный момент недоступен его автору. На основании проведенной имитации выявить слабые и сильные стороны алгоритма и дать рекомендации как достигнуть большей эффективности на большом числе узлов, или же на каком максимальном числе узлов целесообразно использовать данный алгоритм.
 4. Исследована масштабируемость процесса вычисления задач прямого статистического моделирования (методом Монте-Карло):
 - 1) в изначальном виде масштабируемость данного параллельного алгоритма отличалась от теоретической. Это связано с тем, что при увеличении вычислительных узлов сильно возрастает нагрузка на узел сборки и усреднения результатов.
 - 2) предложена модификация алгоритма, предусматривающая многоуровневую иерархию узлов сборщиков, что позволяет повысить эффективность алгоритма.
 - 3) для 10^6 вычислительных узлов экспериментально обоснована оптимальная схема вычислений с 2 уровневой иерархией узлов сборщиков.
 5. Исследована масштабируемость процесса численного моделирования различных задач с помощью сеточных методов. Масштабируемость сеточных методов зависит от схемы декомпозиции расчетной области, количества, типа и объема коммуникаций между вычислительными узлами. Лучшую масштабируемость показывают вычисления на узлах с ускорителями Nvidia и Intel Phi.

Список публикаций по теме диссертации:

Статьи в изданиях, рекомендованных ВАК:

1. Б.М. Глинский, И.В. Куликов, А.В. Снытников, И.Г. Черных, **Д.В. Винс** Многоуровневый подход к разработке алгоритмического и программного обеспечения экзафлопсных суперЭВМ // Вычислительные методы и программирование. – 2015. – С. 600-613.
2. **Д.В. Винс**, Б.М. Глинский, А.С. Родионов. Исследование управляющих процессов в суперкомпьютерных системах на основе мультиагентного моделирования // Вестник СибГУТИ. - № 4 (28). - 2014. - С.35-44.

3. **Д.В. Винс.** Анализ эффективности системы управления потоком заданий для ЦКП в мультиагентной имитационной модели // Вестник НГУ. - т.12. - вып. 2. – 2014. - С. 33-41.
4. Б.М. Глинский, М.А. Марченко, Б.Г. Михайленко, А.С. Родионов, И.Г. Черных, Д.И. Подкорытов, Д.А. Караваев, **Д.В. Винс** Отображения параллельных алгоритмов для суперкомпьютеров экзафлопсной производительности на основе имитационного моделирования//Информационные технологии и вычислительные системы, №4 (2013), С. 3-14 .
5. Б. М. Глинский, А. С. Родионов, М. А. Марченко, Д. А. Караваев, Д. И. Подкорытов, **Д. В. Винс** Использование имитационного моделирования для настройки параметров масштабируемых алгоритмов при высокопроизводительных вычислениях // Вестник УГАТУ, Т. 17, № 5(58) (2013), С. 200–209.
6. Б.М. Глинский, А.С. Родионов, М.А. Марченко, Д.И. Подкорытов, **Д.В. Винс.** Агентно-ориентированный подход к имитационному моделированию суперЭВМ экзафлопсной производительности в приложении к распределенному статистическому моделированию // Вестник ЮУрГУ, 2012. № 18 (277), Вып. 12 С. 93 – 106.

Статьи и тезисы в прочих изданиях:

7. Kulikov, I., Chernykh, I., Glinsky, B., **Weins, D.**, Shmelev, A. Astrophysics simulation on RSC massively parallel architecture //Proc. 2015 IEEE/ACM 15th Int. Symposium on Cluster, Cloud, and Grid Computing, CCGrid 2015. IEEE Press, 2015.1131 – 1134. [SCOPUS]
8. Glinsky B.M., Kulikov I.M., Snytnikov A.V., Chernich I.G. **Weins D.V.** A Multilevel Approach To Algorithm And Software Design For Exaflops Supercomputers // Суперкомпьютерные дни в России: Труды международной конференции -М.: Изд-во МГУ, 2015.- с. 4-17. [SCOPUS]
9. Glinsky, B., Rodionov, A., Marchenko, M., Podkorytov, **D.**, **Weins, D.** Scaling the distributed stochastic simulation to exaflop supercomputers // Proceedings of the 14th IEEE International Conference on High Performance Computing and Communications, HPCC-2012 - 9th IEEE International Conference on Embedded Software and Systems, ICCESS-2012 Washington: IEEE Press, 2012. 1131–1136. [SCOPUS].
10. Глинский Б.М., Родионов А.С., Марченко М.А., Караваев Д.А., Подкорытов Д.И., **Винс Д.В.** Исследование масштабируемости параллельных алгоритмов методом агентно-ориентированного моделирования// Параллельные вычислительные технологии (ПаВТ'2013) Труды международной научной конференции. 2013. с. 318-329. [РИНЦ]
11. Kulikov I.M., Chernykh I.G., Vorobyov E., Snytnikov A.V., **Weins D.V.**, Shmelev A., Protasov V., Serenko A., Katysheva E., Vshivkov V., Rodionov A.S., Glinsky B.M., Tutukov A. High performance computing in astrophysics // Electronic Proc. of ISC'15 conference, Frankfurt, Germany 2015.
12. **Винс Д.В.** Использование мультиагентной модели центра коллективного пользования для принятия управленческих решений// «Управление большими системами» (УБС-10): материалы X Всероссийской школы-конференции молодых ученых. Том 1/ Уфимск. гос. авиац. тех. ун-т. – Уфа: УГАТУ, 2013. с. 151-155.
13. **Винс Д.В.** Мультиагентная модель системы оперативного управления распределением ресурсов для вычислительных центра коллективного пользования// Труды конференции молодых ученых. – Новосибирск: ИВМиМГ СО РАН, 2013. с. X-X.
14. **Винс Д.В.** Агентный подход к моделированию систем управления потоком заданий на вычислительные центры коллективного пользования // Труды конференции молодых ученых. – Новосибирск: ИВМиМГ СО РАН, 2011. с. X-X.